MWDSI
MIDWEST DECISION SCIENCES INSTITUTE

54th MIDWEST DECISION SCIENCES INSTITUTION ANNUAL CONFERENCE

PROCEEDINGS

BALL STATE UNIVERSITY

April 7th - 8th, 2023, Muncie Indiana

Program Chairs

Qiannong Gu
Ball State University

Xiaodong Deng
Oakland University

Sanjay Kumar
Valparaiso University

2023 Annual Conference Proceeding

Midwest Decision Sciences Institute

Muncie, Indiana. April 7th – 8th, 2023

Tracks

- Accounting and Finance
- Global Business Management
- Information Technology and e-Business
- Innovative Education
- Operations Management – Manufacturing
- Operations Management – Services
- Quantitative Methods and Statistics
- Student Papers
- Supply Chain & Marketing Management

TABLE OF CONTENTS

Accounting & Finance

51. IMPACT OF COGNITIVE BIASES IN FORECASTING ECONOMIC VARIABLES

Geoffrey Booth,

Michigan State University

Panayiotis Theodossiou

Ball State University

Using a generalized probabilistic framework, the impact of the cognitive biases of overconfidence, over-precision, and anchoring on experts' forecasts of important economic variables is investigated analytically. Examples of such economic variables include the growth rate of Gross Domestic Product growth rate, the return on a portfolio, and company earnings. The relationships between the cognitive biases and the moments of the perceived distribution of the attribute variables under consideration is investigated mathematically and several testable propositions are derived.

1. Introduction

Tversky and Kahneman (1974) state: "occasionally beliefs concerning uncertain events are expressed in numerical form as odds or subjective probabilities. … People rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to similar judgmental operations. In general, these heuristics are quite useful, but sometimes lead to severe and systematic errors."

Managerial decisions are often out of line with rational expectations, e.g., Shefrin (2001) and (2005) and Baker et al. (2006). This is generally due to cognitive biases, such as, over confidence and under confidence, and anchoring attributed to cognitive shortcuts and emotional errors, e.g., Statman (2017). Overconfident managers generally believe that their firms are undervalued, therefore, tend to hold executive options for longer periods, e.g., Palmon et al., 2008; Palmon and Venezia (2013) and (2015); Malmendier et al. (2011); and Malmendier and Tate (2015). Moreover, because of overestimation, they often invest in negative net present value projects, e.g., Heaton (2002).

According to Moore and Healy (2008) in Psychology literature the term overconfidence is defined in three distinct ways. The first definition relates to the overestimation of one's actual ability, performance, level of control, or probability of success. The second definition evolves around the beliefs of people to be better than others, and third one, the overprecision or excessive certainty regarding the accuracy of beliefs.

Overconfidence has been explained by hubris and miscalibration, e.g., Oberlechner and Osler (2012). Hubris occurs when persons overestimate their own success or the probability of favourable outcomes, e.g., Roll (1986) and Camerer and Lovallo (1999). Calibration measures the accuracy of probabilities and is a probabilistic tool in decision making, e.g., Lichtenstein et al. (1982). Often, the perceived probability distribution of overconfident people is characterized as too tight and that of underconfident people as too loose, e.g., Alpert and Raiffa (1982).

Moore and Healy (2008) describe three main characteristics of overconfident individuals. The first characteristic relates to the overestimation of actual performance due to illusion of control and planning fallacy. Such individuals tend to overestimate their ability of control and underestimate the time needed to complete a task, e.g., Langer (1975). The second characteristic is over-placement, which occurs when overconfident individuals believe that are better than others. An interesting research found that 93% of US drivers believe that are better than average drivers, e.g., Svenson (1981). The final characteristic is over precision. When individuals were asked to estimate the confidence intervals around their answers, the estimated intervals were found to be too narrow. Over the years, the overconfidence bias was explained using mainly conceptual probabilistic statements, such as, "overconfident traders perceive the distribution of portfolio returns to be too tight and underconfident traders too loose", e.g., Kyle and Wang (1997).

In other words, overconfident managers tend to overestimate the probabilities of favorable events and underestimate their range. This miscalibration of subjective probabilities and tails of the distribution may lead to an overestimation of managerial perrformance, expected portfolio returns, a project's expected cash flows and over evaluation of their firm's stock; e.g., Ben-David et al. (2013). In the management literature, overconfidence has been identified with the overestimation of the actual value and the precision of a business outcome. Underconfident managers, on the other hand, exhibit the opposite behavior. They underestimate the actual mean value of an outcome and attach a wider range of confidence intervals around it. Moreover, they tend to underestimate the probablility of favorable events relative to unfavorable events.

To build on existing findings and develop a deeper, broader understanding of these biases, there is a need for additional research to: (1) define overconfidence and under confidence using a probabilistic framework, (2) model and explore the properties of the biases and their impact on managerial perceptions about important decision-making measures such as, the mean, value-at-risk and other risk measures of economic variables and (3) express the biases graphically to improve understanding through visualization.

Section 2 develops a probabilistic framework that accounts for overconfidence and under confidence biases. Section 3 analyzes the impact of behavioral biases on managerial perceptions about the expected value, overall risk, downside risk, value-at-risk and expected shortfall of decision-making economic variables. Section 4 uses monte Carlo simulations to examine the impact behavioural biases in a portfolio and a capital budgeting application. The paper ends with summary and conclusions.

This section develops a probabilistic framework based on the skewed normal distribution (SN) to model and understand the impact of behavioural biases on managerial perceptions about the expected value, the variance, downside risk, value-at-risk and expected shortfall measures of important decision-making variables. Such variables include, among others, the return of a portfolio, a project's cash-flow stream and a company's earnings. The SN is a continuous three-parameter probability distribution. Each of its parameters can be used to describe or proxy managerial biases about future outcomes of economic variables

The use of the SN is motivated by its relative simplicity and the fact that is well-known to social scientists, therefore, its skewed extension is easier to grasp and relate to behavioral biases. It is worth noting that the main results in this paper could be derived albeit in a general fashion using abstract or flexible type skewed distributions, such as, the EGB2 of McDonald and Xu (1995) or the skewed generalized t of Theodossiou (1998) and the skewed generalized error distribution of Theodossiou (2015). In such cases, however, the derivation of the main results would be quite complex not provide any extra benefits to the issues investigated in this paper.

This paper investigates the impact of three cognitive biases related to the prediction accuracy, precision, and persistence on the BEA's three sequential monthly forecasts of the nominal and real annualized quarterly growth rate of the U.S. GDP and the associated implied price deflator (PD). This is accomplished by examining and comparing the distribution of forecast errors associated BEA's estimation approach to the forecast errors that would have occurred if the forecast met the conditions of rational expectations. To analyze the statistical properties of estimation errors, a new analytical framework based on Theodossiou's (1998, 2015) flexible four-parameter Skewed Generalized Error Distribution (SGED). The distribution's moments are derived and linked to each of the three cognitive biases. The nature of these three biases is identified and measured using the statistical concepts of mode and skewness. Consideration to the presence of rational biases is also given.

The analyses of the raw forecast error data, the SGED estimation results and their corresponding bootstrap simulation indicate that BEA forecasts, on average, under-predict nominal and real GDP and over-predict the GDP implied PD that connects the two GDP measures. The results also support the notion that there are instances in which forecasts are not as exact as would be expected if its error distribution was Gaussian. Moreover, there is some evidence supporting the conjecture that forecasters use past forecasts as a reference point when making new forecasts. Together these empirical results are consistent with the notion that over time that the BEA's monthly GDP forecasts do not meet the requirements to be considered rational and appear to exhibit cognitive biases.

There are three cognitive bias continuums that are especially relevant to the claim of irrationality: 1) over-predicting vs. under-predicting, 2) precision vs. imprecision, and 3) persistence vs. non-persistence. In the case of forecasting, these biased are linked to each other. To illustrate, the under-prediction of nominal and real GDP is consistent with the BEA forecasters being pessimistic since GDP growth is thought of as a positive outcome. The over-prediction of implied PD, however, also supports the notion of forecasters being pessimistic because of inflation typically being considered a negative outcome. Forecaster pessimism is also supported by the presence of error confidence bands that are wider than those that would have occurred if the error distribution was Gaussian. Wider bands signal imprecision and is consistent with the forecaster wanting to ensure that in hindsight the forecast does not give the impression of being completely wrong. This also suggests forecaster pessimism. Finally, there is mixed evidence supporting the presence of persistence, which is the tendency to attach a forecast to a reference point. Persistence bias is pessimistic for nominal GDP and implied PD, but it is only pessimistic for first preliminary forecast of real GDP.

2. Cognitive Biases and Calibrated Probabilities

This section employs a generalized probabilistic framework to investigate the impact of the cognitive biases of overconfidence, over-precision, and anchoring on the perceived distribution of important economic variables by managers, investors, traders, and other economic agents, referred, for the purpose of this paper as experts. Examples of such economic variables include the return on a portfolio, analysts earnings forecasts, managerial performance assessments, computation of a project's cash flows, etc. The probabilistic framework employed is that of the class of skewed generalized two-piece distributions analyzed in Theodossiou (2021). The relationships between the cognitive biases and the moments of the perceived distribution of the attribute variables under consideration is investigated mathematically and several testable propositions are derived. The analysis in this paper constitutes an important a generalization and extension of the results presented in Ellina, Mascarenhas, and Theodossiou (2020).

2.1 A General Two-Piece Forecasting Error Distribution

Let $y_{i,t} = x_{i,t} - x_t$, for $i = 1, 2, \ldots, q$, be the error associated with an expert's estimate of the value of an economic variable i periods prior to the realization of its true value at time t. The estimate is denoted by $x_{i,t}$ and the realization or actual value of the variable at time t by $x_{0,t}$ or simply x_t. The estimate is based on an expert's perception (calibration) of the probability distribution of x_t. Without loss of generality, the probability distribution of the error of each estimate is assumed to be the two-piece generalized probability function below:

$$f_y\left(y_{i,t}\right)=\frac{1}{\phi_i}f\left(\frac{y_{i,t}-m_i}{\phi_{1,i}}\right)I_{y_{i,t}\le m_i}+\frac{1}{\phi_i}f\left(\frac{y_{i,t}-m_i}{\phi_{2,i}}\right)I_{y_{i,t}>m_i}, \tag{1}$$

where f is a symmetric unimodal probability density function, m_i is the mode of $y_{i,t}$, $\phi_{1,i}$ and $\phi_{2,i}$ are respectively non-negative left and right to the mode tail parameters, $\phi_i = (\phi_{1,i} + \phi_{2,i}) / 2$, and I is an indicator function that takes the value of one when the specified condition is met and the value of zero otherwise. Eq. (1) can be viewed as a generalization of the double Gaussian distribution presented in Wallis (2014).

When $\phi_{1,i} < \phi_{2,i}$ the values of the estimation error $y_{i,t}$ are more dispersed to the right of its mode m_i. Thus, the distribution of $y_{i,t}$ is positively skewed with a mean greater than the mode, i.e., $E(y_{i,t}) > m_i$. When $\phi_{1,i} > \phi_{2,I}$, the values of $y_{i,t}$ are more dispersed to the left of its mode. The distribution is negatively skewed and $E(y_{i,t}) < m_i$. When $\phi_{1,i} = \phi_{2,i}$ the distribution of $y_{i,t}$ is symmetric and $E(y_{i,t}) = m_i$. It is noted that the distribution of the error f_y reflects the perceptions including the cognitive biases of the individual forming the estimate. The latter subjective distribution may differ significantly from that of the realized value of the underlying economic variable.

As shown in Theodossiou (2021), Eq. (1) can be expressed in terms of the following equivalent probability measure:

$$dF_y\left(y_{i,t}\right)=f_y dy_{i,t}=\left(1+sign\left(w_{i,t}\right)\lambda_i\right)f_w\left(w_{i,t}\right)dw_{i,t}, \tag{2}$$

where

$$w_{i,t}=\frac{y_{i,t}-m_i}{\left(1+sign\left(y_{i,t}-m_i\right)\lambda_i\right)\phi_i} \tag{3}$$

and

$$\lambda_i=\left(\phi_{2,i}-\phi_{1,i}\right)/\left(\phi_{1,i}+\phi_{2,i}\right). \tag{4}$$

It follows easily from Eq. (4) that the range of values for λ_i is the closed interval [−1, 1]. For negatively skewed distributions $\phi_{1,i} > \phi_{2,I}$ and $\lambda_i < 0$. For positively skewed distributions $\phi_{1,i} < \phi_{2,I}$ and $\lambda_i > 0$. Finally, for symmetric distributions $\phi_{1,i} = \phi_{2,I}$ and $\lambda_i = 0$.

The expected value and standard deviation of the error of each estimate are respectively:

$$\mu_i=E\left(y_{i,t}\right)=E\left(x_{i,t}-x_t\right)=m_i+\delta_i\sigma_i, \tag{5}$$

and

$$\sigma_i=\text{std}\left(y_{i,t}\right)=\left(1/\theta_i\right)\phi_i, \tag{6}$$

where

$$\delta_i=2\lambda_i G_{1,i}\theta_i,$$

$$\theta_i = \frac{1}{\sqrt{\left(1+3\lambda_i^2\right)^2 G_2 - 4\lambda_i^2 G_1^2}}$$

and

$$G_s = \int_{-\infty}^{\infty} \left|w_{i,t}\right|^s f_w dw_{i,t} \text{ , for } s = 1, 2,$$

are respectively the Pearson's mode of skewness, a scaling constant for the standard deviation, and the first and second absolute moments of $w_{i,t}$; see Theodossiou (2021) for the derivations. Note that the Pearson's mode of skewness δ_i has the same sign as λ_i. The above equations hold for any unimodal probability density function with at least two moments.

2.2 Cognitive Biases Propositions

Using the rational forecast concept as the benchmark, the impacts of overconfidence, precision, and persistence on the distribution of the estimates' errors are explored in this section. These potential impacts are presented as propositions, which, as Williamson (2002) points out can be positive or negative statements.

Proposition 1. *When sequential estimates of an economic variable incorporate efficiently all available information at the time of their construction, such estimates* (a) *will be unbiased predictors of subsequent estimates,* (b) *will be Martingale processes obeying the variance bounds* $var(x_{0,t}) \geq var(x_{1,t}) \geq var(x_{2,t}) \geq \ldots \geq var(x_{q,t})$, (c) *their estimation errors will have zero means and obey the variance bounds* $var(y_{1,t}) \leq var(y_{2,t}) \ldots \leq var(y_{q,t})$, *where* $y_{i,t} = x_{i,t} - x_{0,t}$, and (d) *subsequent revisions of these estimates will be orthogonal to each other as well as past estimates of the variable*.

A proof of Proposition 1 based on the properties of linear projection operators and nested information sets is provided in the Appendix. Classical statistical estimation methods focus exclusively on the data to estimate the parameters of forecasting model. Estimates generated by well-specified models are unbiased and efficient, and under certain regularity conditions, their estimated parameters are asymptotically normal. In contrast, the errors of models estimated using Bayesian statistical methods may not have a zero mean, minimum expected squared errors, or both because these methods incorporate informative priors and beliefs in the estimation of forecasting models. Their estimated parameters are obtained by maximizing the prior distribution of the parameters given the data, which may be based on the psychological biases of the experts constructing the estimates. (Morris (1974, 1977) and Van den Steen (2004, 2011) provide examples of how Bayesian processes incorporate the beliefs of experts in the estimation process.) Thus, the resulting parameters and forecasts may be biased and inefficient. Nevertheless, regardless of the estimation approach, the distribution of errors of the estimates can be captured by a flexible probability density function with the normal distribution being a special case.

For the propositions to follow is assumed that values of an economic variable larger than its mode are considered favourable outcomes. On the other hand, values smaller than the mode are considered non-favourable outcomes.

Proposition 2: *The mode, being the most likely value of a variable, represents the reference point or the anchor of an economic variable.*

The mode is the maximum point on the probability curve of a random variable. Thus, the probability mass assigned to the random variable at its mode is the largest. In other words, the mode represents the most likely of all values of a random variable. Naturally, the mode represents the reference point or the anchor around which an individual calibrates probabilities. The mode of a forecast is equal to the mode of the realized value plus the mode of the forecast error. That is,

$$\text{mode}(x_{i,t}) = \text{mode}(x_t) + \text{mode}(y_{i,t}).$$

For an individual with perfect foresight, the perceived mode is expected to be equal to the realized value of an economic variable, thus the mode of the forecast error should be zero. Individuals with anchors higher than the realized values of the variable tend to overestimate. In sequential estimation, as new information becomes available, anchors may be updated accordingly. Under the assumption rationality of forecasts, biases are expected to be lower.

Proposition 3: *The precision bias of overconfident individuals is characterized by the tails of the distributions of the estimation errors being thinner relative to the actual tails. For overconfident individuals, the tail parameter of the subjective probability distribution will be smaller.*

The proof of this proposition, presented in the Appendix, relies on the quantile ranges associated with the estimates of overconfident individuals relative to those of rational expectations individuals. If the quantile range of the forecast is smaller than would occur if the forecast was rational, the forecast is more precise than the rational forecast.

Proposition 4: *Overconfident* (*underconfident*) *individuals assign higher* (*lower*) *probabilities to favorable than to unfavorable outcomes. As a result, their subjective probability distributions are positively* (*negatively*) *skewed.* This translates into a positive (negative) value for the asymmetry parameter λ.

An implicit assumption of this proposition is that larger values than the mode of a random variable are preferable to smaller values. Overconfident individuals tend to assigned a larger probability mass to values larger than the mode as compared to those smaller than the mode. As a result, their perceived probability distribution will be right- or positively skewed. The opposite is the case for underconfident individuals.

Proposition 5: *The total estimation bias consists of the anchoring bias and the skewness bias. That is,* $E\left(y_{i,t}\right) = E\left(x_{i,t} - x_{0,t}\right) = m_i + \delta_i \sigma_i$, for i = 1, 2, ..., q.

According to Eq. (5) the expected value of the error of an estimate is equal to the mode m_i of the estimation error and the skewness bias $SB_i = \delta_i \sigma_i$. The equation provides the basis for separating the anchoring bias from the skewness such that mi (mode) represents the anchoring bias and δi σi the skewness bias. Skewness bias is directly linked to the asymmetry parameter λi. Note that the sign of the skewness bias is driven by the sign of the asymmetry parameter λ_ι. For overconfident individuals the skewness bias is expected to be positive and for underconfident individuals, negative.

3. Specification and Estimation of Distributional Parameters

Let x_t for t = 1, 2, …, T be a series of actual values of the economic variable and $x_{i,t}$, for i = 1, 2, …, q a sequence of estimates i periods prior to the realization of the actual value of x_t. The computed error for each estimate is $y_{i,t} = x_{i,t} - x_t$, for i = 1, 2, …, q and t = 1, 2, …, T.

3.1 Distribution of Errors

The errors of the q sequential estimates are modelled using the flexible skewed generalized t (SGT) distribution of Theodossiou (1988) below:

$$f\left(y_{i,t}\right)=\frac{k\,n^{-\frac{1}{k}}}{2\phi_i}B\left(\frac{1}{k},\frac{n}{k}\right)^{-1}\left(1+\frac{1}{n}\left|\frac{y_{i,t}-m_i}{\left(1+sign\left(y_{i,t}-m_i\right)\lambda_i\right)\phi_i}\right|^k\right)^{-\frac{n+1}{k}}, \qquad (12)$$

where m_i, ϕ_i, and λ_i are as defined previously, k and n are distributional shape parameters, *sign* is the sign function taking the value of minus one for negative values of $y_{i,t} - m_i$ and the value of one for positive values, and $B\left(a,b\right)=\int_0^1 x^a\left(1-x\right)^{b-1}dx$ is the beta function.

The shape parameter k controls mainly the peakness of the distribution around the mode. Smaller values result in leptokurtic distributions and larger values to platykurtic distributions. The parameter n controls mainly the tails of the distribution. Empirical evidence show that the SGT, as a flexible type of distribution, provides an excellent fit to the empirical distribution of financial and economic data. The SGT nests several other distributions as special cases. It yields, for $n \to\infty$ the skewed generalized error distribution (SGED) of Theodossiou (2015), for $k_i = 2$ the skewed student's t of Hansen (1994), for $k = 2$ and $n = 1$ the skewed Cauchy, for $n\to\infty$ and $k = 1$ the skewed Laplace distribution, and for $n\to\infty$ and $k = 2$ the skewed normal distribution.

3.2 Parameters of the Distribution, Specification

The mode, tail, and asymmetry parameters of $y_{i,t}$, are specified as:

$$\phi_i=\text{tail}\left(y_{i,t}\middle|EI_{i,t}\right)=\exp\left(\sum_{i=1}^{q}a_iEI_{i,t}\right), \qquad (12a)$$

$$m_i=\text{mode}\left(y_{i,t}\middle|EI_{i,t}\right)=\sum_{i=1}^{q}b_iEI_{i,t}\,, \qquad (13a)$$

and

$$\lambda_i=asym\left(y_{i,t}\middle|EI_{i,t}\right)=1-\frac{2}{1+\exp\left(\sum_{i=1}^{q}c_iEI_{i,t}\right)},$$

(14a)

where $EI_{i,t}$ is an indicator variable for the i^{th} sequential estimate $x_{i,t}$ of the actual value x_t or $x_{0,t}$ for i = 1, 2, …, q, and t = 1, 2, …, T. That is, $EI_{i,t}$ takes the value of one for the i^{th} estimate and zero otherwise. For each estimate i the above parameters simplify to

$$\phi_i = \exp\left(a_i\right) \tag{12b}$$

$$m_i = b_i, \tag{13b}$$

and

$$\lambda_i = 1 - \frac{2}{1+\exp\left(c_i\right)}, \tag{14b}$$

The expected value and standard deviation, if $n > 2$, given by Eqs. (5) to (9) are respectively.

$$\mu_i = E\left(y_{i,t} \middle| EI_{i,t}\right) = m_i + \delta_i\ \sigma_i,$$

and

$$\sigma_i = std\left(y_{i,t} \middle| EI_{i,t}\right) = \phi_i / \theta_i,$$

where

$$\delta_i = 2\lambda_i G_{1,i}\theta_i,$$

$$\theta_i = \frac{1}{\sqrt{\left(1+3\lambda_i^2\right)G_2 - 4\lambda_i^2\, G_1^2}},$$

and

$$G_s = n^{\frac{s}{k}} B\left(\frac{s+1}{k}, \frac{n-s}{k}\right) B\left(\frac{1}{k}, \frac{n}{k}\right)^{-1} \text{ for } s = 1, 2, \ldots < n.$$

For the derivation of Eq. (9) see Theodossiou (2021). Note that in Eq. (6), $\delta_{i,i}\ \sigma_{i,j}$ represents the skewness term of the expected value of $y_{i,j,t}$.

3.3 Estimation

The MLE estimates are obtained from the maximization of

$$\max_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) = \sum_{i=1}^{q}\sum_{t=1}^{T} \log f\left(\boldsymbol{\beta} \middle| y_{i,t}\right) = \sum_{i=1}^{q}\sum_{t=1}^{T} L_{i,t}(\boldsymbol{\beta}), \text{ for } i = 1, 2, 3, \qquad (15)$$

where $f\left(\boldsymbol{\beta} \middle| y_{i,t}\right)$ is the likelihood function of the forecasting error of the estimate constructed *i* periods prior to the realization of the actual value of x_t, $\boldsymbol{\beta} = \left[\boldsymbol{\beta}_1;\boldsymbol{\beta}_2;\ldots;\boldsymbol{\beta}_q;k,n\right]$ and $\boldsymbol{\beta}_i = \left[\mu_i;\sigma_i;\lambda_i\right]'$ includes the estimates for the mean, standard deviation, asymmetry parameter of the *q* estimates along with the shape parameters *k* and *n* of the distribution.

The sample likelihood function is maximized using the Berndt, et al. (1974) algorithm. Robust standard errors for the MLE estimators $\tilde{\boldsymbol{\beta}}_i$ are computed using:

$$\operatorname{var}\left(\tilde{\boldsymbol{\beta}}\right) = -\left(\sum_{i=1}^{q}\sum_{t=1}^{T}\frac{\partial L_{i,t}\left(\tilde{\boldsymbol{\beta}}\right)}{\partial \boldsymbol{\beta}}\frac{\partial L_{i,t}\left(\tilde{\boldsymbol{\beta}}\right)}{\partial \boldsymbol{\beta}'}\right)^{-1}\sum_{i=1}^{q}\sum_{t=1}^{T}\frac{\partial L_{i,t}^{2}\left(\tilde{\boldsymbol{\beta}}\right)}{\partial \boldsymbol{\beta}\,\partial \boldsymbol{\beta}'}\left(\sum_{i=1}^{q}\sum_{t=1}^{T}\frac{\partial L_{i,t}\left(\tilde{\boldsymbol{\beta}}\right)}{\partial \boldsymbol{\beta}}\frac{\partial L_{i,t}\left(\tilde{\boldsymbol{\beta}}\right)}{\partial \boldsymbol{\beta}'}\right)^{-1}. \qquad (16)$$

3.4 Simulations

To understand how the parameters of the perceived probability distribution impact the mean, the standard deviation, and the asymmetry parameter of the estimates of economic variables we contact the following simulation exercise.

Assume that the distribution of an economic variable is normal with mean or mode ten and standard deviation one, i.e., $\mu = 10$ and $\sigma = 1$. Note that for the normal distribution $\mu = m$, and $\sigma = \phi$, $k = 2$, $n = \infty$. Furthermore, consider an overconfident expert with the following initial anchors of the mode and left and right tails of the distribution of the variable *i* periods prior to the realization of its actual value (let $i = 5$):

$$m_5 = 14,\ \phi_{1,5} = 0.4 \text{ and } \phi_{2,5} = 1.1.$$

Furthermore, assume that each period, as new information arrives, the expert updates his/her anchors, using the following adaptive equations:

$$m_{i-1} = 3/4\,\phi + 1/4\,m_i\,,$$

$$\phi_{2,i-1} = 1/2\,\phi + 1/2\,\phi_{2,i}\,,$$

$$\phi_{2,i-1} = 1/2\,\phi + 1/2\,\phi_{2,i}$$

Columns 2, 3, and 4 of Table 1 present the updated values of the parameters starting from five to one period prior to the realization of the true value of the variable. The last columns present respectively the average value of the tail parameter and the implied asymmetry parameter of the perceived distribution of the economic variable. Notice that in the above example the anchors converge to the true parameters of the distribution as we approach the time of realization of the true value of the variable.

The perceived distribution of the above agent in each period is simulated using 800 randomly generated observations. Let $x_{i,t}$ denote the randomly generated values at time t and i period prior to the realization of the actual value of the economic variable. In our case i = 1, 2, …, 5 and t = 1, 2, …, 800, i.e., each sample includes 4,000 (= $q \cdot T = 5 \cdot 800$) randomly generated values of $x_{i,t}$. The perceived standard deviations, means, and asymmetry parameters are estimated using maximum likelihood and the parameter specifications presented in Section 3.3 and 3.4. This process is repeated 500 times. Table 2 presents the arithmetic averages and standard deviations, means, and asymmetry parameters in the five hundred randomly generated samples. Notice that as we move closer to the point of release, the standard deviation converges to one, the mean to ten, and the asymmetry parameter to zero.

Table 1. Distributional Parameters of and Overconfident Agent with a Positive Anchoring Bias and Adaptive Expectations

Periods	m_i	$\phi_{1,i}$	$\phi_{2,i}$	$\phi_i = \frac{\phi_{1,i} + \phi_{2,i}}{2}$	$\lambda_i = \frac{\phi_{2,i} - \phi_{1,i}}{\phi_{1,i} + \phi_{2,i}}$
5	14.0000	0.4000	1.1000	0.7500	0.4667
4	11.0000	0.7000	1.0500	0.8750	0.2000
3	10.2500	0.8500	1.0250	0.9375	0.0933
2	10.0625	0.9250	1.0125	0.9688	0.0452
1	10.0156	0.9625	1.0063	0.9844	0.0222

The distribution of the economic variable is normal with mean ten and standard deviation one (μ = 10 σ = 1). For the normal distribution $\phi = \sigma$ and m = μ. The anchoring bias of the overconfident agent, five periods prior to the realization (q = 5) of the true value is assumed to be fourteen (m_5 = 14 > m = 10) and the left- and right-to-the-mode tail parameter $\phi_{1,5} = 0.4$ and $\phi_{2,5} = 1.1$. As new information arrives, the agents updates his/her the anchors using the adaptive $m_{i-1} = 3/4\phi + 1/4 m_i$, $\phi_{2,i-1} = 1/2\phi + 1/2\phi_{2,i}$, and $\phi_{2,i-1} = 1/2\phi + 1/2\phi_{2,i}$, for i = 2, 3, 4, 5. Columns 2, 3, and 4 of the table present the updated values of the parameters starting from five to one periods prior to the realization of the true value of the variable. The last columns present respectively the average value of the tail parameter and the implied asymmetry parameter of the perceived distribution of the economic variable. Notice that in the above example the anchors converge to the true parameters of the distribution as we approach the time of realization of the true value of the variable.

Table 2. Averages of Standard Deviations, Means, and Asymmetry Parameters Based on Five-Hundred random Samples Generated Using the Parameters in Table 1

σ_5	0.7845	μ_5	14.5573	λ_5	0.4645
	(0.0203)		(0.0270)		(0.0477)
σ_4	0.8826	μ_4	11.2803	λ_4	0.2058
	(0.0230)		(0.0310)		(0.0505)
σ_3	0.9397	μ_3	10.3901	λ_3	0.0934
	(0.0237)		(0.0350)		(0.0513)
σ_2	0.9688	μ_2	10.1295	λ_2	0.0438
	(0.0242)		(0.0331)		(0.0531)
σ_1	0.9829	μ_1	10.0536	λ_1	0.0272
	(0.0253)		(0.0336)		(0.0553)
k	2.0114				
	(0.0704)				

The perceived distribution of the agent of Table 1 in each period is simulated is simulated using 800 observations. The perceived standard deviations, means, and asymmetry parameters are estimated using maximum likelihood and the parameter specifications presented in Section 3.3 and 3.4. This process is repeated 500 times. The table above presents the arithmetic averages and standard deviations, means, and asymmetry parameters in the five-hundred randomly generated samples. Notice that as we move closer to the point of release, the standard deviation converges to one, the mean to ten, and the asymmetry parameter to zero.

APPENDIX – PROOFS FOR PROPOSITIONS 1, 2, AND 4

PROPOSITION 1

The proof is based on the properties of linear projection operators and nested information sets and is a variation of the proof presented in Neftci and Theodossiou (1991). Let $x_{q,t}$, …, $x_{2,t}$, $x_{1,t}$ be a sequence of estimates released at different points in time prior to the realization of the actual value of an economic variable at time t, denoted by $x_{0,t}$ or simply x_t. Each estimate is expressed as $x_{i,t} = x_{i+1,t} + \varepsilon_{i+1,t}$, where x_{i+1} represents the estimate formed one period earlier, $\varepsilon_{i+1,t}$ the revision, and $i = 0, 1, 2, …, q-1$. Recursive substitution of the latter equation into $x_{0,t}$ yields

$$x_{0,t} = x_{1,t} + \varepsilon_{1,t} = x_{2,t} + \varepsilon_{1,t} + \varepsilon_{2,t} = \dots = x_{q,t} + \varepsilon_{1,t} + \varepsilon_{2,t} + \dots + \varepsilon_{q,t}$$

Under the assumption that successive estimates of x_t incorporate all relevant information available at the time of their construction, the sequence of information sets on which these estimates are based, denoted respectively by $\Phi_{1,t}$, $\Phi_{2,t}$, …, $\Phi_{q,t}$, will be nested, thus, $\Phi_{0,t} \supset \Phi_{1,t} \supset \Phi_{2,t} \supset \dots \supset \Phi_{q,t}$. Hence, $\Phi_{i-1,t} = \Phi_{i,t} \oplus \Phi^c_{i,t}$, where $\Phi^c_{i,t}$ is the new information on which the revision $\varepsilon_{i,t}$ is based. Note that $\Phi^c_{i,t}$ is orthogonal to $\Phi_{i,t}$ and previous information sets, i.e., $\Phi^c_{i,t} \perp \Phi_{s,t}$ for $i \le s$ and $i, s = 1, 2, …, q$.

Let $\mathbf{P}\left(\cdot \middle| \Phi_{i,t}\right)$ be a linear projection operator onto space $\Phi_{i,t}$. Because of the orthogonality of $\Phi^c_{i,t}$ and $\Phi_{s,t}$, $\mathbf{P}\left(\varepsilon_{i,t} \middle| \Phi_{s,t}\right) = 0$, and $\mathbf{P}\left(x_{i,t} \middle| \Phi_{s,t}\right) = x_s$, for $i, s = 1, 2,…, q$. Thus, future revisions are not predictable from available information. This proves part (a) of the proposition.

Because of the orthogonality of $\Phi^c_{i,t}$ with $\Phi^c_{s,t}$ and Φ_s, for $i < s$ and $i, s = 1, 2, …, q$, the sequence of revisions $\left\{\varepsilon_{i,t}, i = 1, 2, \dots, q\right\}$ are orthogonal to each other and earlier estimates of x_t.

These results in conjunction with the unbiasedness property of estimates implies that the series of preliminary estimates is a Martingale process with respect to $i = 1, 2, …, q$; part (b) of the proposition. Because

$$x_{0,t} = x_{1,t} + \varepsilon_{1,t} = x_{2,t} + \varepsilon_{1,t} + \varepsilon_{2,t} = \dots = x_{q,t} + \varepsilon_{1,t} + \varepsilon_{2,t} + \dots + \varepsilon_{q,t},$$

the orthogonality of $\varepsilon_{i,t}$'s implies that

$$\operatorname{var}\left(x_{0,t}\right) \ge \operatorname{var}\left(x_{1,t}\right) \ge \operatorname{var}\left(x_{2,t}\right) \dots \ge \operatorname{var}\left(x_{q,t}\right).$$

The recursive equation above implies that the forecast errors of successive estimates of x_t, expressed as functions of the revisions are:

$$y_{1,t} = x_{1,t} - x_{0,t} = -\varepsilon_{1,t}$$

$$y_{2,t} = x_{2,t} - x_{0,t} = -\varepsilon_{1,t} - \varepsilon_{2,t}$$

$$\ldots$$

$$y_{q,t} = x_{q,t} - x_{0,t} = -\varepsilon_{1,t} - \varepsilon_{2,t} - \ldots - \varepsilon_{q,t},$$

thus, due to the orthogonality of $\varepsilon_{i,t}$'s

$$\mathrm{var}\left(y_{1,t}\right) \leq \mathrm{var}\left(y_{2,t}\right) \ldots \leq \mathrm{var}\left(y_{q,t}\right) .$$

PROPOSITION 2

Consider a pair of values positioned symmetrically around their mode $y_1 = m - b$ and $y_2 = m + b$, where $b > 0$ (for simplicity of exposition the subscripts i and t are dropped from the notation). Assume that larger values of y are preferable to smaller values. Their transformed values in terms of w are respectively:

$$w_1 = \frac{y_1 - m}{\left(1 + sign\left(y_1 - m\right)\lambda\right)\phi} = -\frac{b}{\left(1-\lambda\right)\phi}$$

and

$$w_2 = \frac{y_2 - m}{\left(1 + sign\left(y_2 - m\right)\lambda\right)\phi} = \frac{b}{\left(1+\lambda\right)\phi} .$$

For $\lambda > 0$,

$$(1 - \lambda) < (1 + \lambda),$$

$$\left|w_1\right| = \frac{b}{\left(1-\lambda\right)\phi} > \frac{b}{\left(1+\lambda\right)\phi} = w_2 ,$$

$$f_w\left(w_1\right) = f_w\left(\left|w_1\right|\right) < f_w\left(w_2\right),$$

$$\left(1-\lambda\right) f_w\left(w_1\right) dw < \left(1+\lambda\right) f_w\left(w_2\right) dw ,$$

and

$$dF_y\left(y_1\right) < dF_y\left(y_2\right).$$

Thus, the subjective probability distribution of overconfident individuals is positively skewed. The equality $f_w(w_1) = f_w(|w_1|)$ is a by-product of the assumption that f_w is symmetric.

On the other hand, for $\lambda < 0$,

$$(1 - \lambda) > (1 + \lambda),$$

$$|w_1| = \frac{b}{(1-\lambda)\phi} < \frac{b}{(1+\lambda)\phi} = w_2,$$

$$f_w(w_1) = f_w(|w_1|) > f_w(w_2),$$

$$(1-\lambda) f_w(w_1) dw > (1+\lambda) f_w(w_2) dw,$$

and

$$dF_y(y_1) > dF_y(y_2).$$

Thus, the subjective probability distribution of underconfident individuals is negatively skewed.

PROPOSITION 4

The proof employs the properties of quantile values.

For $a < (1 - \lambda) / 2$, the

$$F(y_a) = \int_{-\infty}^{y_a} dF_y = (1-\lambda) \int_{-\infty}^{w_a} dF_w = a$$

and

$$F_w(w_a) = \int_{-\infty}^{w_a} dF_w = \frac{a}{1-\lambda}.$$

The respective left-quantile values of w and y are

$$w_a = F_w^{-1}\left(\frac{a}{1-\lambda}\right) < 0$$

and

$$y_a = m + (1-\lambda) w_a = m - (1-\lambda)|w_a|.$$

For the right-quantile values (corresponding to probability 1 – *a*),

$$F_y\left(y_{1-a}\right)=\int_{-\infty}^{y_{1-a}} dF_y=\int_{-\infty}^{w_{1-a}}\left(1+sign\left(w\right)\lambda\right)dF_w$$

$$=\frac{1}{2}\left(1-\lambda\right)+\left(1+\lambda\right)\int_{-\infty}^{w_{1-a}} dF_w-\frac{1}{2}\left(1+\lambda\right)$$

$$=-\lambda+\left(1+\lambda\right)F_w\left(w_{1-a}\right)=1-a$$

and

$$F_w\left(w_{1-a,\lambda}\right)=1-\frac{a}{1+\lambda}.$$

The right-quantile values of *w* and *y* are respectively:

$$w_{1-a}=F_w^{-1}\left(1-\frac{a}{1+\lambda}\right)>0$$

and

$$y_{1-a}=m+\left(1+\lambda\right)\phi w_{1-a}$$

Let $\lambda_o > 0$ and ϕ_o denote the asymmetry and tail parameters of the subjective distributions of overconfident individuals. The range between their left and right quantile values is

$$r_o=y_{o,1-a}-y_{o,a}=\left(w_{o,1-a}+\left|w_{o,a}\right|+\lambda_o\left(w_{o,1-a}-\left|w_{o,a}\right|\right)\right)\phi_o.$$

The range for rational expectations individuals with $\lambda_c = 0$ and ϕ_c is

$$r_c=y_{c,1-a}-y_{c,a}=\left(w_{c,1-a}+\left|w_{c,a}\right|\right)\phi_c.$$

According to the theory, overconfident individuals narrow down the confidence intervals of estimates, thus $r_o < r_c$. Therefore,

$$\frac{r_c}{r_o}=\frac{w_{c,1-a}+\left|w_{c,a}\right|\phi_c}{\left(w_{o.1-a}+\left|w_{o,a}\right|+\lambda_o\left(w_{o,1-a}-\left|w_{o,a}\right|\right)\right)\phi_o}>1$$

and

$$\phi_o<\frac{w_{c,1-a}+\left|w_{c,a}\right|}{w_{o.1-a}+\left|w_{o,a}\right|+\lambda_o\left(w_{o,1-a}-\left|w_{o,a}\right|\right)}\phi_c<\phi_c,$$

because $w_{c,1-a} + |w_{c,a}| < w_{o.1-a} + |w_{o,a}| + \lambda_o \left(w_{o,1-a} - |w_{o,a}| \right)$.

which proves the proposition. If $\lambda_o > 0$, $w_{c,1-a} + |w_{c,a}| < w_{o.1-a} + |w_{o,a}| + \lambda_o \left(w_{o,1-a} - |w_{o,a}| \right)$. Pessimistic forecasters widen the forecast confidence intervals, therefore, r_c / $r_p < 1$ and $\phi_p > \phi_c$.

References

Theodossiou, P. (2021). "Class of Skewed Distributions Based on a Two-Piece Generalized Framework," SSRN, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3763625.

54.

A Study of Female Executive Compensation and Performance

Jiangxia Liu

Valparaiso University, Valparaiso, USA

Abstract

In this study, we start with human capital and social capital theory and analyses female top executive's compensation and performance over a twenty-year period. A possibility suggested by organizational research is that women executives are paid less because they do not have the same opportunities for networking and building social capital that men executives have. we find that in the industry with higher percentage of female top executives, the gender pay gap narrows significantly. We also observe that women executives are paid less than male executives however female leadership apparently enhances corporate performance.

Categories

Accounting

56.

China's Luckin Coffee: From Star to Scam

Lin Chen [1], Dongfang Nie [2]

[1]Valparaiso University, Valparaiso, USA. [2]University of Texas Permian Basin, Odessa, USA

Abstract

In this study, we analyze the fraudulent activities of Luckin Inc., a Chinese company cross-listed in the U.S. stock exchange. We document various parties' responses to the fraud. We further analyze the fraud from four aspects of culture: Guanxi culture, scapegoat culture, absolute authoritarian leadership culture, and elite university culture. Our analysis contributes to the ongoing debate on whether the U.S. should require foreign firms to follow the U.S. accounting and auditing rules.

Categories

Accounting

57.

The Effect of Investor Sentiment on Earnings Management

Lin Chen [1], Dongfang Nie [2], Yuan Shi [3], Chaoping Li [4]

[1]Valparaiso University, Valparaiso, USA. [2]University of Texas Permian Basin, Odessa, USA. [3]Penn State Great Valley, Malvern, USA. [4]Skidmore College, Saratoga Springs, USA

Abstract

We examine earnings management (EM), as an important outcome of corporate reporting decisions, may be affected by investor sentiment. We use Michigan Consumer Confidence Index as the investor sentiment proxy and find a significant and positive association between investor sentiment and the propensity of (1) accrual earnings management (AEM), (2) the overall measure of real earnings management (REM), and (3) the specific REM mechanism through accelerating sales. Additional analyses after controlling for the effect of the Sarbanes-Oxley Act of 2002 (SOX) indicate that the association between investor sentiment and the propensity of AEM and REM are not affected by fundamental economic factors which are related to investor sentiment.

Categories

Accounting

Business Analytics

7.

Effect of predictive analytics in automotive industry - Review of the state of the art

Shubham Singh

Purdue University Fort Wayne, Fort Wayne, USA

Abstract

Predictive analytics could have an outsized role when it comes to quality control in the automotive industry. Traditional uses of business analytics for statistical process control and design of experiments methods have been studied and there are several use cases that are present in industries all around us. However, predictive analytics happens to provide avenues of improving efficiencies in the said processes which have an outsized impact on maintenance, understanding patterns of defect, identification of prospective sales opportunity, and understanding the reliability of car parts to predict cost of warranty. These instances of automotive market are supported with the help of current actions in the industry.

Keywords: predictive analytics, automotive industry, maintenance, defects, sales, warranty, analytics, business insight

Categories

Business Analytics

32.

Why did Your Reviews Receive Thumbs-Up Votes but Mine Receive Thumbs-Down Votes?

Yixiu Yu

Ball State University, Muncie, USA

Abstract

It is common for online review systems to have a voting feature to highlight helpful/unhelpful reviews. Users increasingly rely on online reviews before they make software download decisions or purchase decisions. The helpfulness/unhelpfulness voting feature can help users during their decision-making, especially when they need to deal with information overload. Therefore, it is important to examine the underlying factors of users' helpfulness/unhelpfulness voting. In this study, we extracted and analyzed online reviews of security software and browsers and focused on both the helpfulness ("thumbs-up") votes and unhelpfulness ("thumbs-down") votes. We found that the degree of ambivalence in a review, review rating extremity, length of pros reviews, length of cons reviews, length of review summary, and days since posting have significant impacts on the helpfulness votes and unhelpfulness votes, respectively, across two different types of software: security software and browsers. This study can advance the understanding of software users' voting behavior and provide recommendations for practitioners.

Categories

Business Analytics

40.

Community-Based Seeding Strategy for Viral Marketing in Social Networks

Wenjun Wang

University of Arkansas at Little Rock, Little Rock, USA

Abstract

With the widespread use of social media and mobile phones, word-of-mouth (WOM) promotion has become a powerful resource for both marketers and consumers in business. Viral marketing is an important and cost-effective marketing technique driven by WOM promotion. One of the fundamental problems in viral marketing is to find a small set of initial adopters (seeds) who may trigger the largest further adoptions through WOM messaging and influence diffusion in a social network. Community structure is a prominent property of social networks, which facilitate WOM communication and influence propagation within communities, but may also hinder influence spread between communities. In this paper, a novel community-based heuristic is proposed as the seeding strategy for viral marketing in social networks. This algorithm considers the underlying community structure of social networks, including the size and cohesiveness of individual communities. Experimental results on two real-life networks demonstrate its superior performance on both influence spread and time efficiency.

Categories

Business Analytics

2023 Midwest Decision Sciences Institute Annual Conference

47. Use of Process Mining Analytics to Overcome Deficiencies of Process Modeling Practice

Julian Koech, BORINARA PARK

Illinois State University, Normal, USA

ABSTRACT

This paper presents a combination of process modeling with process mining techniques to improve the quality of input data used for process modeling. Modelers often use insights from traditional sources such as process documentations, observations, and interviews with stakeholders. Obtained information from these sources might present a biased and inaccurate view of the real process. The proposed combination has been applied to a case study of a purchasing requisition process. The process's components such as activities, relationships, durations, and resources are extracted in process mining and used as input data for the creation of a process model that closely reflects reality.

KEYWORDS: Process mining; Process modeling; Discrete Event Simulation; Process Mapping

1 INTRODUCTION

Process modeling is a technique used in the analysis of complex business processes. Such business processes contain a series of connected entities that use the organization's resources to produce a specified output for a particular customer or market (Doomun & Vunka Jungum, 2008). Process modeling technique is therefore used to combine and define the relationship between these critical entities and resources, while its analysis produces key performance data outcomes that can be used to understand their cause and effects relationships. It allows managers to understand the fundamentals of business processes, identify opportunities for change and assess the impact of proposed changes (Aguirre et al., 2013). It is mostly used in industries like manufacturing, healthcare, transportation, and defense, to optimize processes, improve efficiency, and decision making (Solding, P., & Gullander, P. 2019). As such, it has become an essential tool for businesses and organizations looking to better their operations and gain a competitive advantage in their respective industries. Process modeling technique, however, relies heavily on the type of information used to create it. Making it susceptible to the "garbage in garbage out" phenomenon.

Most often, process modelers create models based on insights from documenting processes, interviews with stakeholders, and directly observing a process (Rozinat et al., 2009). Obtained modeling information through these sources might present a biased and inaccurate view of the actual process, resulting in disparity between the actual behavior of the process and the developed model (Abohamad et al.2017). Lynch et al. (2014) supports this by stating that modelers are often challenged to accurately model a system that closely reflect the reality. However, advancements in technology have enabled business processes to be supported by information systems such as Enterprise Resource Planning (ERP) and Supply Chain Management (SCM) which extract organizational data in the form of event logs that can be analyzed in process mining packages (Aguirre et al., 2013). Process mining is a recent discipline that generates useful information that organizations can use to understand and improve their business processes (van der Aalst, W.M.P., et al., 2011). It involves analyzing event log data to gain insight to patterns and trends in a process and understand how this impacts current performance (Van Der Aalst et al.,2004). Such knowledge from event logs can support modeling activities in the conceptual process modeling phase (Abohamad et al. 2017).

Research from Aguirre et al., (2013), Van Der Aalst, (2011), Weske, (2012) and Jadrić et al., (2020) has shown that the integration of process modeling with other techniques such as process mining would help solve process modeling’s challenge on input data needed to model accurate processes. Despite the growing number of studies relating to the coupling of process modeling with process mining, challenges persist in the research area which needs to be addressed (Van Der Aalst, 2012, Zakarija et al., 2020). For instance, the focus of most studies lacks clear guidelines on how process mining data can be extracted to support the construction of process models. The methods used to extract event log data are general and abstract, without results from real case study (Wang et al., 2018).

This paper presents the development of a process model using the process mining information to support the creation of the model that is less biased and has a close representation of the real process. The focus is to illustrate how process mining information can be extracted and used as input for process model development. The primary aim is to overcome the limitation of the process modeling practice where subjective information from the modeler is the main source of model inputs. A case study of a purchasing requisition is used to demonstrate how the actual data about the process has been extracted and transformed to create a process model with objective inputs.

2 LITERATURE REVIEW

The literature review focuses on process modeling previous practices, and combination efforts of process modeling and process mining.

2.1 Process Modeling Practices

A process model is an abstract description of an actual or proposed process. It represents selected process elements considered important in the model's creation. Abohamad et al.

(2017) stated that, to develop a validated process model, input data used in the creation of the model is a basic entity. Robinson et al (2012) created a Discrete-event simulation (DES) model of a hospital process that modeled the flow of patients through an outpatient theater over a day. They obtained information from interviewing senior executive managers, senior medical consultants, departmental managers, nurses, support workers, ward clerks and staff.

This data was used to validate the integration of Discrete-event simulation (DES) with lean methods. An approach they called SimLean. They modeled the hospital process which generated results on the use of resources, size of queues and the number of patients in the system (including arrivals and discharges). The model was based on the SimLean model in which its success was measured through a meeting with the lean team. They mentioned that their aim was not to develop an exact representation of the real process, but to set process names and the model data (quantities of resources and timings) to approximately the right level.

Mahmood et al. (2019) used simulation process modeling software, AnyLogic to build and simulate dynamic models for engineering systems. Input data for the process model was collected from a software called Simuland and from observations of how the process worked. Abohamad et al. (2017), objects this method of data collection by stating that modeling processes based on insights from observation might present biased and inaccurate view of the process.

Doomun & Vunka (2008) also developed a process model of a call center. They used input data recorded by computers called Automatic Call Distributor (ACD) and Interactive Voice Response (IVR) which used data to attribute the process flow of calls. The computers recorded the call's identification number, action taken, time elapsed since the earlier action, numbers of arrivals and abandonment, average service times, agent utilization, and the distribution of delays in a queue. They used quantitative analysis of the data obtained to measure the performance of the system. They faced a challenge where they lacked modeling data on the relationship between call-by-call stored at the IVR level. They recommended that data mining techniques can be applied to help mitigate this problem.

In support of this, process mining founder, der et al (2013) confirmed that event log data from process mining can support process modeling activities. He explained that event logs from information systems can be analyzed using process mining techniques, which help handle discrepancy between how processes operate in real life and process guidelines. Abohamad et al. (2017) adds that process knowledge discovered from analysis of event data can support process modeling activities. The process knowledge can then be cross-checked against traditional sources of information.

2.2 Attempts to combine Process Modeling and Process Mining

Abohamad et al. (2017) presented a hybrid framework within a hospital's emergency department that integrated process mining to support process modeling activities as part of the steps in developing simulation models. They used historical data, in an event log structure, provided by managers over one year. Using ProM and Disco software, they obtained statistical data in the conceptual modeling phase. They then used AnyLogic software to develop a discrete event simulation model, from which they found satisfactory results in comparison to the key performance indicators of the historical data. He recommended that efforts are needed to capture interaction between event log cases and extraction of resource schedules, queues, and activity duration for process modeling.

Măruşter & van Beest, (2009) proposed a process redesign methodology based on coupling process mining and process modeling. From three case studies, they used Colored Petri Net (CPN) simulation to simulate resource usage optimization and throughput time. The focus on this study, however, was on CPN simulation, which emphasized understanding of the first stage of process redesign project. Aguirre et al., (2013) complemented this work by putting emphasis on project understanding phase. They used a procurement process as a case study in which event log data was extracted from Enterprise Resource Planning (ERP) system. The information supplied by the log included case id, time stamps, activities and performers. Through the application of process mining algorithms such as alpha mining (van der Aalst et al., 2004), it was possible for them to automatically discover the actual process model using the ProM software functionality, and later represented it in a Petri Net.

They then simulated and evaluated different process improvement options based on waiting times calculated through the statistical analysis of the event log data. In their findings, the ProM petri nets used were not easy to interpret, especially for business users. They recommended other process mining tools such as Disco from Fluxicon that would supply more understandable visualization and animation of event log data. A general approach was proposed by Jadrić et al. (2020) where they focused on possible issues and advantages of the detection of student behavior and processes based on the data from a standard learning management system. He proposed the potential of using process mining results as input data for discrete-event simulation modeling in an educational sector.

3 PROBLEM STATEMENT

Despite preliminary research efforts to integrate process mining and process modeling, there is lack of clear guidelines as to how process mining data can be extracted and used as input data for constructing process models. Martin et al. (2015) attempts to bridge this gap by suggesting a structured overview of process modeling and how process mining can support the modeling tasks. They do not, however, show how these guidelines can be practically applied. Abohamad et al. (2017) even adds that most reported research neglects the internal structure of process modeling and how it is related to the output of process mining algorithms. This paper's main goal is to explore and demonstrate specific ways to improve the process modeling practice by integrating it with process mining information, ultimately resulting in better process models that better reflect the reality.

4 METHODOLOGY

A purchasing requisition process is used as a case study for the proposed framework. The event log of the actual execution of the purchasing requisition is obtained in an excel file. It is then imported to the process mining software, Disco, for analysis. This software is preferred due to the complexity of the proposed case study (having 9,120 cases) and its unstructured nature. Disco produces fuzzy miner algorithms or process maps which are suitable for analyzing such complex processes at various levels of granularity. It also helps provide meaningful abstraction and different views of the process (Günther & van der Aalst, 2007).

For a successful import of the log data in the process mining tool for analysis, the data is required to have at least a Case ID, Activity name, Timestamp and Resources (Rozinat et al., 2009). The Case ID associates an activity with a particular case, for example, client ID or purchase ID. The timestamp specifies when an event occurred while the resources are the human or machine resources that perform an activity (Figure 1). This information is taken and used as input data for process modeling, in a discrete-event simulation tool called AnyLogic. The process is broken down into three phases:

Case ID	Start Time	Complete	Activity	Resource	Role
339	31:00.0	23:00.0	Create Pur	Nico Ojenl	Requester
339	34:00.0	40:00.0	Analyze Pu	Maris Free	Requester Manager
339	29:00.0	52:00.0	Amend Pu	Elvira Lore	Requester
339	24:00.0	30:00.0	Analyze Pu	Heinz Guts	Requester Manager
339	36:00.0	38:00.0	Create Rec	Francis Od	Requester Manager
339	34:00.0	58:00.0	Analyze Re	Magdalena	Purchasing Agent
339	50:00.0	03:00.0	Amend Re	Penn Oste	Requester Manager
339	10:00.0	34:00.0	Analyze Re	Francois d	Purchasing Agent
940	31:00.0	08:00.0	Create Pur	Immanuel	Requester

Figure 1: Event log data of a purchasing requisition process

***Case study phase I:* Modeling of the main process flow using the frequency and performance feature data –** the goal of this phase is to use information in process mining frequency and performance features to create a process model of the main process flow. This phase explains how process mining elements such as activities, control flow, and paths are extracted and used to create a simple process model.

***Case study phase II*: Modeling of Resource units using information from tables in Process mining as input data–** this phase is intended to illustrate how resource units' allocation and types in process mining are added to the development of process models. ***Case study phase III*: Modeling process variations in process mining**- the goal of this phase is to zoom in to all other activities in process mining and use this information to model process deviations, loops and conditional paths.

4. 1 Case study phase I: Modeling of the main process flow using the frequency and performance feature data

After importing the event log data onto the process mining tool, Disco, a process map is generated, from which analysis of the process is done. These analyses are then used to extract useful input data for process modeling. The frequency and performance analysis features in process mining are used.

1) **A process map model of the main flow using frequency feature**: The map generated from event log data in process mining shows the top-down control-flow of the overall process (Figure 2). The map shows a set of activities, which are 18 that

the purchasing requisition goes through. The map is set to visualize the frequency of these activities (seen on the right side in figure 2). Frequency displays the number of units moving from one activity to the next (numbers beside the arrows in fig 2) and the number of units being worked on in each activity (numbers inside the rectangles). This information alone suffices as input data to develop the overall basic process model.

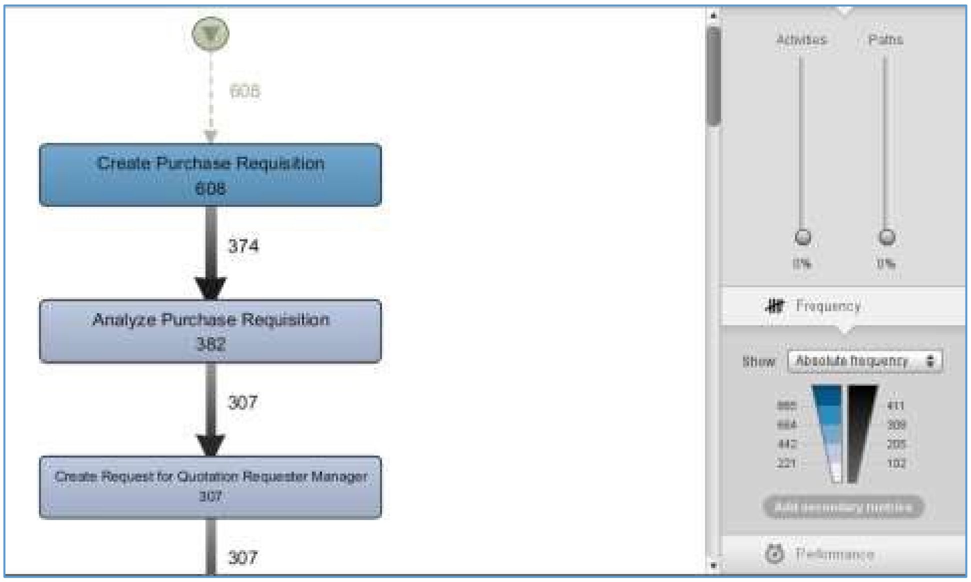


Figure 2: Main process of the purchase requisition on frequency metric in process mining

From Figure 2 above, the process begins with the creation of a purchase requisition. This is represented in process modeling as ‘purchasingRequest’ agent, the main agent. The main agent is a fundamental concept in process modeling that represents the objects or entities that move through the modeled system. It is the primary entity that performs actions, changes state, and interacts with other objects in the system. To create the purchase requisition in process modeling, the source block is used. This block is the starting point of a process model, it is used to generate model agents. Therefore, since there are observed 608 purchase requisitions created in process mining, the same number is used as input to represent the number of ‘purchasingRequest’ agents entering the process. The number is entered as ‘maximum number of arrivals’ in the source block (figure 3.a). Succeeding activities (Figure 3.b.1) in process mining are also used as input data for modeling. The processes are modeled using the Queue block in process modeling tool, AnyLogic to represent activities in the rectangles (Figure 3.b.2). The queue block is a buffer of agents waiting to be accepted by the next block(s) in the process flow. It is therefore the recommended block for modeling activities moving from one rectangle (in process mining), or one queue block (in process modeling) to the next.

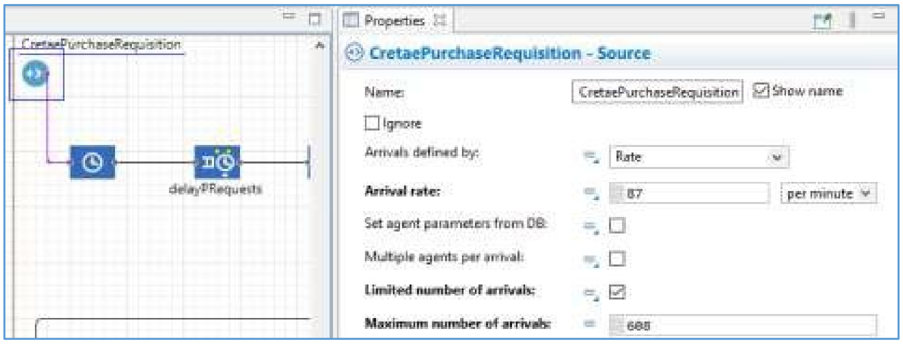


Figure 3(a): Modeling of the main agent in the source block

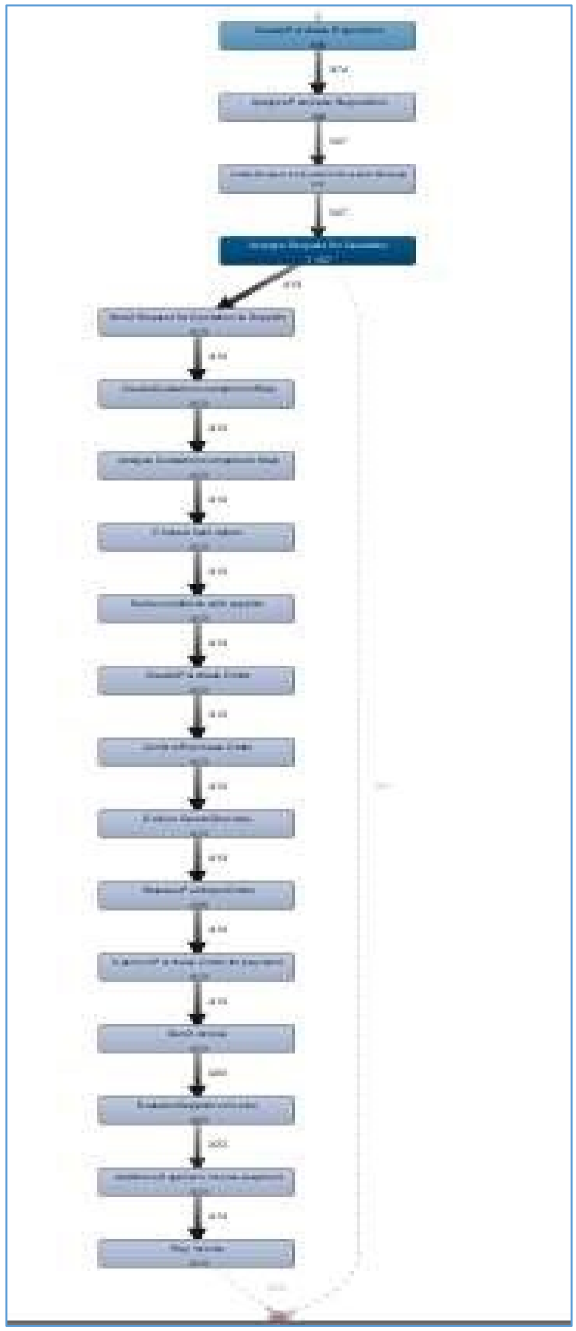


Figure 3(b.1) Process mining map of the overall process

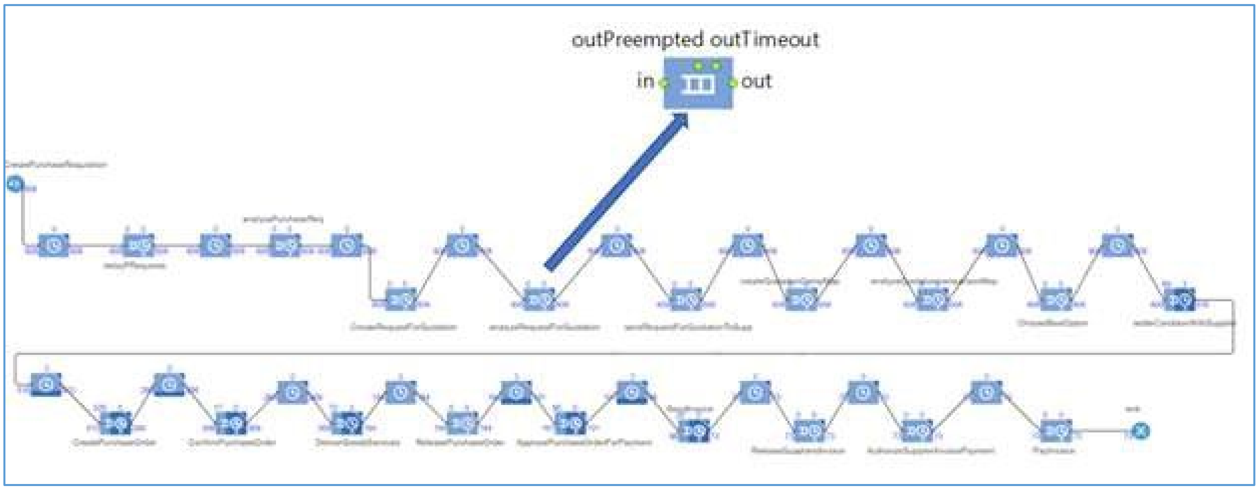


Figure 3(b.2): Process modeling logic of the overall process

After modeling the frequency of the process, the results obtained are a transition of the process flow from the process mining map in figure 3(b.1) to the process modeling logic of the same process in figure 3(b.2). However, these results only show activities, transitions, and paths which the main agent goes through. The frequency feature lacks the capability to provide any information on the actual performance metrics of the process, such as the time it takes for activities to transition to the next activity. This information is important in identifying delays and other inefficiencies in the process. Therefore, the process map's performance feature is applied to display key indicators of process efficiency such as cycle time, lead time, and processing time.

2) **A process map of the main flow using performance feature:** After applying the performance feature to the process map, more detailed and granular insights into process performance are obtained. This information helps in identifying bottlenecks, inefficiencies, and areas for improvement. These are key insights that Managers who use process analysis techniques look for. Therefore, it is important to add this information as input data for developing process models. Figure 4 shows the visualization of the process map when the performance feature is applied.

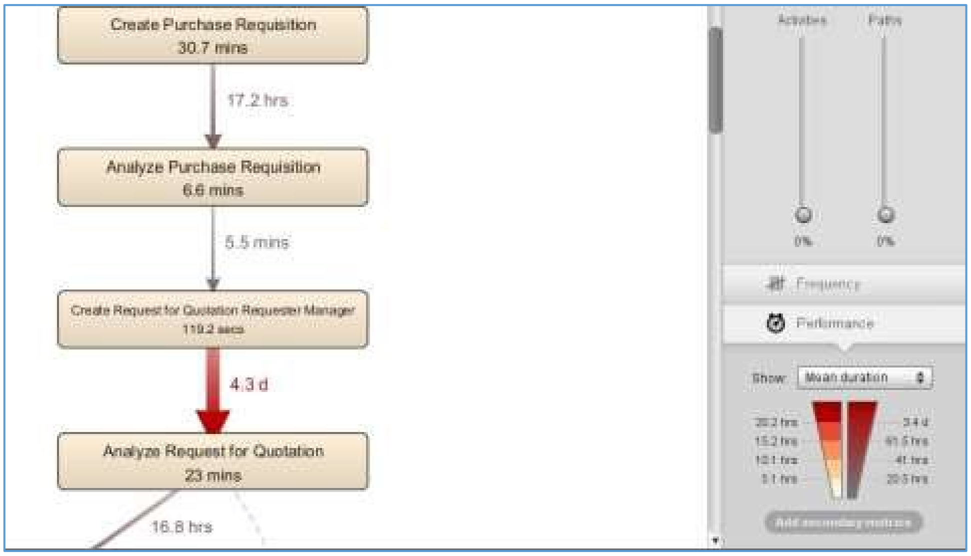


Figure 4: Main process of the purchase requisition on performance metric in process mining

The numbers inside the rectangles represent time that was taken by a certain resource to perform the activity, while the number beside the arrow indicates the time taken for the main agent (the purchasing requisition) to move from one activity to the next. These numbers are used as input data in process modeling, where the Delay block in process modeling is used to represent the time in which the agent takes to move from one activity to another. The delay block can delay agents for a given amount of time. It evaluates the agent's delay time dynamically, may be stochastic and depends on the agent and any other conditions. It also allows for multiple agents (up to the given Delay capacity) to be delayed simultaneously and independently. For example, on the process map, the agent took 17.2 hours to move from the first activity 'Create Purchase Requisition' to the second activity, 'Analyze Purchase Requisition'. This is represented in the delay block's parameter (Figure 5), where the delay time is set to 17.2 hours. This will make the transition of the agent from the first activity to the second, to take exactly this amount of time when the model is run (figure 5.1). This performance or behavior can be observed in the developer panel (on the right side in figure 5.1) which provides access to the model data during the model run time.

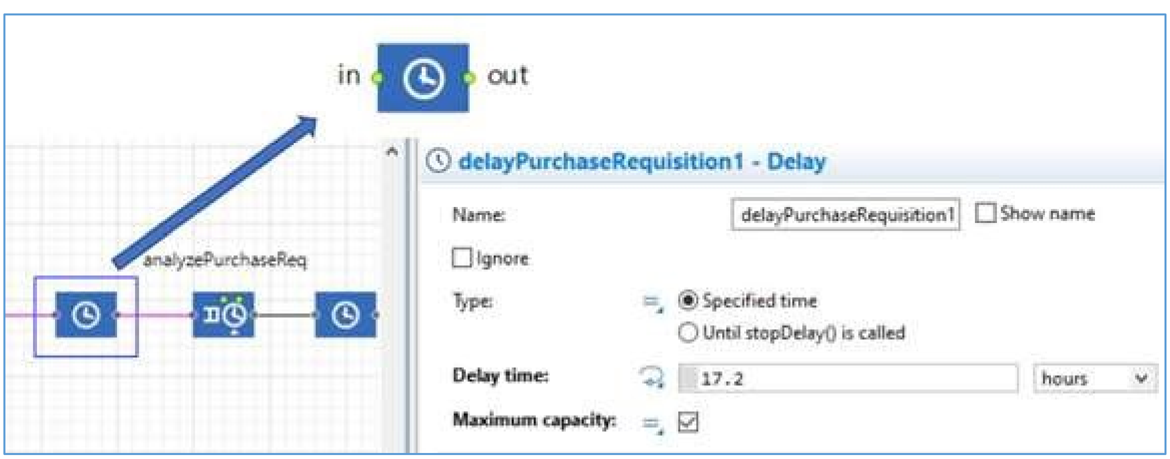


Figure 5: The delay parameter in process modeling, representing the time taken from one activity to the next in process mining

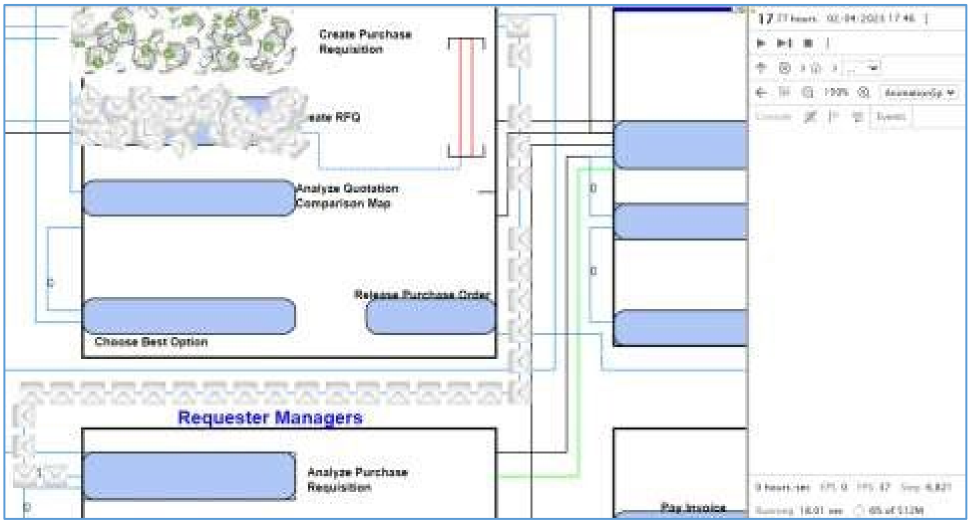


Figure 5.1: Process modeling output showing transition from activity 'Create Purchase Requisition' to activity, 'Analyze Purchase Requisition' taking 17.2 hours

Additionally, the process mining data also shows the duration of time that each activity's resource agent took to work on the purchase requisition (the numbers in the rectangular shapes). For example, activity three, 'Create Request for Quotation Requester Manager' took 119.2 seconds (about 2 minutes) to be processed by a particular resource agent. This information is used as input in process modeling, where the Service block is used to model this process. The function of the service block is to delay the agent, seize a given number of resource units, and release the seized units. Since there are no resource units up to this point, the seize block is used to delay the agent. The 119.2 seconds (about 2 minutes) was used in the delay time parameter area of the service block (Figure 6), to delay the agent in activity three, for this period.

Figure 6: The Service block parameter in process modeling representing the time an agent takes in an activity

After running the model, activity 3 was observed in the developer panel to take 119.2 seconds (about 2 minutes) before moving to the next block. After developing the main process flow's activity frequency and performance, the main flow is seen to lack insight into the behavior of individual process participants, or the interactions between them. This is particularly an important aspect of any business process. As mentioned by Doomun &

Vunka Jungum, (2008) that a business process is a series of connected entities that use the organization's resources to produce a specified output for a particular customer or market. Therefore, the process mining tool was further analyzed to identify the resources, and their relationships.

4.2 Case study phase II: Modeling of Resource units using information from tables in

Process mining as input data

Among the output information displayed in Disco, process mining tool, 'cases' are one of them. They are represented as rows in the event log (Figure 7). Each row contains information about a specific event in the process, such as the activity that was performed, the time at which it occurred, the resource that executed it, the date the activity was executed, and duration it took for each activity (Rozinat et al., 2009).

Since the model with specific activities and times has been created, the focus of this phase is to implement the resource agents that oversaw these activities. The following resource agents were identified in the 'cases' section of process mining: Requesters, Request managers, Purchasing agents, Financial managers, and Suppliers. 608 cases were also observed, but only case 335 was selected to represent the main flow. Case 335 was seen to have a similar number of activities as in the main process map. Using the table function under cases in process mining, it is possible to identify the major roles of the resource agents and the activities they were involved in at a particular point in time.

	Activity	Resource	Date	Time	Duration	Role
1	Create Purchase Requisition	Kim Passa	02/15/2011	05:46:00	7 mins	Requester
2	Analyze Purchase Requisition	Maris Freeman	02/16/2011	16:55:00	6 mins	Requester Manager
3	Create Request for Quotation Requester Manager	Francis Odell	02/16/2011	17:05:00	1 min	Requester Manager
4	Analyze Request for Quotation	Karel de Groot	02/20/2011	11:40:00	27 mins	Purchasing Agent
5	Send Request for Quotation to Supplier	Francois de Perrier	02/20/2011	17:30:00	29 mins	Purchasing Agent
6	Create Quotation comparison Map	Francois de Perrier	02/20/2011	18:24:00	5 hours, 6 mins	Purchasing Agent
7	Analyze Quotation comparison Map	Immanuel Karagianni	02/21/2011	06:07:00	28 mins	Requester
8	Choose best option	Alberto Duport	02/21/2011	06:35:00	0 millis	Requester
9	Settle conditions with supplier	Karel de Groot	02/21/2011	09:26:00	13 hours, 48 mins	Purchasing Agent
10	Create Purchase Order	Magdalena Predutta	02/22/2011	09:26:00	8 mins	Purchasing Agent
11	Confirm Purchase Order	Karen Clarens	02/23/2011	02:48:00	5 mins	Supplier
12	Deliver Goods Services	Esmeralda Clay	02/23/2011	21:39:00	1 day, 20 hours	Supplier
13	Release Purchase Order	Miu Hanwan	02/26/2011	01:43:00	1 min	Requester
14	Approve Purchase Order for payment	Francois de Perrier	02/28/2011	08:34:00	1 min	Purchasing Agent
15	Send invoice	Kiu Kan	02/28/2011	20:12:00	0 millis	Supplier
16	Release Supplier's Invoice	Pedro Alvares	03/01/2011	10:29:00	5 mins	Financial Manager
17	Authorize Supplier's Invoice payment	Pedro Alvares	03/01/2011	10:34:00	0 millis	Financial Manager

Figure 7: Resource agents in process mining

This information was then used to create a process model, containing resources. In the process modeling tool, the Resource Pool block was used. It is used to define a set of resource units that can be seized and released by agents. It can be of three types: 1) Static resources which are bound to a particular location, 2) Moving resources which can move on their own, they can represent staff, vehicles, etc. And 3) Portable resources which can be moved by agents or by moving resources. In this case study, the resource agents' type was static as they did not require movement to carry out a task. The resource pool block can define the capacity of the resource agents in the process. In this case, requesters in process mining were counted to have carried out activities 4 times, the request managers 3 times, Purchasing agents 3 times,

financial managers 3 times, and suppliers 3 times. These numbers were then used in the process modeling to represent the capacity of the resource agents using parameters (Figure 8)

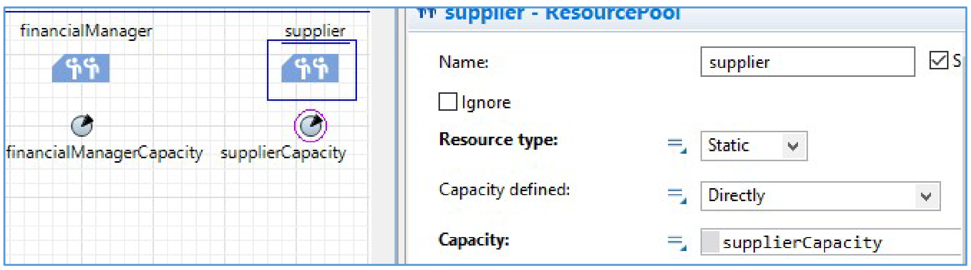


Figure 8: Resource pool block in process modeling

After introducing the resource agents in the process model, the resources were then connected to the activities they carried out. This is done in the service block's resource pool parameter from where the resources are requested for a specific activity. After modeling the resources, their utilization was measured using the key performance indicators in process modeling to help identify areas where resources are being over-utilized or under-utilized. Table 1 shows the analysis results of resource utilization after the model is run for one month. This information can be used to optimize resource usage and improve the overall efficiency of the system (Abohamad et al.2017).

Table 1: Resource Utilization Key Performance Indicators (KPI) from Process Modeling

Resource	%Utilization
Requester	34%
Requester Manager	97%
Purchasing Agent	70%
Financial Manager	10%
Supplier	10%

After completion of the model with resources included, it was important to realize that modeling only one case out of 608 cases does not really represent the overall process reality as variations in processes are inevitable. The next phase attempts to bridge this gap by modeling subprocesses identified in process mining.

4.3 Case study phase III: Modeling process variations in process mining

Real processes are complicated, and often complex such that it is not possible to look at every detail at once (Rozinat et al., 2009). From the process mining information, there are 98 process variations uncovered. These are specific instances or occurrences of the process when activities were executed multiple times. Process variants exhibit different variations in terms of the activities performed, the sequence in which they are performed, and the time taken to complete them. The variants are captured as different process instances in process mining (figure 9a)

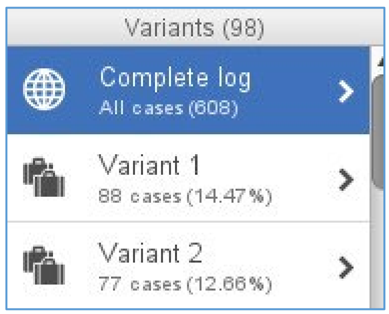


Figure 9a: 98 Variants uncovered from event log data in process mining

Process mining algorithms, however, have the capability to choose the level of detail to look at processes (Rozinat et al., 2009). Using the paths and activities features in process mining, it is possible to identify the variant that represents the overall process flow map. When activities and paths sliders are put to their lowest point, 0%, only activities performed in the most frequent process variant are observed, which results in visualization of the main flow of the process. When moved to the highest point, 100%, every activity and path ever performed is visualized.

(Figure 9b)

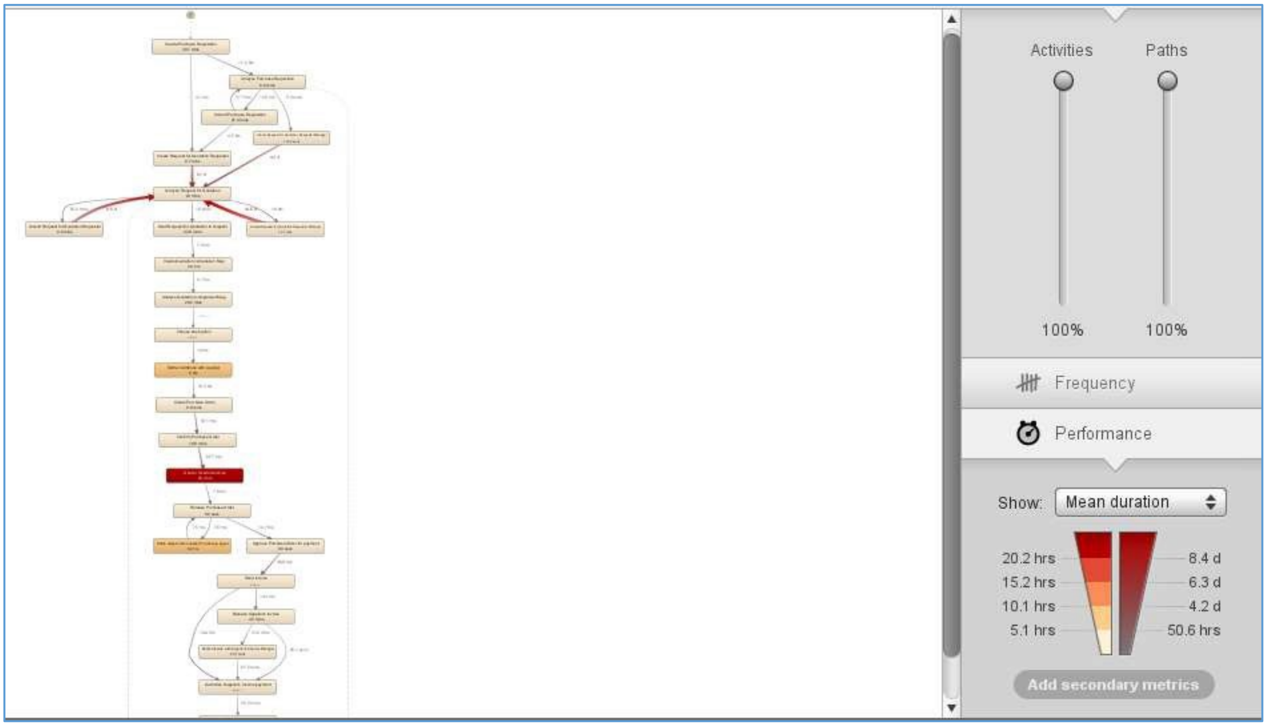


Figure 9b: Sub process of all the activities and paths

The results in the visualization show 24 activities done. This additional information is then used as input data to create an advanced process model with more routes and sub processes. A variant that represents 24 activities is chosen at random. It is represented in table format in process mining (Figure 9c). This variant contains information such as activities, resources, time and duration.

	Activity	Resource	Date	Time	Duration	Role
8	Analyze Request for Quotation	Karel de Groot	07/22/2011	00:09:00	18 mins	Purchasing Agent
9	Send Request for Quotation to Supplier	Magdalena Predutta	07/22/2011	13:46:00	22 mins	Purchasing Agent
10	Create Quotation comparison Map	Francois de Perrier	07/22/2011	18:07:00	2 hours, 28 mins	Purchasing Agent
11	Analyze Quotation comparison Map	Anna Kaufmann	07/23/2011	03:27:00	27 mins	Requester
12	Choose best option	Fjodor Kowalski	07/23/2011	03:54:00	0 millis	Requester
13	Settle conditions with supplier	Magdalena Predutta	07/23/2011	17:00:00	14 hours, 28 mins	Purchasing Agent
14	Create Purchase Order	Magdalena Predutta	07/24/2011	10:00:00	10 mins	Purchasing Agent
15	Confirm Purchase Order	Carmen Finacse	07/24/2011	22:16:00	7 mins	Supplier
16	Deliver Goods Services	Carmen Finacse	07/26/2011	11:05:00	1 day, 16 hours	Supplier
17	Release Purchase Order	Esmana Liubiata	07/28/2011	10:23:00	1 min	Requester
18	Settle dispute with supplier Purchasing Agent	Francois de Perrier	07/28/2011	13:07:00	9 hours	Purchasing Agent
19	Release Purchase Order	Anne Olwada	07/29/2011	05:37:00	1 min	Requester
20	Approve Purchase Order for payment	Karel de Groot	07/29/2011	15:32:00	1 min	Purchasing Agent
21	Send invoice	Sean Manney	08/01/2011	01:11:00	0 millis	Supplier
22	Release Supplier's Invoice	Karalda Nimwada	08/01/2011	15:23:00	7 mins	Financial Manager
23	Authorize Supplier's Invoice payment	Pedro Alvares	08/01/2011	15:30:00	0 millis	Financial Manager
24	Pay invoice	Karalda Nimwada	08/01/2011	16:31:00	13 mins	Financial Manager

Figure 9c: Table showing variant details that have 24 activities

When observed closely in the process map, some activities are seen to be branching off from activities in the main path. For example, the activity 'Amend Request for Quotation Requester' (Figure 9.1a) is one of the activities branching off from the main process activity, 'Analyze

Request for Quotation'. It takes 50.2 minutes from the main activity and 8.5 days back. Other examples of such activities are shown in Figures 9.1a and 9.2b. All the information about these activities is represented in figure 9c above.

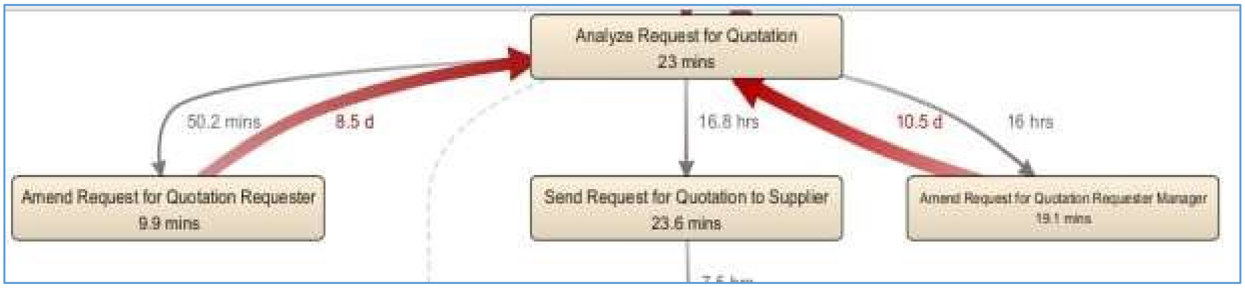


Figure 9.1a Activities branching off from the main process

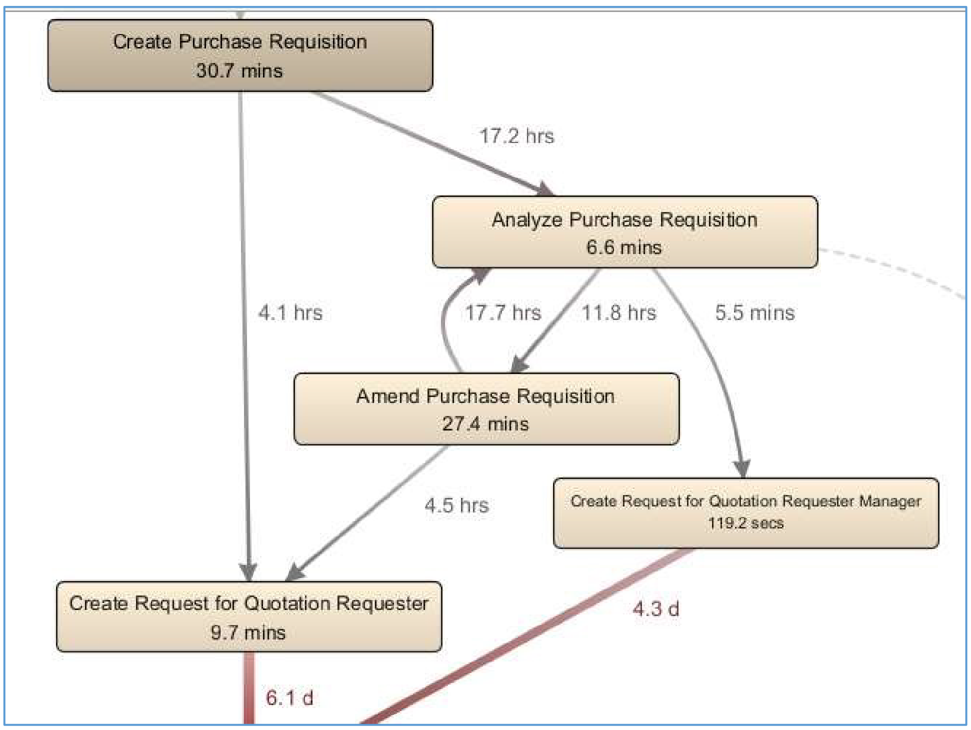


Figure 9.2b Additional activities branching off from the main process

To model such activities, an additional block in process modeling is introduced, the select output block (Figure 10). It is typically used when there are multiple output paths available from a process, and the process model must decide which path to take. It is used here as a tool for making decisions on where the activities are routed to. Since this information is provided in process mining, the select output is then used at all instances where the branching of activities is taking place.

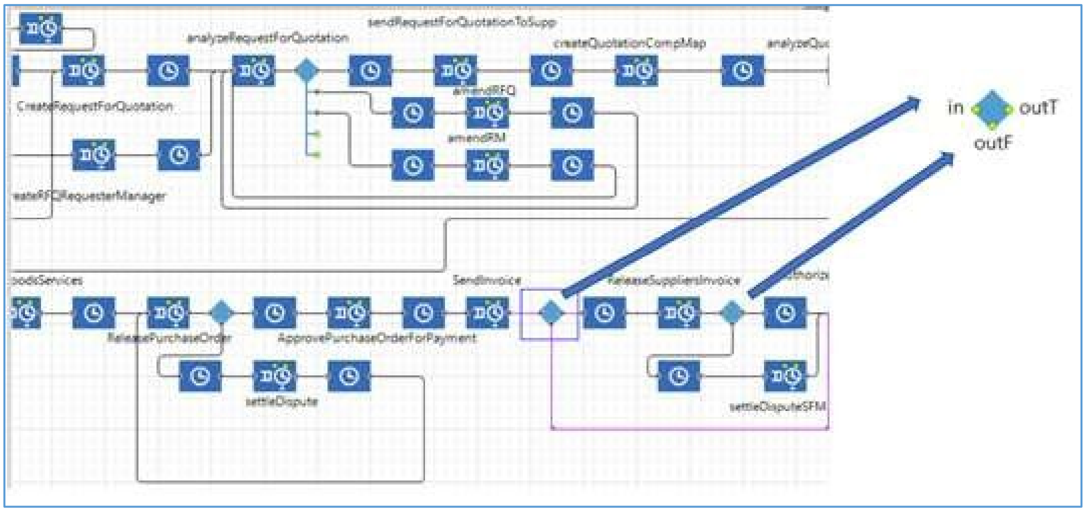


Figure 10: Subprocesses representation in process modeling

One challenge with modeling this information however is that, usually there is a reason for activities to move back and forth between resource agents. This kind of information is obtained from interviewing process owners or managers. However, since there is no record for reasons why the activities are branching off to different resource agents, the model used the select output default probability to branch off these activities. Had there been information on why activities branch off, the select output block would have been set to route each individual activity to its succeeding activity based on a condition created in java programming. After modeling the subprocesses, the duration of time taken from the start of the process up to the end was measured and found to be 1004.63 (Figure 10.1a). To measure duration in process modeling, 'TimeMeasureStart' as well as 'TimeMeasureEnd' blocks are used (Figure 10.1b). The blocks measure the time the agents spend between activities by remembering the time when an agent went through the 'TimeMeasureStart'. Later, when the agent goes through TimeMeasureEnd block, the time the agent spent between these two "marker" blocks is measured.

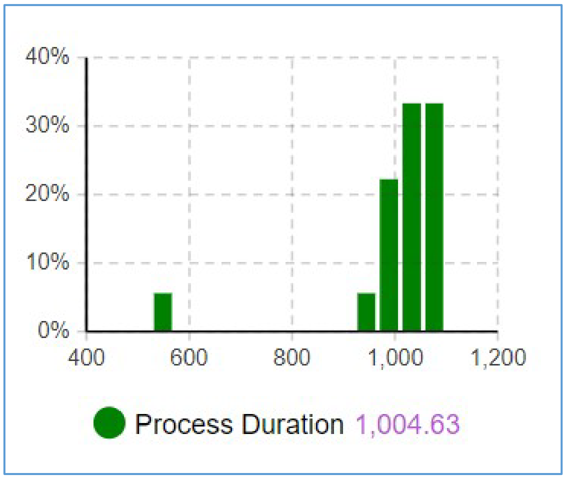


Figure 10.1a: Process duration from start to finish

Figure 10.1 b: Time Measure Start block in process modeling

This information provides valuable insight into the performance and efficiency of the system. By measuring the time taken to complete a process or a set of processes, we can identify

bottlenecks, inefficiencies, and areas for improvement. This information can then be used to optimize the system, increase productivity, and reduce costs. For example, in the process

mining output, an observation is made where activity 'Deliver goods Services' is highlighted in

red, to show that this activity is a pain point in the process, a main bottleneck due to the length of time it takes to execute this process (Figure 10.2). This issue is also observed after the process is modeled, and the output results show overcrowding at this activity (Figure 10.2a).

Figure 10.2: Illustration of the problems and inefficiencies in process mining.

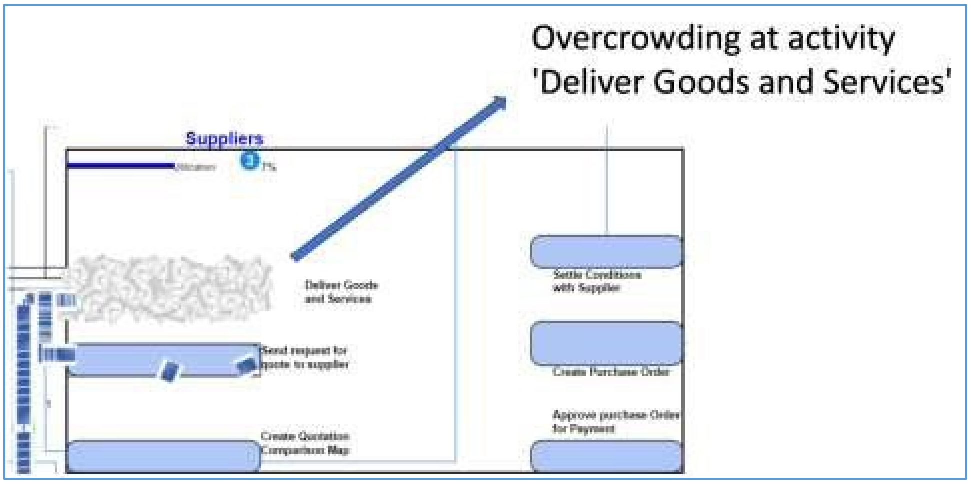


Figure 10.2a: Illustration of the problems and inefficiencies in process modeling.

Using process modeling, a what-if scenario is created to mitigate this issue. An improvement suggestion is made to omit activity 'Amend Request for Quotation Requester Manager', as it is seen to take 16 hours and 10.5 days (about 1 and a half weeks) to move back and forth from activity 'Analyze Request for Quotation' (Figure 10.3).

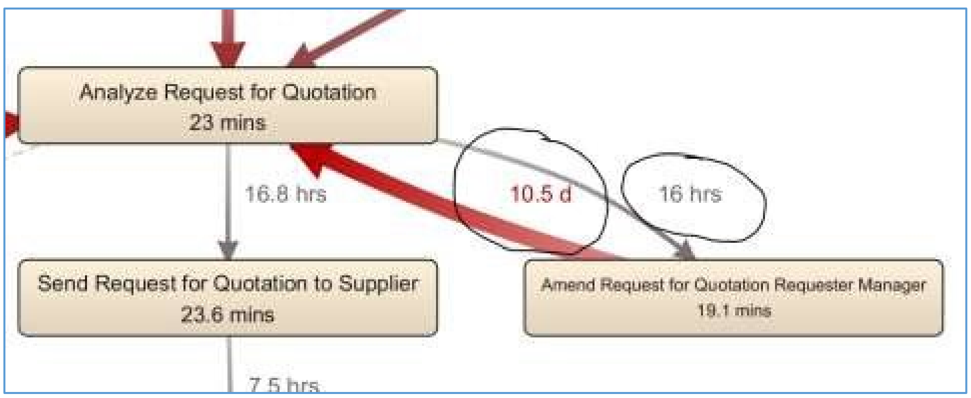


Figure 10.3: An activity assumed to cause the root problem

Another improvement suggestion is to assign the task 'Amend Request for Quotation 'to the Requester only as the task is being performed by the Requester Manager Resource as well, causing a waste in the utilization of resources. The utilization for the Requester had been observed to be 34% initially, while that of the Requester Manager was 97%. After the application of the suggested improvements, the results show that utilization of the Requester manager improved from 97% to 70% (Figure 10.4) while the cycle time of the process is significantly reduced from 1082.86 to 844.67 (figure 10.5a and 10.5b).

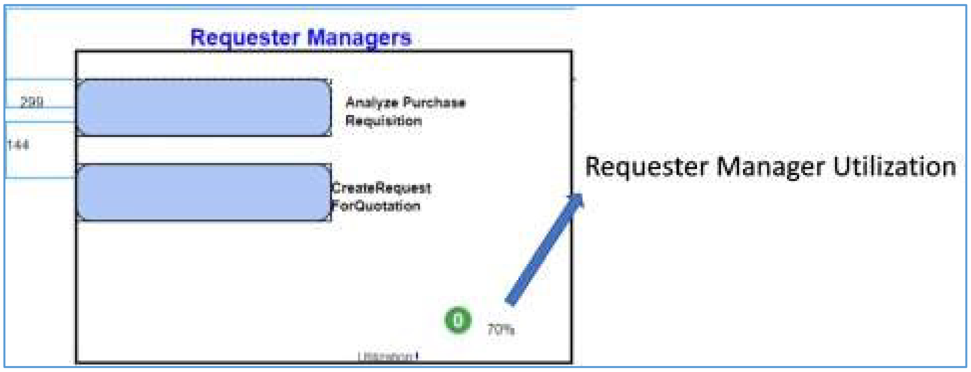


Figure 10.4: Observed change in resource utilization of the Requester Manager

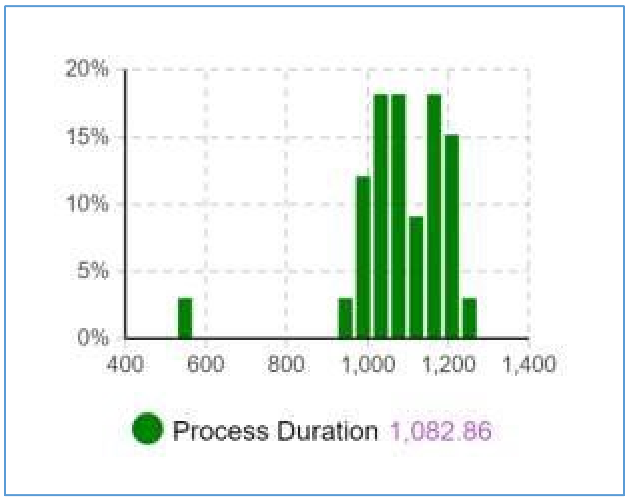


Figure 10.5a: Process Duration without improvement suggestions

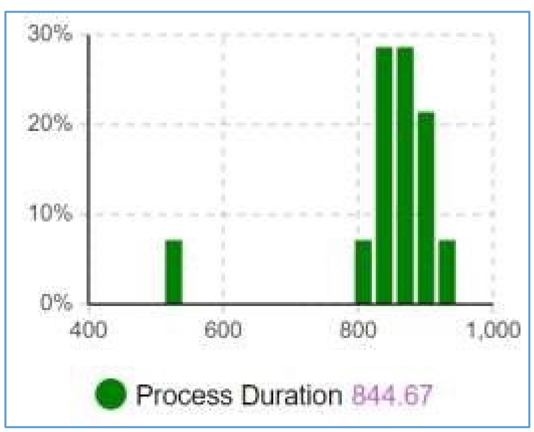


Figure 10.5b: Process Duration with improvement suggestions

After applying these improvement suggestions, process modeling enabled the evaluation of the impact of adding or removing tasks to resources or changing the workflow. Making process modeling technique more optimizable and capable of reducing the cycle time, increasing throughput and improving the quality of the process.

5 LESSONS LEARNED

In the attempts to apply process mining analytics information as input data for more accurate process models, the major lessons learned are summarized as follows:

1. Process mining and process modeling techniques are complementary. They provide different perspectives on the same process where process mining provides visualization of the entire process information about the process flow, timestamps, activity names, and resource assignments.
2. Process modeling, however, uses the same information to build models that reflect the actual process flow.
3. Process mining can easily help identify areas of the process that need improvement
4. While process modeling can be used to test potential changes to the process before they are implemented.
5. Process modeling also provides more information on process elements like resource utilization and duration. This is achievable through its key performance indicators that allow for easy analysis and measurement of process performance.
6. With process modeling capability to create what-if scenarios, continuous improvement of the system is achieved.

By combining the insights gained from both approaches, organizations can make data-driven decisions to improve their processes and achieve better performance outcomes.

6 CONCLUSION AND LIMITATIONS

In this paper, the process mining analytics was instrumental in uncovering the system's real process flow and related information, which would not have been acquired through interviews and subjective observations. Extracting such real process information from the process mining gives tangible insight on the major activities, transitions, and paths that the main agent goes through. This, however, presents a limitation where information on the process's performance metrics is not represented. To fill this gap, the performance feature is used which uncovers key information such as cycle time, lead time, and processing time. While creating a process model with this information, it is also realized that the main flow lacks insight into the behavior of individual process participants, or the interactions between them. This led to the use of the roles feature in process mining which has information on human resources that perform certain activities. Modeling this information provided more insight into the utilization of such resources using the key performance indicators feature in process modeling. It was also clear that processes are complex. The process mining information had different activities and sequences that were performed at various times. This is captured in process mining in the form of variants. There were 98 variants uncovered from this case study, from which only one that stood for the entire flow of the process was modeled. Selecting one out of 98 variants presents a challenge as the process model might not accurately represent the real process.

Nevertheless, the results from modeling the activities resulted in the same output seen in process mining flow, where an event of a bottleneck that was visualized in red in process mining, had the same behavior where the same instance was seen to overcrowd in process modeling. This acted as validation that process mining data can be used as input data for process modeling. Additionally, by comparing the results of the process model with the actual process flow in process mining, users can find any discrepancies and make necessary adjustments to the model. This was an advantage realized from using process modeling. It's capability to edit the model and to create what-if scenarios allowing for experimentations. This led to a suggestion for a solution to the main process flow. The solution was then implemented, and actual improvement results were obtained. However, for any major changes in the process flow, confirmation is needed from process managers or process owners. Modelers need to conduct interviews or meetings to discuss the changes before implementing them. This was not captured in the model as it relied on information from process mining alone and the modeler's assumption.

Another limitation was that once processes were modeled, there was no way to confirm that it produced the same information as that in process mining. However, the model's behavior, the same as that of the process mining output, acted as validation. In conclusion, by combining these two approaches, organizations can find process variations and inefficiencies and then use process modeling to design and implement more efficient processes. This approach has been successfully applied in this paper. The overall combination of process mining and process modeling is seen to be powerful for improving organizational processes and achieving greater efficiency and effectiveness. Further research is needed to explore process mining techniques which lack the capability to create hypothetical scenarios and observe different outcomes. Future work is recommended to study how process mining predictive capability can be improved using process modeling.

REFERENCE

Aguirre, S., Parra, C., & Alvarado, J. (2013). Combination of process mining and simulation techniques for business process redesign: A methodological approach. *Lecture Notes in Business Information Processing*, 24–43. https://doi.org/10.1007/978-3-642-40919-6_2

Augusto, V., Xie, X., Prodel, M., Jouaneton, B., & Lamarsalle, L. (2016). Evaluation of discovered clinical pathways using process mining and joint agent-based discrete-event simulation. 2016 Winter Simulation Conference (WSC).

https://doi.org/10.1109/wsc.2016.7822256

Banks, J. (1998). *Handbook of Simulation: Principles, methodology, advances, applications and practice*. Wiley.

Brailsford, S. C. (2014). *Discrete-event simulation and system dynamics for Management Decision making*. Wiley.

Curtis, B., Kellner, M. I., & Over, J. (1992). Process modeling. *Communications of ACM*, *35*(9), 75–90. https://doi.org/10.1145/130994.130998

Doomun, R., & Vunka Jungum, N. (2008). Business process modelling, simulation and reengineering: Call Centers. *Business Process Management Journal*, *14*(6), 838–848. https://doi.org/10.1108/14637150810916017

Greasley, A. (2003). Using business-process simulation within a business-process reengineering approach. *Business Process Management Journal*, *9*(4), 408–420. https://doi.org/10.1108/14637150310484481

Günther, C. W., & van der Aalst, W. M. (2007). Fuzzy mining – adaptive process simplification based on multi-perspective metrics. *Lecture Notes in Computer Science*, 328–343. https://doi.org/10.1007/978-3-540-75183-0_24

Jadrić, M., Pašalić, I. N., & Ćukušić, M. (2020). Process mining contributions to discrete-event simulation modelling. *Business Systems Research Journal*, *11*(2), 51–72.

https://doi.org/10.2478/bsrj-2020-0015

Mahmood, I., Kausar, T., Sarjoughian, H. S., Malik, A. W., & Riaz, N. (2019). Integrated modeling, simulation, and Analysis Framework for Engineering Complex Systems. *IEEE*

(Institute of Electrical & Electronic Engineers) (Institute of Electrical & Electronic

Engineers) (Institute of Electrical & Electronic Engineers) (Institute of Electrical & Electronic Engineers) (Institute of Electrical & Electronic Engineers) (Institute of Electrical & Electronic Engineers) Access, *7*, 67497–67514.

https://doi.org/10.1109/access.2019.2917652

Martin, N., Depaire, B., & Caris, A. (2015). The use of process mining in business process simulation model construction. Business & Information Systems Engineering, 58(1), 73– 87. https://doi.org/10.1007/s12599-015-0410-4

Mărușter, L., & van Beest, N. R. (2009). Redesigning business processes: A methodology based on simulation and process mining techniques. Knowledge and Information Systems, 21(3), 267–297. https://doi.org/10.1007/s10115-009-0224-0

Robinson, S., Radnor, Z. J., Burgess, N., & Worthington, C. (2012). Sim Lean: Utilising simulation in the implementation of Lean in Healthcare. *European Journal of Operational Research*, *219*(1), 188–197. https://doi.org/10.1016/j.ejor.2011.12.029

Rozinat, A., Mans, R. S., Song, M., & van der Aalst, W. M. P. (2009). Discovering simulation models. *Information Systems*, *34*(3), 305–327. https://doi.org/10.1016/j.is.2008.09.002 van der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering, 16(9), 1128–1142. https://doi.org/10.1109/tkde.2004.47

Wang, Y., Zacharewicz, G., Traoré, M. K., & Chen, D. (2018). An integrative approach to Simulation Model Discovery: Combining system theory, process mining and Fuzzy Logic. *Journal of Intelligent & Fuzzy Systems*, *34*(1), 477–490. https://doi.org/10.3233/jifs-17403

58. EXPLORATORY INVESTIGATION OF DIGITAL INNOVATION: USERS' REACTIONS TO EMERGING DIGITAL TECHNOLOGY—CHATGPT

Peiwei Li, College of Informatics, Northern Kentucky University, Louie B Nunn Dr, Highland Heights, KY 41099, 859-572-6408, lip1@nku.edu
Haoqiang Jiang, College of Informatics, Northern Kentucky University, Louie B Nunn Dr, Highland Heights, KY 41099, 859-572-5992, jiangh1@nku.edu
Xiaoni Zhang, Collat School of Business, The University of Alabama at Birmingham, CSB 371 | 710 13th Street South, Birmingham, AL 35294, 205-934-8890, xzhang2@uab.edu

ABSTRACT

Digital innovation is driving progress and change across industries, and the emergence of advanced AI technologies (e.g., ChatGPT) is revolutionizing how we work, communicate, and interact. ChatGPT has been praised for its intelligent capability to recognize natural language and respond like a human, making it an invaluable tool in various industries. One of the most significant impacts of ChatGPT has been in the field of business analytics. It has enabled businesses to extract valuable insights from unstructured data, perform sentiment analysis, and automate routine tasks, among other things. While there are concerns about the potential displacement of jobs by automation and AI technologies, there is also a growing interest and excitement among professionals in the business analytics domain toward ChatGPT and its potential applications. To explore the potential usage of ChatGPT in business analytics, this study aims to understand users' reactions to ChatGPT and how these reactions influence its continuous development and management. The study is grounded in the existing literature on ChatGPT and digital innovation and will contribute to understanding the potential usage and impacts of ChatGPT on businesses

Keywords: Digital Innovation, ChatGPT, Business Analytics, Natural Language Processing (NLP), Users' Reactions

INTRODUCTION

Digital innovation involves using various technological tools and methods during the ideation process (Nambisan et al., 2017). It is one of the most crucial factors that drive progress and change (Boudreau, 2010). As the industry environments become more dynamic and IT competencies are essential for organizational success (Eesley & Wu, 2019), fast-paced digital innovations—from advanced AI technologies to sophisticated communication tools—have revolutihowthe way we work, communicate, and interact with each other. One of the most prominent tools used in the digital innovation process is the ChatGPT, which is a large language model developed by OpenAI (Introducing ChatGPT, 2022). ChatGPT signifies an innovation in natural language processing (NLP) and has been widely acknowledged as a groundbreaking innovation in the field, and has rapidly become a widely known question-and-answer dialogue system (Frieder et al., 2023). ChatGPT has been featured in various traditional media outlets and referred to as a tool for digital innovation (Roose, 2022).

ChatGPT has been praised for its intelligent capability to recognize natural language and respond like a human, making it an invaluable tool in various industries, such like computer coding, finance analysis, business analytics, and healthcare. According to a study by Absolutdata, the demand for NLP solutions is rapidly increasing, driven by the growing need for intelligent automation and improved customer experiences (Absolutdata, 2019). This trend has paved the way for innovations like ChatGPT, which has been specifically designed to improve communication and collaboration between humans and machines. As such, ChatGPT could help people to solve complex problems and provide a comprehensive guide to specific tasks such as coding and debugging (Sobania et al., 2023), math problems (Frieder et al., 2023), high education (Rudolph et al., 2023), medical research and writing (Gilson et al., 2023; Kitamura, 2023; Liebrenz et al., 2023), accounting professionals (Alshurafat, 2023), and financial research (Dowling & Lucey, 2023).

Building upon the advanced emerging digital technology (i.e., ChatGPT), the field of business analytics—a comprehensive work that required both math skills and computer programming skills (Chen & Jiang, 2020)—has seen a rapid transformation. Extant research has identified the ability of language models to perform sentiment analysis on unstructured data (Hirschberg & Manning, 2015; Yue et al., 2019), such as social media data. As a cutting-edge language model, ChatGPT uses advanced NLP techniques to analyze and understand large amounts of data. By providing businesses with the ability to extract valuable insights from vast amounts of data, ChatGPT has the potential to become a game-changer in the business analytics discipline. Furthermore, existing research identifies the great potential of ChatGPT in dealing with Data Visualizations (Maddigan & Susnjak, 2023) and data literacy (Lund & Agbaji, 2023).

As various digital innovations continue to evolve and improve, these emerging digital technologies (e.g., ChatGPT) will likely become an essential tool and, eventually, replace some job positions in the business analytics domain by automating tasks, performing sentiment analysis, and analyzing unstructured data. Considering the capabilities of ChatGPT and other advanced AI technologies, there are concerns among people that they may become less competitive in the job market, while ChatGPT can perform specific tasks more efficiently and at a lower cost than human workers. For example, a recent report estimates that up to 800 million jobs could be displaced by automation and AI technologies by 2030 (McKinsey Global Institute, 2017). However, there is a growing interest and excitement among professionals working in the business analytics domain toward ChatGPT and its potential applications. Many see it as a powerful tool that can help them extract valuable insights from large amounts of data and gain a deeper understanding of customer feedback and opinions. As such, while there are concerns about ChatGPT, there is also growing excitement and interest among professionals toward ChatGPT.

To better understand the emerging digital technology—ChatGPT, we investigate its impact on the business analytics discipline from users' perspectives. Particularly, we focused on two specific research questions:

1. What are the users' reactions to ChatGPT?
2. How do various users' reactions to ChatGPT influence its continuous development and management?

Related Literature

The study is grounded in the existing literature on digital innovation and ChatGPT and will contribute to understanding ChatGPT from users' perspectives. Research on digital innovation focuses on various themes, such as digital innovation management (Nambisan et al., 2017; Yoo et al., 2010, 2012), the environment and drivers of digital innovation (Gawer & Cusumano, 2014; LaValle et al., 2010), digital innovation transform the way that businesses operate (Westerman et al., 2014), digital innovation in business analytics (Brynjolfsson & McAfee, 2014; Davenport & Patil, 2012), and digital innovation on the job market (Autor, 2015; Brynjolfsson & McAfee, 2014; Frey & Osborne, 2017). In this study, we focused on the social impacts of digital innovation on the job market. In other words, we try to understand how digital innovation can transform the job market by creating new methods and opportunities for communication, resulting in various users' responses to these advanced innovations. Brynjolfsson & McAfee (2014) examines the impact of digital technologies on the labor market and argues that while these technologies have the potential to create new jobs and increase productivity, they are also likely to displace workers in specific industries. Frey & Osborne (2017) explores the potential impact of computerization on employment. He estimates that as many as 47% of jobs in the United States might be replaced by automation or AI in the future. Building upon the divergent impact of digital innovation on the job market, it is obvious that users hold both optimistic (e.g., excitement) and pessimistic (e.g., anxieties) attitudes toward digital innovation.

Considering ChatGPT is a unique digital innovation attracting tremendous attention in academia, we narrow our scope down to grasp the emerging knowledge of ChatGPT. The current literature on ChatGPT is mainly focused on its use and performance in various disciplines (Dowling & Lucey, 2023; Gilson et al., 2023; Jiao et al., 2023; Kung et al., 2023; Qin et al., 2023). For example, previous research has investigated how ChatGPT perform on medical licensure exam (Kung et al., 2023); ChatGPT can significantly assist with financial research (Dowling & Lucey, 2023), ChatGPT is supportive of information literacy and privacy literacy skills (Lund & Agbaji, 2023), and simplify data visualizing process (Maddigan & Susnjak, 2023). However, extant literature shares limited information and knowledge on ChatGPT from users' perspectives. For example, users may express either negative or positive emotions about using and adopting ChatGPT. Liebrenz et al. identified ethical issues raised by potentially misleading or inaccurate content (Liebrenz et al., 2023). Users' responses are essential for the continuous development and management of ChatGPT.

Combining the merits of the previous literature and our current understanding of this emerging phenomenon, it is necessary to investigate the emerging digital technology (i.e., ChatGPT) via an ignored perspective—users' responses to ChatGPT and add valuable knowledge to both the digital innovation and ChatGPT literature. Next, we propose our methodology, research design, and theoretical framework for this study.

Methodology and Design Research

Content analysis is a decent approach to analyze and interpret textual, audio, or visual data (Elo & Kyngäs, 2008; Krippendorff, 2018). It systematically codes and categorizes data to identify patterns, themes, and relationships (Graneheim & Lundman, 2004). This research method is a appropriate to investigate emerging digital technologies like Blockchain (Lohmer et al., 2022) and COVID-19 (Jiang et al., 2023). As ChatGPT is an innovation in artificial intelligence, we adopt the methodology of content analysis to explore themes and sentiments.

Methodology

We use online content analysis to investigate users' responses to the emerging digital technology—ChatGPT. Emerging digital technologies always raise important questions and concerns. For example, there may be concerns about the ethical and social implications of ChatGPT. In addition, emerging technologies may disrupt existing industries and employment patterns, leading to job losses and economic dislocation. As such, the responses of users should be thoroughly understood whenever an emerging digital innovation is released to the market. To better analyze the users' responses, online content analysis can be used to study a wide range of topics and research questions, from social media trends and online behaviors to political discourse and public opinion (Seale, 2017). Furthermore, online content analysis can provide access to large amounts of data that would be difficult or impossible to collect through other means. With the rise of social media and other digital platforms, there is an abundance of user-generated content available online, providing researchers with a rich source of data to analyze. By carefully analyzing online content related to users' responses to ChatGPT, we expect to gain insights into ChatGPT and contribute to the existing literature.

Preliminary Date Collection

We collected the data from a popular high-tech social media platform, a platform designed for people to discuss various topics, specifically in digital business. The reason we try to investigate this emerging topic—how ChatGPT will reform the business analytic discipline—on this platform is that this platform is a well-known social platform for people to discuss popular innovation and technology (Anderson, 2015). The targeted topics are narrowed down by using the keywords: "ChatGPT" and "Business Analytics", which will enable us to capture the most updated trend in this joint domain—business analytics and ChatGPT. We select the most relevant topics discussed on Reddit and present the collected data below.

Preliminary Results and Proposed Theoretical Framework

Based on the preliminary dataset, we identified five main categories that people are interested in discussing ChatGPT on this social media platform: ChatGPT functionality, job elimination, job creation, plagiarism, and misinformation. The majority of discussions are focused on ChatGPT functionality and Job Elimination. Specifically, ChatGPT functionality refers to the various features included in this artificial intelligence (AI) technology; job elimination indicates the potential threats of ChatGPT on the existing job positions (i.e., business analytic job positions); job creation means the opportunity of ChatGPT to generate new job positions in the business analytic domain; plagiarism points out the concern of people are cheating by using this AI technology; misinformation reveals the fact that people are not confident on the accuracy of the results provided by this AI technology (i.e., ChatGPT).

Furthermore, we found that people are expressing different emotions toward ChatGPT (i.e., negative and positive feelings). The negative emotion refers that people are not holding much confidence in the capacity of ChatGPT. Meanwhile, the positive emotion refers that people are optimistic about the capability of ChaptGPT.

Combining the merits of our understanding of the dataset and the established coding scheme, we propose a theoretical framework to better understand how the adoption level of this newly introduced AI technology (i.e., ChatGPT) on social media platforms impacts the development and management of ChatGPT. First, ChatGPT is an ongoing project. The continuous development and management of ChatGPT in various categories are essential and critical. On the one hand, the continuous development and management of ChatGPT are consistently constrained and improved by either negative or positive emotional expressions from the five identified categories: ChatGPT functionality, job elimination, job creation, plagiarism, and misinformation. On the other hand, the continuous development and management of ChatGPT will focus more on exploitations to incrementally mitigate the negative emotions expressed by critics, and simultaneously, ChatGPT will engage in more explorations to reinforce the positive emotions expressed by supporters.

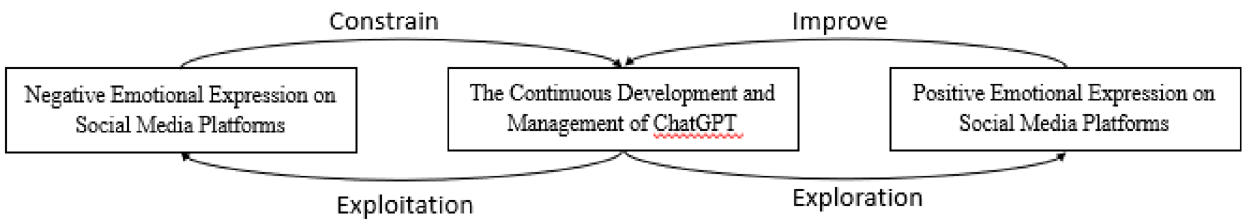


Figure 1. Potential Theoretical Model

CONCLUSION

This paper represents a work-in-progress, and we intend to conduct further analysis to produce a complete research paper in the near future. Our research has thus far indicated that the continuous development and management of ChatGPT in various categories are essential and critical and are constrained and improved by negative or positive emotional expressions from five identified categories. These categories are ChatGPT functionality, job elimination, job creation, plagiarism, and misinformation. The continuous development and management of ChatGPT can incrementally mitigate the negative emotions of critics and reinforce the positive onessentiments expressed by supporters. We expect to contribute to the existing knowledge on digital innovation and ChatGPT from the users' perspectives by illustrating the importance of users' responses on developing and managing emerging digital innovation—ChatGPT.

REFERENCE

Absolutdata. (2019, April 9). Unlocking Data Insights: Natural Language Processing Applications for Business and Marketing. *Absolutdata*. https://www.absolutdata.com/blog/unlocking-data-insights-natural-language-processing-applications-for-business-and-marketing-absolutdata/

Alshurafat, H. (2023). The Usefulness and Challenges of Chatbots for Accounting Professionals: Application On ChatGPT. *Available at SSRN 4345921*.

Anderson, K. E. (2015). Ask me anything: What is Reddit? *Library Hi Tech News*, *32*(5), 8–11.

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, *29*(3), 3–30.

Boudreau, K. (2010). Open platform strategies and innovation: Granting access vs. devolving control. *Management Science*, *56*(10), 1849–1872.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

Chen, C., & Jiang, H. (2020). Important skills for data scientists in china: Two Delphi studies. *Journal of Computer Information Systems*, *60*(3), 287–296.

Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard Business Review*, *90*(5), 70–76.

Dowling, M., & Lucey, B. (2023). ChatGPT for (Finance) research: The Bananarama Conjecture. *Finance Research Letters*, 103662. https://doi.org/10.1016/j.frl.2023.103662

Eesley, C. E., & Wu, L. (2019). For startups, adaptability and mentor network diversity can be pivotal: Evidence from a randomized experiment on a mooc platform. *MIS Quarterly, Forthcoming*.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*, 254–280.

Frieder, S., Pinchetti, L., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Chevalier, A., & Berner, J. (2023). Mathematical capabilities of chatgpt. *ArXiv Preprint ArXiv:2301.13867*.

Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. *Journal of Product Innovation Management*, *31*(3), 417–433.

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Medical Education*, *9*(1), e45312. https://doi.org/10.2196/45312

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*(2), 105–112.

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, *349*(6245), 261–266.

Introducing ChatGPT. (2022, November 30). https://openai.com/blog/chatgpt

Jiang, H., Castellanos, A., Castillo, A., Gomes, P. J., Li, J., & VanderMeer, D. (2023). Nurses' Work Concerns and Disenchantment During the COVID-19 Pandemic: Machine Learning Analysis of Web-Based Discussions. *JMIR Nursing*, *6*, e40676.

Jiao, W., Wang, W., Huang, J., Wang, X., & Tu, Z. (2023). Is ChatGPT a good translator? A preliminary study. *ArXiv Preprint ArXiv:2301.08745*.

Kitamura, F. C. (2023). ChatGPT is shaping the future of medical writing but still requires human judgment. In *Radiology* (p. 230171). Radiological Society of North America.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., & Maningo, J. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, *2*(2), e0000198.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2010). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*.

Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with ChatGPT: Ethical challenges for medical publishing. *The Lancet Digital Health*.

Lohmer, J., Ribeiro da Silva, E., & Lasch, R. (2022). Blockchain technology in operations & supply chain management: A content analysis. *Sustainability*, *14*(10), 6192.

Lund, B., & Agbaji, D. (2023). Information Literacy, Data Literacy, Privacy Literacy, and ChatGPT: Technology Literacies Align with Perspectives on Emerging Technology Adoption within Communities. *Data Literacy, Privacy Literacy, and ChatGPT: Technology Literacies Align with Perspectives on Emerging Technology Adoption within Communities (January 14, 2023)*.

Maddigan, P., & Susnjak, T. (2023). Chat2vis: Generating data visualisations via natural language using chatgpt, codex and gpt-3 large language models. *ArXiv Preprint ArXiv:2302.02094*.

Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital Innovation Management: Reinventing Innovation Management Research in a Digital World. *MIS Quarterly*, *41*(1), 223–238. https://doi.org/10.25300/MISQ/2017/41:1.03

Qin, C., Zhang, A., Zhang, Z., Chen, J., Yasunaga, M., & Yang, D. (2023). Is ChatGPT a General-Purpose Natural Language Processing Task Solver? *ArXiv Preprint ArXiv:2302.06476*.

Roose, K. (2022, December 5). The Brilliance and Weirdness of ChatGPT. *The New York Times*. https://www.nytimes.com/2022/12/05/technology/chatgpt-ai-twitter.html

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, *6*(1).

Seale, C. (2017). Researching society and culture. *Researching Society and Culture*, 1–664.

Sobania, D., Briesch, M., Hanna, C., & Petke, J. (2023). An analysis of the automatic bug fixing performance of chatgpt. *ArXiv Preprint ArXiv:2301.08653*.

Westerman, G., Bonnet, D., & McAfee, A. (2014). *Leading digital: Turning technology into business transformation*. Harvard Business Press.

Yoo, Y., Boland, R. J., Lyytinen, K., & Majchrzak, A. (2012). Organizing for Innovation in the Digitized World. *Organization Science*, *23*(5), 1398–1408. https://doi.org/10.1287/orsc.1120.0771

Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). **Research Commentary** —The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. *Information Systems Research*, *21*(4), 724–735. https://doi.org/10.1287/isre.1100.0322

Yue, L., Chen, W., Li, X., Zuo, W., & Yin, M. (2019). A survey of sentiment analysis in social media. *Knowledge and Information Systems*, *60*, 617–663.

66. Topic Mining-based Paper Abstract Analysis for Block Chain in Healthcare

Yangyang Tao [1], Junxiu Zhou [1], Haoqiang Jiang [1], and Juanjuan Li [2], Zhuojun Duan [3]

[1] Northern Kentucky University, Highland Heights KY 41099, USA
[2] Florida International University, 11200 SW 8th St, Miami, FL 33199, USA
[3] James Madison University, Harrisonburg VA 22807, USA
{ taoy1, zhouj2, jiangh1}@nku.edu, jli071@fiu.edu, duanzx@jmu.edu

Abstract.

Blockchain technology has gained significant attention due to its decentralized and secure nature. As the number of research papers on blockchain in the healthcare domain continues to grow, it becomes increasingly challenging to extract meaningful insights from them. Topic mining, a technique for identifying key topics in a collection of texts, can help in this regard. In this paper, we attempt to apply text mining to extract valuable information about blockchain from abstract of existing literature. Our method involves collecting blockchain in health-related abstracts and preprocessing the abstracts to remove stop words and tokenize the remaining words. We then use Latent Dirichlet Allocation (LDA) to identify the most relevant topics in the abstracts. We evaluate our approach on a dataset of blockchain research papers and demonstrate its effectiveness in identifying the key topics and subtopics in the field. Our findings can be useful for researchers and practitioners in understanding the current state of the art in blockchain research and identifying future research directions. The next task is to involve human experts to conduct a more thorough examination of each topic and produce narrative summaries based on the current findings.

Keywords: Text Mining, Topic Modeling, Block Chain, Healthcare.

1 Introduction

Blockchain is a pioneering technology that originated from Bitcoin (Nakamoto, 2008). It employs digital ledgers comprised of blocks to record and monitor transactions between various users or systems. The information stored in blocks is hashed to ensure data integrity and prevent unauthorized modifications. The decentralized, immutable, and transparent nature of blockchain makes it a promising platform for various applications such as finance (Nguyen, 2016), supply chain (Azzi et al., 2019), accounting (Dai & Vasarhelyi, 2017), and healthcare (Hölbl et al., 2018).

As Blockchain is increasingly implemented across various industries (Casino et al., 2019), researchers have provided systematic literature reviews in different domains including the Internet of Things (Conoscenti et al., 2016), FinTech (Rabbani et al., 2020), and supply chain (Queiroz et al., 2020). Recently, academics have turned to text mining and topic modeling to explore the potential of blockchain in areas such as consumer trust (da Silva & Moro, 2021), cybersecurity (Prakash et al., 2022), and Bioeconomy (Willrich et al., 2019). However, despite existing literature summarizing blockchain applications in healthcare (Tandon et al., 2020), limited studies focus on using natural language processing techniques such as topic modeling to extract meaningful insights from existing blockchain studies in the healthcare domain.

In this research in progress, we gathered data resources by collecting abstracts related to blockchain in the health industry. Inspired by previous study (Brown et al., 2019), we use keywords "(blockchain* OR "block chain") AND (*health* OR *medic* OR patient OR hospital OR EMR OR EHR OR disease OR diagnosis OR client OR drug OR doctor OR physician)" to retrieve approximately 5,710 abstracts from various databases such as PubMed, IEEE Xplore, Web of Science, and Medline EBSCO. After removing duplicate and missing abstracts, we get 4,532 valid abstracts as our dataset. We preprocessed the abstracts to remove stop words and tokenize the remaining words, and then use Latent Dirichlet Allocation (LDA) to identify the most relevant topics in the abstracts. Our results showed that the topics generated by the LDA uncovered topics from different angles of blockchain implementation in healthcare, such as blockchain applications in Telemedicine, patient data privacy, and transmission of IOT devices. Our next step is to use human-in-the-loop to further explore each topic deeper and provide narrative summaries based on the current results (Jiang et al., 2023).

1 Related Work

1.1 Block Chain

Blockchain technology (Zheng et al., 2018) has gained significant attention in recent years, leading to a proliferation of research in this area. Several studies have focused on the technical aspects of blockchain, including consensus algorithms, security, and scalability. For instance, the original Bitcoin paper (Nakamoto, 2008) introduced the concept of proof-of-work consensus (Szalachowski et al., 2019), which has since been extended to other blockchain platforms. Other studies have focused on improving blockchain scalability, such as the Lightning Network (Poon & Dryja, 2016) and sharing techniques.

In addition to technical aspects, researchers have also explored the various applications of blockchain in various domains, including finance, healthcare, and supply chain management. Moreover, researchers have explored the social and economic implications of blockchain, such as its impact on privacy, governance, and regulation. Some studies have highlighted the potential for blockchain to empower individuals and communities (Thomason et al., 2018), while others have raised concerns about its potential for centralization and concentration of power (Sai et al., 2021). This work focuses on exploring the application of blockchain in the healthcare field. Specifically, we use topic modeling technology of text mining to analyze literature on blockchain in healthcare in order to have a comprehensive understanding of the current research status.

1.2 Topic Modeling

Topic modeling is a popular research area in natural language processing and machine learning and has been applied in a wide range of domains. Latent Dirichlet Allocation (LDA) is a widely used algorithm for topic modeling (Blei et al., 2003), which has been used in many studies to extract meaningful topics from text corpora. Many studies have focused on improving the accuracy and efficiency of LDA and other topic modeling algorithms, such as Gibbs sampling and variational inference. Other studies have explored the application of topic modeling in various domains, including social media analysis, sentiment analysis, and news categorization.

Moreover, researchers have explored the combination of topic modeling with other techniques, such as deep learning and graph analysis, to improve the quality of topic extraction and the interpretation of topics. Some studies have also investigated the evaluation of topic models, including measures of coherence and interpretability. Additionally, researchers have examined the applications of topic modeling in practical settings, such as recommender systems and personalized search (Brownfield & Zhou, 2020). Overall, the related work in topic modeling research is extensive and reflects the versatility and importance of this technique in various fields. Therefore, this work leverages the LDA method to discover keywords to help us capture insights of current research work on blockchain in healthcare domain.

2 Methodology

The proposed work follows a framework illustrated in Fig. 1, which comprises three stages: dataset cleaning, topic modeling, and visualization and analysis of the results.

2.1 Dataset Cleaning

In order to prepare the dataset for topic modeling, we first removed noisy abstracts and any erroneous content that contained illegal coding. Then do the preprocessing to purify the dataset. The abstract cleaning and preprocess aim to clean the dataset so that it can fit into the topic modeling step. The resulting cleaned dataset was then visualized using a word cloud, as shown in Fig. 2.

In the data cleaning process, firstly, all capitalized letters are converted to lowercase, and any punctuation, brackets, and words containing numbers are removed. After removing noise abstracts and erroneous content, the next step in preparing the dataset for topic modeling is lemmatization. This process involves mapping each word to its base root form, which helps group words with similar meanings together. To further focus on relevant words, stop words such as "background," "aim," "we," "however," "result," and "a" are eliminated. Once this is done, the abstracts are tokenized by splitting them into individual words, and any tokens that appear in fewer than 20 documents or more than 20% of the documents are removed. These additional preprocessing steps enable the efficient application of topic modeling techniques to the resulting corpus.

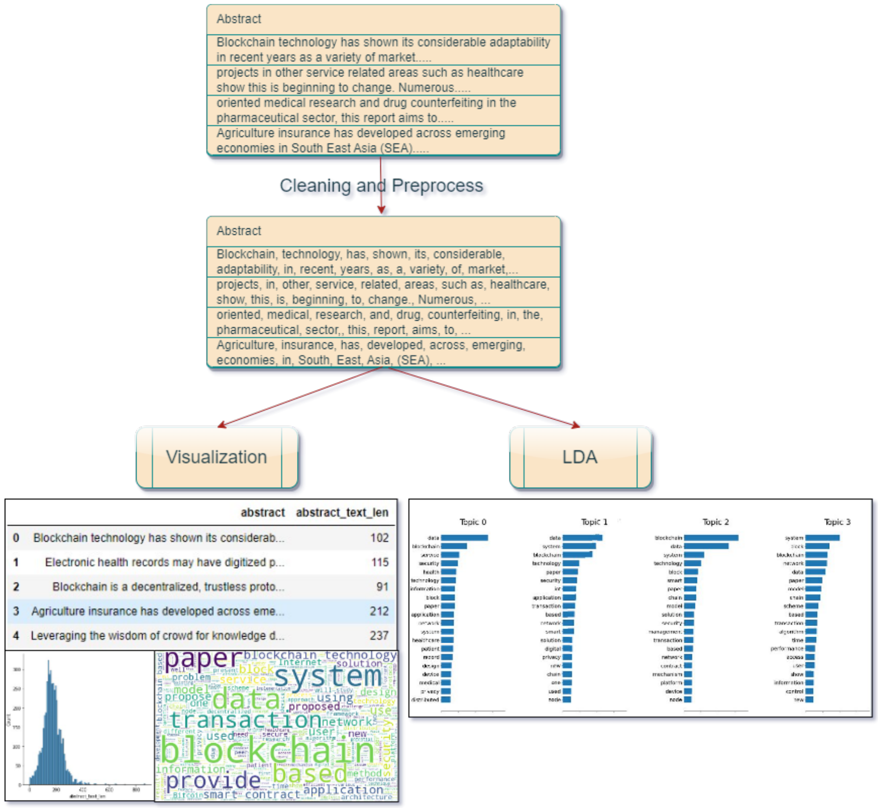


Fig. 1. The framework of the proposed method.

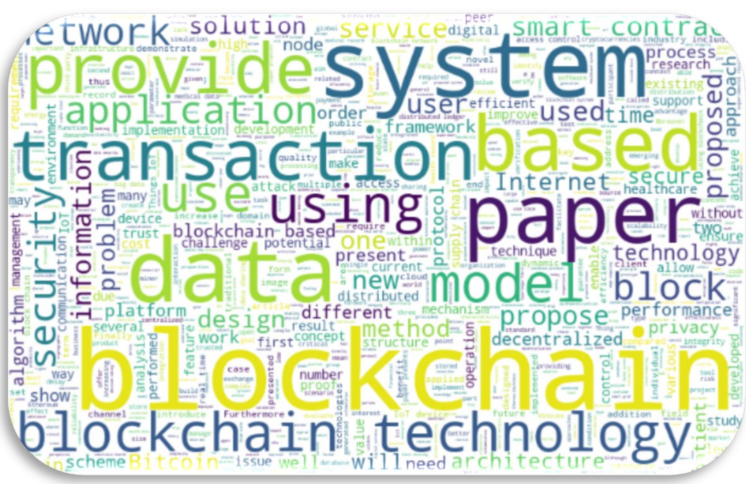


Fig. 2. Word clouds of abstracts in the dataset.

2.2 LDA

Latent Dirichlet Allocation (LDA) (Blei et al., 2003) is a widely used unsupervised machine learning algorithm that is commonly used for topic modeling. It is a generative probabilistic model that assumes that there is an underlying, unobserved probability distribution that generates the observed data. The goal of LDA is to identify the latent variables that govern this distribution. Specifically, LDA aims to identify the distribution of topics in a corpus and the distribution of words within each topic. These distributions are represented as probability distributions over a fixed vocabulary.

The LDA algorithm works as follows. First, it randomly assigns each word in the corpus to one of K topics. Next, for each document in the corpus, LDA generates a topic distribution by randomly sampling a mixture of K topics. Finally, for each word in each document, LDA generates a word distribution by randomly selecting a word from the assigned topic. During training, the algorithm updates the assignment of topics to words based on the likelihood of the observed data. This update involves iteratively estimating the posterior distribution of the latent variables given the observed data, and then re-estimating the parameters of the model using this distribution. This process continues until convergence is achieved. LDA analyzes each document d to determine the probability p of each topic t given the words w in the document, denoted as p(topic t | document d). The algorithm also calculates the probability of each word w being associated with topic t, represented as p(word w | topic t), by examining the frequency of topic assignments to that word across all documents. These probabilities are then utilized to update the probability of the word w being associated with topic t, indicated by p(word w with topic t) = p(topic t | document d) * p(word w | topic t). This process is iterated until the algorithm converges to a steady state where topic assignments no longer change. Finally, based on the final topic assignments, the proportion of topics in each document can be determined.

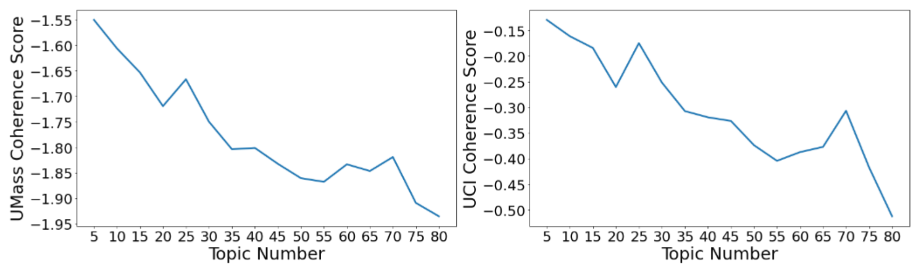
3 Experimental Results

This section demonstrates the LDA topic modeling results on the collected dataset to effectively provide a literature overview of blockchain and health-related research. The experiments were performed on a laptop equipped with an Intel(R) i7-Core CPU, using the Python programming language and relevant packages, including sklearn, genism, and wordcloud.

3.1 Topic Generation

The aim of the topic modeling method is to identify the themes that are present within a given corpus. These themes are represented as the top N words that are most likely to be associated with a specific topic. However, since the LDA method is an unsupervised topic modeling approach, determining the optimal number of topics can be challenging. To address this, researchers use coherence score measurements (Zhou et al., 2023), which assess how similar the words in a particular topic are to each other. Evaluating the interpretability of topics generated by a topic modeling model can be accomplished using coherence scores, which are metrics that gauge the coherence of the topics from a human interpretation standpoint.

Similar to previous work (Zhou et al., 2023), we tested different topic numbers using LDA (5, 10, 15, and increments of five up to 80 and calculated two coherence scores, i.e., UMass coherence score and UCI coherence score, for each topic number. Based on Fig. 3, the optimal topic numbers for UMass and UCI coherence scores are 5. However, to obtain a better understanding of the literature, we also considered the top four coherence score-based topic numbers (5, 10, 15, and 25) and applied these four different topic numbers to identify the most appropriate number of topics.



UMass score of LDA UCI score of LDA

Fig. 3. Different coherence scores of LDA with the optimal number of topics

3.2 Topic Analysis

The aim of this part is to examine the topics produced by LDA and assess their ability to capture relevant information from the literature. To achieve this, we used coherence scores to identify the four optimal topic numbers for generating topics.

We displayed the topics generated by LDA using topic numbers 5, 10, 15, and 25 from Table 1 to Table 4. These tables show the top 15 words of topics discovered by LDA under different topic numbers. Each topic is comprised of a list of words, with the order of the words indicating their relative importance within the topic. In general, the more prominent a word is in the list, the more significant it is to the topic. Note that, due to space limitations, only the top five popular topics are listed under different topic numbers. Topic number starts from 0 in labeling.

Table 1. Topics Generated by LDA with topic number = 5

Topic ID	**weights and related words of each topic**
Topic 3	'system', 'block', 'blockchain', 'network', 'data', 'paper', 'model', 'chain', 'scheme', 'based', 'transaction', 'algorithm', 'time', 'performance', 'access'
Topic 4	'blockchain', 'technology', 'paper', 'application', 'smart', 'contract', 'system', 'security', 'transaction', 'network', 'based', 'information', 'data', 'service', 'bitcoin'
Topic 1	'data', 'system', 'blockchain', 'technology', 'paper', 'security', 'iot', 'application', 'transaction', 'based', 'network', 'smart', 'solution', 'digital', 'privacy'

Topic 2	'blockchain', 'data', 'system', 'technology', 'block', 'smart', 'paper', 'chain', 'model', 'solution', 'security', 'management', 'transaction', 'based', 'network'
Topic 0	'data', 'blockchain', 'service', 'security', 'health', 'technology', 'information', 'block', 'paper', 'application', 'network', 'system', 'healthcare', 'patient', 'record'

Table 2. Top five Topics Generated by LDA with topic number = 10

Topic ID	weights and related words of each topic
Topic 3	'system', 'block', 'network', 'chain', 'blockchain', 'paper', 'model', 'data', 'algorithm', 'scheme', 'based', 'performance', 'transaction', 'time', 'user'
Topic 4	'blockchain', 'technology', 'paper', 'application', 'smart', 'security', 'contract', 'research', 'industry', 'based', 'system', 'transaction', 'new', 'information', 'healthcare'
Topic 5	'data', 'system', 'blockchain', 'medical', 'block', 'image', 'cloud', 'technology', 'chain', 'service', 'information', 'network', 'model', 'paper', 'scheme'
Topic 7	'blockchain', 'iot', 'data', 'transaction', 'device', 'network', 'paper', 'technology', 'privacy', 'access', 'smart', 'security', 'user', 'framework', 'decentralized'
Topic 9	'blockchain', 'system', 'technology', 'data', 'bitcoin', 'network', 'application', 'security', 'smart', 'service', 'block', 'paper', 'information', 'transaction', 'based'

Table 3. Top five Topics Generated by LDA with topic number = 15

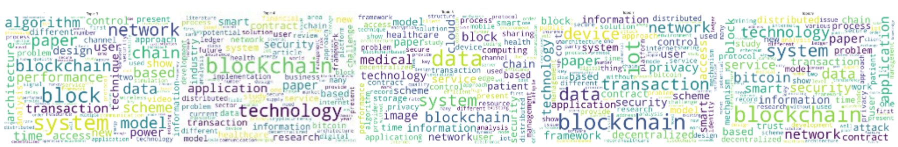
Topic ID	weights and related words of each topic
Topic 3	'system', 'network', 'block', 'blockchain', 'model', 'paper', 'data', 'based', 'user', 'transaction', 'chain', 'algorithm', 'access', 'time', 'power'
Topic 5	'data', 'system', 'medical', 'blockchain', 'cloud', 'technology', 'patient', 'service', 'image', 'healthcare', 'information', 'model', 'network', 'block', 'security'
Topic 9	'blockchain', 'system', 'data', 'technology', 'bitcoin', 'application', 'service', 'smart', 'security', 'network', 'information', 'contract', 'transaction', 'paper', 'based'
Topic 12	'block', 'data', 'scheme', 'chain', 'blockchain', 'paper', 'system', 'network', 'performance', 'technique', 'information', 'security', 'image', 'based', 'key'
Topic 13	'blockchain', 'technology', 'application', 'paper', 'research', 'security', 'smart', 'network', 'data', 'service', 'study', 'challenge', 'bitcoin', 'transaction', 'new'

Table 4. Top five Topics Generated by LDA with topic number = 25

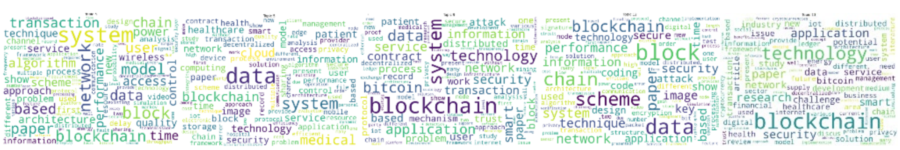
Topic ID	weights and related words of each topic
Topic 5	'data', 'system', 'medical', 'blockchain', 'cloud', 'service', 'technology', 'patient', 'healthcare', 'image', 'information', 'model', 'network', 'time', 'security'
Topic 12	'data', 'scheme', 'block', 'chain', 'blockchain', 'paper', 'image', 'technique', 'information', 'system', 'security', 'network', 'performance', 'coding', 'attack'
Topic 13	'blockchain', 'technology', 'application', 'research', 'paper', 'security', 'smart', 'network', 'bitcoin', 'data', 'service', 'industry', 'challenge', 'study', 'article'
Topic 15	'data', 'block', 'user', 'chain', 'blockchain', 'system', 'paper', 'transaction', 'scheme', 'structure', 'based', 'technology', 'design', 'code', 'channel'
Topic 20	'blockchain', 'transaction', 'system', 'technology', 'paper', 'protocol', 'new', 'security', 'application', 'bitcoin', 'consensus', 'distributed', 'used', 'network', 'algorithm'

As observed, the identified topics effectively showcase the prominent keywords across all the abstracts. As the number of topics increases, more specific keywords are revealed. To gain a better understanding of the most important terms in each topic, we utilized word clouds to depict the top five prevalent topics for four different topic numbers. Word clouds visually represent text data by displaying the first 100 words of a topic in Fig. 4, based on their probability within the related topic.

(a) topic number = 5 (Topic_IDs 3, 4, 1, 2, 0)



(b) topic number = 10 (Topic_IDs 3, 4, 5, 7, 9)



(c) topic number = 15 (Topic_IDs 3, 5, 9, 12, 13)

(d) topic number = 25 (Topic_IDs 5, 12, 13, 15, 20)

Fig. 4. Word clouds of the top 5 topics in descending order by LDA method with different topic numbers.

By utilizing the topics identified through topic modeling methods, we can annotate each abstract record in the collected dataset with relevant topics. Table 5 illustrates how each abstract consists of multiple topics, with each topic associated with a specific probability value. However, in most cases, only one topic is dominant in an abstract. As a result, we can assign each abstract to the dominant topic by selecting the topic with the highest probability value.

For instance, the Abstract_14 talks about Telemedicine under the blockchain technology. It has larger chance to be classified to one specific dominant topic [0, 5, 5, 5] under topic numbers of [5, 10, 15, 25]. Moreover, those topics cover similar keywords, such as, 'data', 'system', 'blockchain', 'technology', 'information', 'service', etc. However, with the increase of topic number, more detailed keywords jump out such as 'time', 'image', 'medical'. As we can see, LDA is able to keep consistent in generating topics even under different topic numbers. While Abstract_22 mainly focuses on the keywords related to 'iot', 'system', 'blockchain', 'technology', 'device', etc. As can be observed, with the increase of topic number, this topic may have decreased probability belongs to one specific topic. When topic number equals to 5, it has 0.521 probability belongs to Topic_ID 1, while it has 0.3506 probability belongs to Topic_ID_24 when topic number equals to 25.

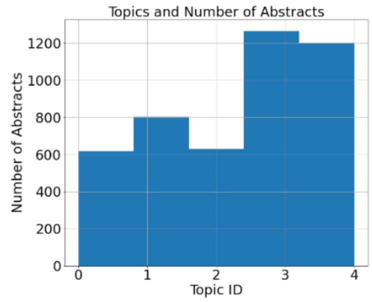
Abstract_14: "*Telemedicine offers a medical-on-demand (MoD) service from a distance …… We construct an independent-update key policy ABE scheme in the distributed telemedicine system that aims to updates patient's keys separately, and there are multiple authorities to manage this system altogether which is more similar to the real situation. On the other hand, by using blockchain and distributed database technologies, the private healthcare data stored in public cloud is protected in integrity, ……*"

Abstract_22: "*Blockchain is an emerging technology that uses distributed ledgers for transparent, reliable, and traceable information exchange among network nodes. …… In this paper, we propose an optimized policy for sampling rate by IoT sensors that utilize blockchain and Tangle technologies for their transmission with the goal of minimizing the age of information (AoI) experienced by the end-users, considering both processing and networking resource constraints……* "

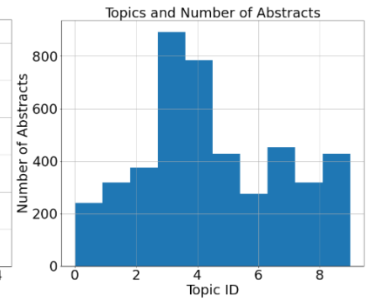
Table 5. Examples of abstracts and related topics distribution

	Topic number	Top 1 (Topic_ID, Probability)	Top 2 (Topic_ID, Probability)	Top 3 (Topic_ID, Probability)
Abstract_14	5	(0, 0.9725)	(4, 0.001)	(2, 0.000)
	10	(5, 0.993)	(0, 0.001)	(4, 0.001)
	15	(5, 0.8697)	(11, 0.124)	(2, 0.000)
	25	(5, 0.8101)	(11, 0.1828)	(1, 0.0005)
Abstract_22	5	(1, 0.521)	(3, 0.3358)	(4, 0.1379)
	10	(7, 0.728)	(3, 0.2186)	(4, 0.0435)
	15	(7, 0.5234)	(3, 0.2837)	(1, 0.1820)

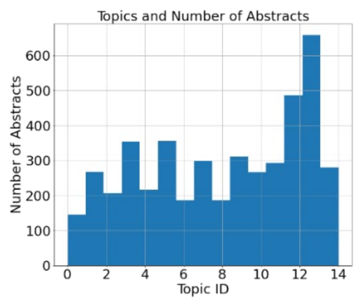
	25	(24, 0.3506)	(7, 0.2733)	(16, 0.2279)



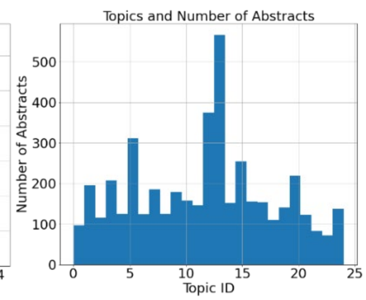
(a) topic number = 5



(b) topic number = 10



(c) topic number = 15



(d) topic number = 25

Fig. 5. Histogram of abstract and dominant topic distribution.

Next, we assigned each abstract in the dataset to its corresponding dominant topic. The resulting distribution of topics within the dataset is presented in Fig. 5.

Figure 5(a) shows that the topics generated by LDA are divided over all 5 topics. By referring to the corresponding topic keywords in Table 1, we can see that approximately 13.33% of the abstracts in the dataset are related to 'health', 'healthcare', and 'patient' associated with blockchain in Topic_0. Topic_3, which concerns 'algorithm', 'time', and 'performance', has about 26.67% of the abstracts. Meanwhile, Topic_2 focuses on 'smart', 'paper', and 'management'. Topic_1 is related to 'digital', 'privacy', and Topic_4 is associated with 'smart', 'contract', and 'bitcoin'. Abstracts are evenly distributed across these topics. Fig. 5(b) shows that topics 3, 4, 5, 7, and 9 cover approximately 3,000 abstracts in the dataset over 10 topics, comprising about two-thirds of all abstracts. We see that most of the collected abstracts are focused on “system”, “network”, “model” “data”, “technology”, “smart”, and “security” perspectives of blockchain in healthcare. As can be observed from Fig. 5(a) and Fig. 5(b), the LDA method can uncover several main trends from the blockchain area, including the performance of algorithms, healthcare and medical applications, technology, and Bitcoin.

Figure 5(c) illustrates the distribution of fifteen topics generated by the LDA method. We can see that Topic_13 has many abstracts compared to other topics, and it focuses on “technology”, “application”, “security” and “smart”. Topic_3 is about “model”, “data”, and “performance”, Topic_5 concerns “medical” and “patient”, and “cloud”, and Topic_9 is about bitcoin. Topic_12 is related to system performance. Therefore, with topic number equals to 15, more detailed concerns about blockchain in healthcare are revealed. Fig. 5(d) shows the distribution of topics generated by the LDA method with a topic number of 25. The top three topics are Topic_5, Topic_12, and Topic_13, covering system security, Bitcoin, and healthcare, respectively. Topics_15 and Topic_20 cover the performance of both systems and algorithms. The results presented in Fig. 5(c) and Fig. 5(d) reveal that the majority of blockchain-related research focuses on system security, Bitcoin, and healthcare.

4 Conclusion

This work applied LDA, a text-mining method, to provide a comprehensive analysis of published articles on the blockchain. The results of LDA were thoroughly examined to extract the trends in blockchain-related research. Our findings indicate that the LDA method was effective in identifying various topics from different angles. The discovered topics reveal that blockchain research is focused on system data, model, performance, security, application, bitcoin, and patient. This study provides valuable insights into the current state of blockchain research, which could be useful for researchers and practitioners in the field. We plan to utilize human-in-the-loop methods to delve deeper into each identified topic and generate narrative summaries based on the current findings. By involving domain experts and utilizing their knowledge and expertise, we aim to provide a more comprehensive understanding of the topics discovered through the topic modeling process. This approach will allow us to validate and refine our results and gain new insights into the application of blockchain in healthcare.

References

Azzi, R., Chamoun, R. K., & Sokhn, M. (2019). The power of a blockchain-based supply chain. *Computers & Industrial Engineering*, *135*, 582–592.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Brown, E. L., Ruggiano, N., Li, J., Clarke, P. J., Kay, E. S., & Hristidis, V. (2019). Smartphone-based health technologies for dementia care: Opportunities, challenges, and current practices. *Journal of Applied Gerontology*, *38*(1), 73–91.

Brownfield, S., & Zhou, J. (2020). Sentiment analysis of Amazon product reviews. *Software Engineering Perspectives in Intelligent Systems: Proceedings of 4th Computational Methods in Systems and Software 2020, Vol. 2 4*, 739–750.

Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. *Telematics and Informatics*, *36*, 55–81.

Conoscenti, M., Vetro, A., & De Martin, J. C. (2016). Blockchain for the Internet of Things: A systematic literature review. *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, 1–6.

da Silva, C. F., & Moro, S. (2021). Blockchain technology as an enabler of consumer trust: A text mining literature analysis. *Telematics and Informatics*, *60*, 101593.

Dai, J., & Vasarhelyi, M. A. (2017). Toward blockchain-based accounting and assurance. *Journal of Information Systems*, *31*(3), 5–21.

Hölbl, M., Kompara, M., Kamišalić, A., & Nemec Zlatolas, L. (2018). A systematic review of the use of blockchain in healthcare. *Symmetry*, *10*(10), 470.

Jiang, H., Castellanos, A., Castillo, A., Gomes, P. J., Li, J., & VanderMeer, D. (2023). Nurses' Work Concerns and Disenchantment During the COVID-19 Pandemic: Machine Learning Analysis of Web-Based Discussions. *JMIR Nursing*, *6*, e40676.

Nakamoto, S. (2008). Bitcoin whitepaper. *URL: Https://Bitcoin. Org/Bitcoin. Pdf-(: 17.07. 2019)*.

Nguyen, Q. K. (2016). Blockchain-a financial technology for future sustainable development. *2016 3rd International Conference on Green Technology and Sustainable Development (GTSD)*, 51–54.

Poon, J., & Dryja, T. (2016). *The bitcoin lightning network: Scalable off-chain instant payments*.

Prakash, R., Anoop, V. S., & Asharaf, S. (2022). Blockchain technology for cybersecurity: A text mining literature analysis. *International Journal of Information Management Data Insights*, *2*(2), 100112.

Queiroz, M. M., Telles, R., & Bonilla, S. H. (2020). Blockchain and supply chain management integration: A systematic review of the literature. *Supply Chain Management: An International Journal*, *25*(2), 241–254.

Rabbani, M. R., Khan, S., & Thalassinos, E. I. (2020). *FinTech, blockchain and Islamic finance: An extensive literature review*.

Sai, A. R., Buckley, J., Fitzgerald, B., & Le Gear, A. (2021). Taxonomy of centralization in public blockchain systems: A systematic literature review. *Information Processing & Management*, *58*(4), 102584.

Szalachowski, P., Reijsbergen, D., Homoliak, I., & Sun, S. (2019). StrongChain: Transparent and Collaborative Proof-of-Work Consensus. *USENIX Security Symposium*, 819–836.

Tandon, A., Dhir, A., Islam, A. N., & Mäntymäki, M. (2020). Blockchain in healthcare: A systematic literature review, synthesizing framework and future research agenda. *Computers in Industry*, *122*, 103290.
Thomason, J., Ahmad, M., Bronder, P., Hoyt, E., Pocock, S., Bouteloupe, J., Donaghy, K., Huysman, D., Willenberg, T., & Joakim, B. (2018). Blockchain—Powering and empowering the poor in developing countries. In *Transforming climate finance and green investment with blockchains* (pp. 137–152). Elsevier.
Willrich, S., Melcher, F., Straub, T., & Weinhardt, C. (2019). Towards More Sustainability: A Literature Review Where Bioeconomy Meets Blockchain. *ICETE (1)*, 113–120.
Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services*, *14*(4), 352. https://doi.org/10.1504/IJWGS.2018.095647
Zhou, J., Tao, Y., Duan, Z., & Gautam, S. (2023). Topic Mining Based Literature Analysis for Face Recognition. *Data Science and Algorithms in Systems: Proceedings of 6th Computational Methods in Systems and Software 2022, Vol. 2*, 807–818.

69.

Effect of Cultural Diversity on Organizational Performance: A Moderating Effect of Remote Work Environment

OMAR FARUQ OSAMA

Emporia State University, Emporia, USA

Abstract

Nowadays, individuals from all over the world can easily find employment in any industry, anywhere in the world. This allows businesses to benefit from a wider range of cultural perspectives. The goal of this research is to examine how distant work situations mitigate the effects of cultural diversity on employee productivity. The study will use structured questionnaires to collect information from 200 eligible individuals at 50 corporate headquarters of well-known U.S. companies. As part of the technique, we will consider about using Structural Equation Modeling (SEM) and a confirmatory factor analysis. The results will complete the gaps in the literature on cultural diversity by showing how cultural conflict caused by cultural diversity affects organizational performance in terms of employee retention, attitude towards work, and output, and thus provide a valuable contribution to the study of how remote work environments are implemented.

Categories

Others

70. Data Mining using Logistic Regression with VALORANT Esports Data

Jacob Bishop

Undergraduate Student

Miller College of Business

Ball State University
jbishop@bsu.edu

Abstract. This paper explores the field of data mining, which involves extracting insights and knowledge from large datasets. Data mining is a fast-growing field creating value in the majority of industries, including esports, or competitive video games. Because this industry is new and generates lots of data, there is untapped potential to reveal valuable insights with data mining techniques. First, there is an examination of data mining techniques and their applications in esports. Then, the creative project focuses on the challenges involved in data preprocessing, feature selection, and algorithm selection, and applies these discussions to the data mining project with real esports data. Then, logistic regression models and data visualizations are created and evaluated, completing the data mining workflow. Finally, a brief conclusion discusses the implications of the project's results and gives anecdotes about data mining in esports. Overall, the paper contributes to the exploration of data mining applications in esports.

Introduction

VALORANT is a free-to-play A 5v5 character-based tactical first-person shooter (FPS) game developed and published by Riot Games, the company behind the popular Multiplayer Online Battle Arena (MOBA) game League of Legends. VALORANT was released in June 2020 and quickly gained popularity as an esports title. It is a direct competitor to Valve Corporation's Counter-Strike: Global Offensive. Riot Games and Valve Corporation are two of an entire industry of companies that have long been competitors as developers of esport titles (Roach & Caswell, 2021).

In VALORANT, two teams of five players each compete against each other in a series of rounds. One team plays as the attacking side, while the other team plays as the defending side.

The attacking team must plant a bomb (called the "spike") at one of the two to three designated sites on the map, while the defending team tries to prevent them from doing so. If the spike is successfully planted, the defending team must defuse it before it detonates. The winner of the match is the team who gets to 13 rounds won first and wins by two rounds. If two complete halves are played and the end score is 12-12, then there will be a series of overtimes where each team plays one round as the attackers and the other as the defenders. This will repeat until one team wins by two rounds (Landa, 2020).

The game features a variety of different characters, called agents, each with their own unique abilities and play styles. Players choose which agent to play as at the beginning of each match, and no two players on the same team can play the same agent. Each agent is classified as either a duelist, an initiator, a sentinel, or a controller (Roach & Caswell, 2021). The different agent combinations allow each team to formulate an infinite amount unique strategies to win rounds and therefore the match.

VALORANT esports tournaments typically follow a best-of-three or best-of-five format. Tournaments can feature both online and offline (meaning Local Area Network, or LAN) play, with top teams competing for prize pools that can range from thousands to over one million dollars. Riot hosts international tournaments with millions in prize money twice a year. They have various smaller tournaments where winners qualify for these. This Riot-sponsored system is called the VALORANT Champions Tour, or VCT for short (Webster, 2020).

VALORANT has quickly become one of the most popular esports titles, with a growing competitive scene and dedicated player base. Its combination of tactical gameplay, diverse character abilities, and high skill ceiling make it an exciting game to watch and play. The heart of the community can be found in social media, where content creators, professional players, casual players, commentators, analysts, coaches, Riot employees, developers, and more all interact on social platforms like Twitch, YouTube, Twitter, Instagram, TikTok, and more. Riot is also adding content regularly to keep interests high with new characters, new maps, and various game balances. Riot has balanced supporting the competitive and casual scenes of the game, offering diverse ways to play and connect with the community (Khosla, 2021).

One contribution this paper offers is an in-depth example of using the data mining workflow pattern with VALORANT data, which is such a new esport there is little literature written on topic, and almost none in the field of data mining. Another contribution is the example utilizing nested object data to derive multiple datasets from different perspectives, allowing more possibilities and opportunities for data mining and data visualization. This contribution is especially useful when there are few match samples to work with, but the matches themselves have significant data.

Related Work

It was initially difficult to find academic literature focused on the esports industry because it is so new and evolving so quickly. Esports have seen a rapid rise in popularity in recent years due to a variety of factors. First, the advancements in technology have allowed for more engaging and immersive gaming experiences, which have attracted a larger and more global audience. Second, the increasing accessibility of high-speed Internet and streaming platforms has made it easier for fans to watch and follow their favorite teams, players, and content creators. Third, the growing acceptance and recognition of esports as a legitimate form of competition has led to increased investment from sponsors and advertisers. Finally, COVID-19 has accelerated growth by pushing people to seek alternatives to real-life activities, including esports. In the last 10 years, the esports industry has transformed from being a niche interest into becoming a mainstream form of entertainment, with millions of fans tuning in to watch live tournaments and competitions all around the globe.

The author of *Predicting the Outcome of NFL Games Using Logistic Regression* used the closest methodology to this paper, but the industry is slightly different. They used NFL game data to predict the outcome of games. They chose to build a model for each team in the NFL and used feature generation to increase performance and reduce multicollinearity. Using feature generation, the author was able to look for nonlinear relationships and generated over 500 features to test in different combinations for optimal performance. Despite having so many features, the author concluded around 20 features to be the "sweet spot" for high accuracy. The author was worried about overfitting the data because they had 333 games worth of data and used around 300 (~90%) of those for training. The results varied by team, with the lowest having about 25% accuracy and the highest having about 75% accuracy. The models which performed well tended to be on the teams which performed either consistently well or consistently poorly. The results were not optimal, so the author gave recommendations for other data mining techniques which could be used instead for predictions similar to this, such as k-nearest neighbors and neural networks (Bouzianis, 2019).

The authors of *Performance of Machine Learning Algorithms in Predicting Game Outcome from Drafts in Dota 2* used various machine learning techniques to predict game outcomes in the esport Dota 2, a Multiplayer Online Battle Arena (MOBA) game developed and published by Valve Corporation. Dota 2 is a hugely different game from VALORANT because there are over 120 characters vs. 22 in VALORANT, and each character in Dota 2 can buy up to 10 of 208 available items in the store. This creates endless combinations to consider and therefore endless more potential features for prediction. This author had success in using an API because Valve is more open to sharing data publicly than Riot Games. They cited K-Nearest Neighbor (kNN), Principal Component Analysis (PCA), Forward Stepwise Regression, random forest regression, and logistic regression with and without additional helper algorithms. The dataset used was an astounding 5 million matches collected via the API mentioned earlier. The author used sophisticated data science techniques for data preprocessing, such as classifying character drafts as 2 x N binary vectors and factorization machine preparation. Then they used Naïve Bayes classifier, logistic regression, factorization machines, and gradient boosting of decision trees. Another important note is the models worked better for lower skilled players, which is great news since the data collected for VALORANT was entirely matches with below average skilled players (Semenov et al., 2018).

The authors of *Real-time eSports Match Result Prediction* wanted to take a different approach to predicting the outcomes of esports matches: real-time prediction of who will win a match. Again using Dota 2 as the esport to focus on, they gathered data using dota2api, an open source project on GitHub, to crawl more than 75 thousand matches. They only gathered data from matches where players were classified as having a "very high" skill level. They also used other third party APIs to collect statistics of players for past match data. When cleaning the data and picking features, the author used averages to fill in missing gaps for players with private profiles. They used vectors to represent data as input for logistic regression and neural networks. The author used these various techniques to build a model that makes a prediction of which team will win that begins with character draft data, and in real-time adjusts it based on in-game statistics, which is a bit out of scope for this paper but is a fitting example of how using esports data effectively with data mining workflows can lead to valuable and interesting products. The author's best performing pre-match model used logistic regression character, player, and character-player stats and had 71.5% accuracy. However, as the match continued, eventually the real-time data became more important for the overall model. The author's flagship model used a pre-match model in tandem with a time-series set of models using logistic regression which predicted the match outcome every 5 minutes, having an impressive 95% accuracy at the 40minute mark. This real-time predictor model could not only predict outcome matches, but also provide live predictions during important tournament matches watched by hundreds of thousands around the globe (Yang et al., 2016).

Methodology

In this section, two models are built from two different datasets generated from VALORANT match data. The steps required in order for this to happen are collecting the data, preprocessing the data, engineering features, and building the logistic regression model. Each section covers the steps taken to achieve these steps. There are useful figures in these sections to visualize these steps as well, and there are supplementary ones in the appendix for further detail about the logistic regression models.

Data Collection

The original idea for finding VALORANT match data was to use the Riot Games Application Programming Interface (API) free tier and get as much data as possible. However, after discovering API keys were only given to those with fully-functional production applications for VALORANT match data, it became quickly apparent it was going to be more of a challenge to get data than originally conceived. After discussing with community members on the Riot Games Third Party Developer Community server on Discord, one community administrator shared a large sample they had collected with me which ended up being the sole source of data for the project. They claimed it was similar in format to the official API output but could have slight formatting differences.

Data Preprocessing

The first step necessary for any data mining project is to clean and prepare data for use in data mining techniques. The data collected from section 3.1 started with around 550MB of VALORANT match data in raw JSON format. These data were separated into 11 files and contained extensive data of around 1100 VALORANT ranked queue matches played during Episode 3, Act 3, which was between the dates of November 2, 2021, and January 11, 2022. Each file contained around 1.5 million lines of JSON in pretty-printed form, which equates to over 17 million lines of JSON between all of the files.

Before cleaning the data it is important to try to understand the raw data at large to brainstorm ways to break it down and make it more usable for data mining techniques. In this case, the data had nested objects which could be interpreted as having a one-to-many relationship with the larger objects that nested them. There are three main "levels" that came from using this technique. The first is matches, which are top-level objects and encapsulate the other layers. These represent one of the matches played from the dataset, so around 1100 records of these were collected. This makes the first clean dataset derived from the raw data. The second is rounds, of which each match can have between 1 (which represents a voted 'remake' which cancels the game when a player leaves after the first round) and theoretically infinity (which represents an infinite number of overtimes if no team ever wins by two rounds). In this dataset, the minimum is 1 round played, or a remake, and the maximum is 32 (a quadruple overtime game). Since there are multiple rounds played per game, around 22,600 records were collected on this level. This makes the second clean dataset derived from the raw data. Finally, the third is fights, of which each round can have between 0 (nobody fought each other to win the round, or the round was forfeited due to one team surrendering) and 12 (all 10 players fought and killed each other, and there were also two player resurrections from each team's Sage who were both killed again as well). In this dataset, the minimum is 0 fights, and the maximum is 12 fights. Since there are multiple fights per round, a staggering 165,000 records were collected on this final layer. This makes the third clean dataset derived from the raw data.

The first and most prominent problem when cleaning the dataset was the data being multidimensional and not flat, which was an issue because logistic regression requires twodimensional datasets and therefore cannot use nested objects as attributes. This is the two-sided coin of the JSON format: on one end it is versatile and has loose rules/syntax which makes it easy to build, but on the other end this also makes it difficult to deconstruct, transform, and interpret at times. However, using the power of Python and for-loops for iterating through the data, the deconstruction of the large, nested objects went smoothly. This is mostly due to the fact that the data was formatted very consistently.

Feature Selection

The unflattening of the large JSON objects allowed for attributes to be extracted on the match, round, and fight level immediately. The more nested the objects were, the easier it was to identify more attributes because they were able to be derived from the layers they were nested in. Various categorical, numerical, Boolean, and binary attributes were extracted, and there were enough features on both the round and fight levels for logistic regression. Due to the top-most match level not having many features readily apparent and having only a small sample size of around 1100, it was decided to not move on with feature selection at that level.

The round layer dataset had a few readily apparent features retrieved from the JSON files directly. However, more features are inferred from other data in JSON, such as pistolRound (see table below). Derived features such as these saved an otherwise unfit dataset for logistic regression, due to the low amount of numerical, Boolean, and binary features which were readily apparent. The dataset ended up having 19 features, with six being categorical, four being Boolean, two being binary, and seven being numerical. It is worth noting while plantLocationX and plantLocationY are technically numerical, they correspond to in-game coordinates on maps, and are therefore not appropriate for predicting the dependent variable of attackerWon. They should be thought of instead as two-dimensional map top-down map data on the in-game maps instead.

Variable Name	Data Type	Variable Description
matchId	Categorical	The unique ID of the match played
attackerWon	Boolean	Whether the attackingTeam won the round
roundNum	Numerical	The round number of the game, starting from 0
attackingTeam	Binary	Which team is attacking (RED or BLUE)
winningTeam	Binary	Which team won the round (RED or BLUE)
roundResult	Categorical	The result of the round (Bomb detonated, Eliminated, Bomb defused, Round timer expired, or Surrendered)
roundCeremony	Categorical	Which ceremony animation will play after the round ends (NONE, DEFAULT, THRIFTY, CLOSER, FLAWLESS, TEAM_ACE, CLUTCH, ACE)
bombPlanter	Categorical	The unique player ID of the player who planted the Spike
plantRoundTime	Numerical	How long into the round the bombPlanter planted the Spike, in milliseconds
Variable Name	**Data Type**	**Variable Description**
plantLocationX	Numerical	The x-coordinate of where the bombPlanter planted the Spike. The coordinate units are specific to Riot's specifications, and vary by map
plantLocationY	Numerical	The y-coordinate of where the bombPlanter planted the Spike. The coordinate units are specific to Riot's specifications, and vary by map
plantSite	Categorical	Which site the bombPlanter planted the Spike (A, B, C, or None)

defuseRoundTime	Numerical	How long into the round a defender defuses the Spike, in milliseconds. Equals 0 if no Spike was planted during the round
roundResultCode	Categorical	The code which represents roundResult (DETONATE, ELIMINATION, DEFUSE, TIMEOUT, SURRENDERED)
attackerMoneySpent	Numerical	The sum of the loadout value of the attacking team
defenderMoneySpent	Numerical	The sum of the loadout value of the defending team
bombPlanted	Boolean	Whether or not the attacking team planted the Spike
pistolRound	Boolean	Whether or not it is the first round of each half, also known as pistol round (each player has only 800 credits to spend)
overtimeRound	Boolean	Whether or not it is an overtime round, which means roundNum is greater than 24

Table 1: Variable Description Table for Round Layer Dataset

The fight layer dataset had the largest amount of readily apparent features retrieved from the JSON files directly. However, there are also features inferred from other data in JSON, such as loadoutValueDiffFromEnemy (see table below). Derived features such as these allowed for a good portion of the features are numerical or Boolean, which made it a higher quality dataset for logistic regression. There are also some features which are derived from the round level in some way, such as loadoutValue, which is determined at the beginning of the round and does not change between fights within the round. The dataset ended up having 19 features, with six being categorical, two being Boolean, one being ordinal, and 10 being numerical. With the creation of the features, this concludes the data preprocessing portion of the data mining workflow.

Variable Name	Data Type	Variable Description
matchId	Categorical	The unique ID of the match played
wonFight	Boolean	Whether the player ID under 'player' won the fight
player	Categorical	The unique player ID of the player
enemy	Categorical	The unique player ID of the enemy the player fought
timeFightEndedMillis	Numerical	The time the fight ended between player and enemy in milliseconds since the start of the round
characterName	Categorical	The character the player selected at the beginning of the match
isDuelist	Boolean	Whether or not the character the player is using is categorized has utility used for taking fights (AKA a duelist)
loadoutValue	Numerical	The loadout value of the player
loadoutValueDiffFromEnemy	Numerical	The difference between the loadout value of the player and the loadout value of the enemy
moneyRemaining	Numerical	How much money the player has left after the shop closes and the round begins
boughtWeapon	Categorical	The unique ID of the weapon the player has after the shop closes and the round starts
armorStatus	Ordinal	The armor status (NONE, LIGHT_SHIELD, or HEAVY_SHIELD) the player has after the shop closes and the round starts
weaponFoughtAgainst	Categorical	The unique ID of the weapon the used by the enemy when the player and the enemy fight
Variable Name	**Data Type**	**Variable Description**

damageGiven	Numerical	The total damage given to the enemy by the player
damageTaken	Numerical	The total damage given to the player by the enemy
headshotsGiven	Numerical	The total headshots given to the enemy by the player
headshotsTaken	Numerical	The total headshots given to the player by the enemy
totalHitsGiven	Numerical	The total amount of hits given to the enemy by the player
totalHitsTaken	Numerical	The total hits given to the player by the enemy

Table 2: Variable Description Table for fight layer dataset

Logistic Regression

The data mining method chosen was logistic regression. The model gives a normalized output between 0 and 1 which is modeled with:

$$\text{P(Dependent Variable)} = \frac{1}{1+e^{-U}} \text{ where } U = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n$$

Where n is the number of predictors used to predict the dependent variable. Since logistic regression can be used in a predictive capacity, it is used in this project to classify the winner of a fight and the winner of a round in VALORANT. The literature review also suggests this to be an appropriate use of logistic regression.

Python was used to make the logistic regression models for this project. Additionally, the *pandas* and *sklearn* libraries were used. *Pandas* was used to transform the processed data into a data frame, drop irrelevant categorical and numerical variables, and display descriptive statistics of the data. *Sklearn* was used to split the data into training and testing sets, create the logistic regression model, display a classification report, display a confusion matrix, and visualize the confusion matrix.

There are two logistic regression models built for this paper. The first one is a predictive model in which the fight winner was predicted with 11 features, and the second one is a predictive model in which the round winner was predicted with seven features. Both are written in Python scripts which follow the same structure. First, the respective dataset was imported into a data frame using *pandas*, a common data science Python library. Then, the mean, median, and standard deviation are displayed to see basic information about each feature. Next, the dependent variable called wonFight is defined, and the other variables were defined as independent variables. Then, the data frame was split into a training set (70% of the total records) and a testing set (remaining 30% of the total records). Now the data is ready, so the logistic regression model is created and fit using the training data. Now the model is trained, so next it attempts to predict the dependent variable on the test data set. This concludes the data mining portion of the data mining workflow. Now it is time to complete the last step of the data mining workflow: post-processing the results, which is done by examining the classification report and confusion matrix of the model.

Results

In order to complete the post-processing step of the data mining workflow, the classification reports and confusion matrices from both models should be examined and interpreted. The round layer logistic regression model turned out to have around 73% accuracy, which is achieved by having good, but not the best, features and dataset size. According to the classification report in the appendix, the model has little imbalance between false positives and false negatives, which is a signal the model is equally dependable predicting when the attackerWon variable is equal to one or equal to zero. Since the dependent variable is a symmetric Boolean variable, this behavior is ideal because one value of wonFight is not more important than the other. In fact, overall accuracy is the only important metric for evaluating this model because it is predictive and not attempting to explain historical data. Because this model's performance is not optimal, it may be improved by including more records in the dataset. It may also improve with better feature selection either by means of retrieving more helpful raw data or deriving more useful features from the raw data used for this paper. Out of 6,789 test records, 1,832 were incorrectly classified, as the visualization of the confusion matrix below shows:

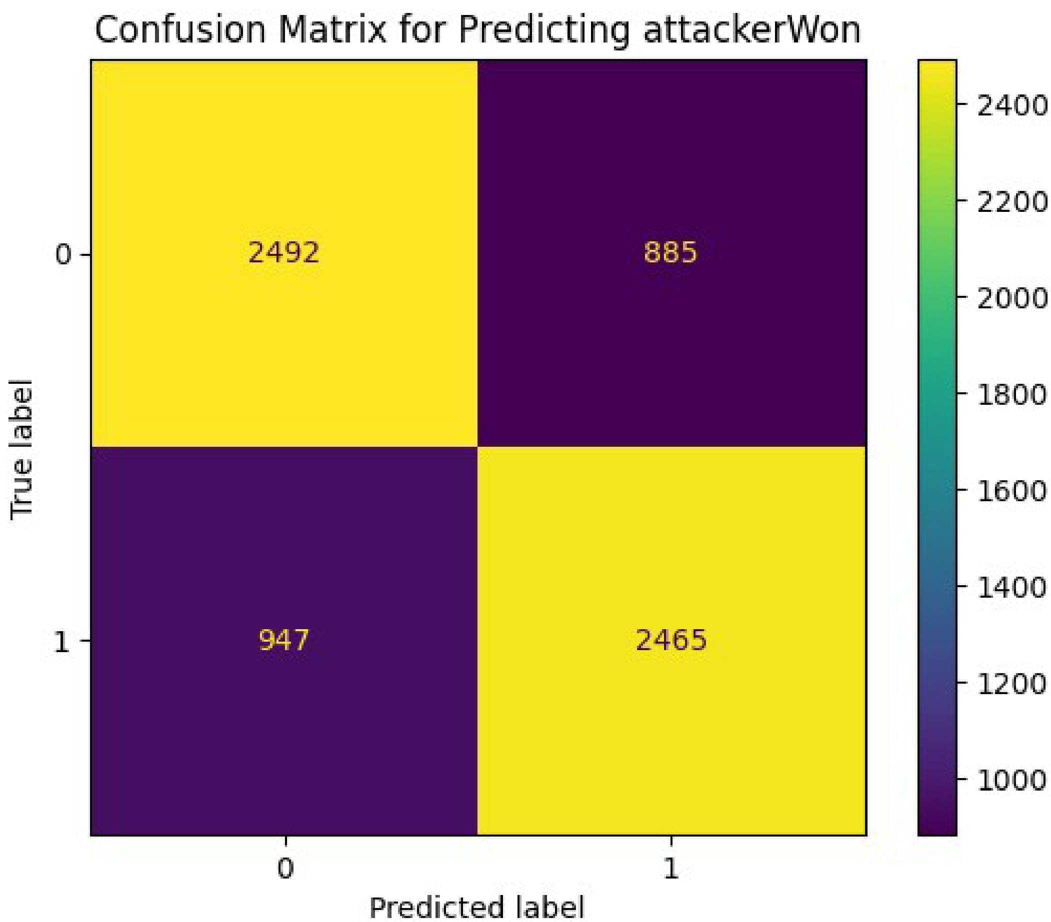


Figure 1: Visualization of Confusion Matrix for Round Layer Logistic Regression Model

The fight layer logistic regression model turned out to have around 97% accuracy, which is achieved due to the quality of the features and the large amount of training data. According to the classification report in the appendix, the model had no imbalance between false positives and false negatives, which is a signal the data was symmetrical. Since the dependent variable is a symmetric Boolean variable, this behavior is ideal because one value of wonFight is not more important than the other. In fact, overall accuracy is the only important metric for evaluating this model because it is predictive and not attempting to explain historical data. Because this model's performance was so high, there was no need to change it. Out of 49,630 test records, only 1,303 were incorrectly classified, as the visualization of the confusion matrix below shows:

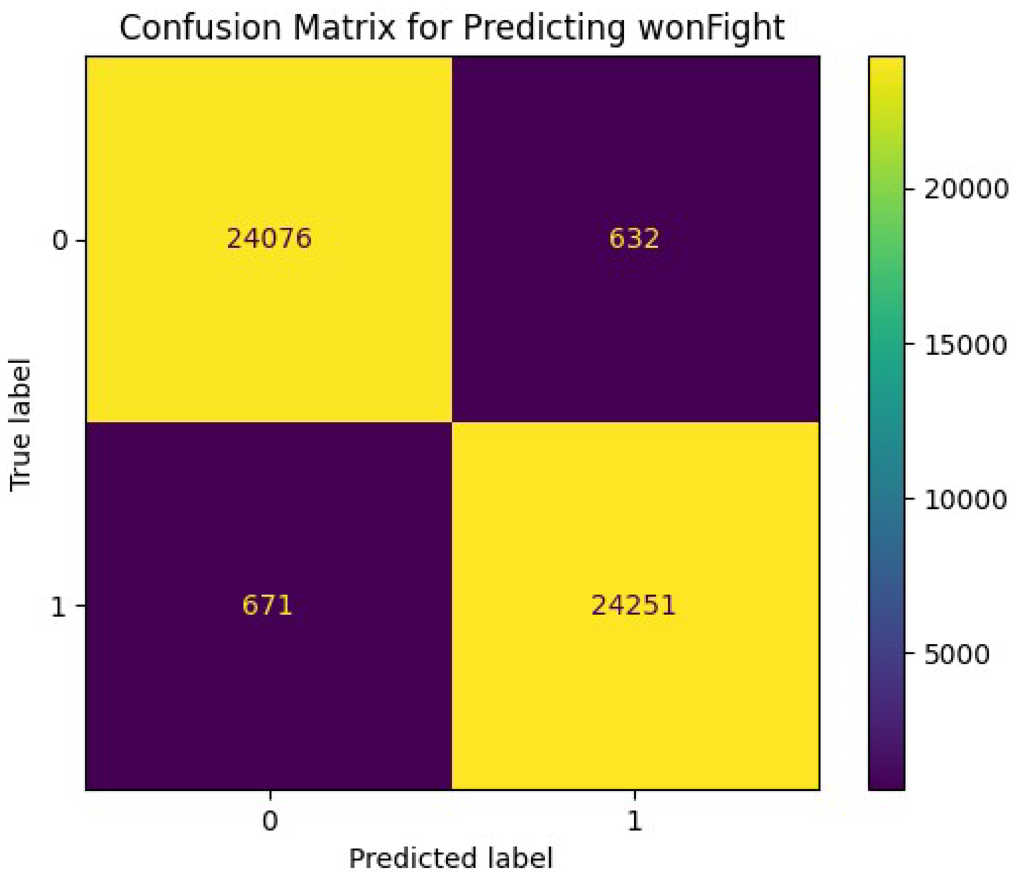


Figure 2: Visualization of Confusion Matrix for Fight Layer Logistic Regression Model

One thing to note is the splitting of training and testing data and the training and testing of the logistic regression model itself were both completed using a random state equal to 5. In *sklearn*, a random state ensures the models are reproducible, meaning when running the Python scripts over and over again they will output the exact same models. Other random states can yield slightly better and slightly worse results, but the differences are negligible. When a random state is not chosen, the output will be slightly different each time. Measures were taken in both logistic regression Python scripts to ensure reproducibility of both models.

Conclusions and Future Work

The two models created in this paper varied slightly in their success in predicting the dependent variables. When predicting which player will win a fight, 97% accuracy is as high as one can hope to achieve with just logistic regression. One of the model's possible uses could be providing insight to Riot Games employees as to how interactions between players in the game play out. This is valuable because future game balances and updates can be informed by these data insights. Another use could be providing beginner players with helpful tips for how to maximize their chances of being competitive during minute-to-minute gameplay. This may increase player engagement since it will help them with the difficult learning curve the game has for beginners.

On the other hand, when predicting which team will win each round the model had 73% accuracy, which is a good model but far from perfect. The model can be improved by more feature engineering and gathering a larger dataset than the one for this paper. It could also be improved by using another machine learning algorithm, such as neural networks. The model's possible uses could be providing insights for team analysis for lower level teams. An addition to the game which analyzes team strengths and weaknesses and provides feedback would be an excellent feature for the new Premier competitive system in which players build their own roster of five players and play in pre-scheduled matches against one another (Cropley, 2023). It would encourage teams to become better together and make the community more tightly knit.

As for future work, there are more data mining workflow possibilities with the datasets generated for this paper. Another direction the work can be taken is towards data visualization, which would help audiences understand stories behind esports data and inspire creativity for future programming-powered projects, including data mining ones. In addition, this paper shows the perspective of deriving multiple layers of data for an esports match. Like in regular sports, hundreds or thousands of small events during matches contribute to victories and defeats, and creating multiple datasets like in this paper opens up more dimensions when deriving useful information from data. The technique of analyzing these small events as well as the matches themselves should be explored in other areas of the esports industry.

Acknowledgements

I would like to thank Dr. Yixiu Yu for her advice and help throughout the process of completing the creative project, authoring this paper, and for teaching me about data mining and data visualization in previous courses. The paper would not have been possible without her help.

References

Bouzianis, Stephen, "Predicting the Outcome of NFL Games Using Logistic Regression" (2019). *Honors Theses and Capstones*. 474. https://scholars.unh.edu/honors/474

Cropley, S. (2023, March 3). *Riot reveals Valorant premier competitive system beta -- an ingame path to pro*. ONE Esports. Retrieved March 15, 2023, from https://www.oneesports.gg/valorant/valorant-premier-competitive-system/

Khosla, K. (2021, November 9). *Why Valorant? A look into the community behind Riot Games New Shooter*. DBLTAP. Retrieved March 15, 2023, from https://www.dbltap.com/posts/why-valorant-a-look-into-the-community-behind-riotgames-new-shooter-01fkpa738g4g

Landa, J. (2020, June 1). *Beginner's Guide*. VALORANT. Retrieved March 15, 2023, from https://playvalorant.com/en-us/news/announcements/beginners-guide/

Roach, J., & Caswell, T. (2021, March 15). *What is valorant? everything we know about riot's shooter*. Digital Trends. Retrieved March 15, 2023, from https://www.digitaltrends.com/gaming/what-is-valorant/

Semenov, A., Romov, P., Korolev, S., Yashkov, D., & Neklyudov, K. (2018). Performance of Machine Learning Algorithms in Predicting Game Outcome from Drafts in Dota 2. *Communications in Computer and Information Science*, 1–12. https://doi.org/10.1007/9783-319-52920-2_3

Webster, A. (2020, November 24). *Valorant Champions Tour is a global season-long esports competition for 2021*. The Verge. Retrieved March 15, 2023, from https://www.theverge.com/2020/11/24/21612562/valorant-champions-tour-esports-2021

Yang, Y., Qin, T., & Lei, Y.-H. (2016). Real-time eSports Match Result Prediction, 1–9. https://doi.org/10.48550/ARXIV.1701.03162

Appendix: Additional Outputs for Logistic Regression Models

In this section there are various raw outputs generated with the Python scripts used to create the two logistic regression models for this paper. All variable descriptions are found in the feature selection section (3.3).

Logistic Regression for Predicting wonFight

Descriptive Statistics:

Variable Name	Mean	Median	Standard Deviation
wonFight	0.4994167	0	0.500001171
timeFightEndedMillis	35748.266	31796	21736.50406
isDuelist	0.4326577	0	0.495445761
loadoutValue	3581.2136	4200	1623.061987
loadoutValueDiffFromEnemy	-2.229026	0	1717.68347
moneyRemaining	1763.1333	1200	1859.684212
damageGiven	91.986883	80	104.5699814
damageTaken	92.07527	80	104.4382111
headshotsGiven	0.2924361	0	0.502983454
headshotsTaken	0.2925268	0	0.502479802
totalHitsGiven	1.6311876	1	1.685282773
totalHitsTaken	1.4720155	1	1.733871971

Classification Report:

	Precision	Recall	F1-Score	Support
True Value 0	0.97	0.97	0.97	24708
True Value 1	0.97	0.97	0.97	24922
Accuracy			0.97	49630
Macro Average	0.97	0.97	0.97	49630
Weighted Average	0.97	0.97	0.97	49630

Other Information:

Number of Features	11
Number of Iterations	54

Logistic Regression for predicting attackerWon

Descriptive Statistics:

Variable Name	Mean	Median	Standard Deviation
attackerWon	0.497922927	0	0.500006734
roundNum	10.11578575	10	6.459602261
plantRoundTime	24971.551	23933.5	25874.34562
attackerMoneySpent	12543.70691	13250	5911.272414
defenderMoneySpent	12011.07256	12400	5755.434437
bombPlanted	0.582906134	1	0.493089564
pistolRound	0.102395263	0	0.3031741
overtimeRound	0.015467562	0	0.123405793

Classification Report:

	Precision	Recall	F1-Score	Support
True Value 0	0.72	0.74	0.73	3377
True Value 1	0.74	0.72	0.73	3412
Accuracy			0.73	6789
Macro Average	0.73	0.73	0.73	6789
Weighted Average	0.73	0.73	0.73	6789

Other Information:

Number of Features	7
Number of Iterations	11

Entrepreneurship

18.

E-Residency in Estonia: A Financial Analysis

Muhammad Umar Farooq

New Mexico State University, Las Cruces, USA

Abstract

The introduction of the concept of e-residency by Estonia has exciting implications for businesses, entrepreneurs, and the future of digital governance. E-Residency enables entrepreneurs to establish a company in Estonia, by providing them with a digital identity, irrespective of their physical location. Digital identity and entrepreneurship have grown considerably in the small Baltic country since the introduction of e-residency in 2014. The author proposed that digital location is starting to play a role in decision-making for entrepreneurs. Entrepreneurs are beginning to realize the value of establishing and managing companies online in a European Union country. The program continues to be profitable for Estonia and the e-residents who have taken advantage of it to establish companies in the country. The program's success has led other countries, such as Portugal and Lithuania, to initiate their own e-residency programs.

Keywords: E-Residency, Entrepreneurship, Financial Analysis, E-Estonia, Digital Identity

Categories

Entrepreneurship

Healthcare

6.

Data Driven Decision Making in Healthcare Sector: A Case Study Approach

Shubham Singh

Purdue University Fort Wayne, Fort Wayne, USA

Abstract

The aim of this research is to study the impact of data-driven decision making on outcomes of healthcare delivery. Business analytics tools can help provide healthcare professionals effective ways to treat based on appropriate markers of the patient and therefore provide efficient pathways towards recovery. This research would survey case studies of healthcare providing institution to understand their current use of analytical technology and its effectiveness for patient satisfaction. The results of this research will shed light on current uses of analytics use in healthcare setting to improve healthcare delivery outcomes.

Keywords: Business analytics, healthcare delivery, personalized medicine, case study, data-driven, decision-making.

Categories

Healthcare

12.

Queueing inspired feature engineering to improve and simplify patient flow simulation metamodels

Mark Isken, Osman Aydas, Yazan Roumani

Oakland University, Rochester, USA

Abstract

We explore the interplay between domain informed feature engineering, model performance and model interpretability. This is a hybrid modeling and simulation study that merges the application of discrete-event simulation with alternative metamodeling techniques for modeling patient flow in healthcare. We consider two cases: A tandem queueing system of obstetric hospital units and a transient analysis (a 4-hour clinic block) of an outpatient clinic in which a finite number of scheduled patients arrive for care. We utilize several metamodels including linear models, random forest and a neural network. We evaluate the performance improvement of metamodel estimations when empowered with supplementary queueing theory knowledge. We consider three knowledge levels: No knowledge (no queueing inspired features), basic (simple queueing features) and advanced (sophisticated queueing approximations). Our results show that queueing related inputs improve the accuracy for the metamodels, independent of the model type. Moreover, the queueing related inputs improve model explain ability and can lead to more parsimonious models. This has positive practical implications for implementing these types of models in actual healthcare analytics projects.

Categories

Healthcare

25.

Global Patent and Citation Data Analysis to Inform Knowledge Directives

Abigail Richard [1], Fred Ahrens [2], Benjamin George [2]

[1]University of Indianapolis, Indianapolis, USA. [2]The University of Toledo, Toledo, USA

Abstract

It is well-accepted that innovation has the potential to be greatly beneficial to firms. However, despite the competitive advantage that innovation can provide, the innovation process tends to be risky. This is especially true in the pharmaceutical industry, where the drug development failure rate is high. As a consequence, it is important for firms to strategically consider potential avenues for innovation research, and selectively choose among plausible research objectives. To aid R&D pharmaceutical partners, we examine the consequences of various relative proportions of knowledge on innovation, and the corresponding role of these knowledge proportions in partnerships. In particular, we examine how a firm's knowledge changes on a year-to-year basis and how that rate of change influences its innovation success. Using patents from authorities throughout the world, we further examine the innovation of each firm on a global scale – allowing us to gain insight into the global innovation process for the firms in the study. We extend the results of our study to a multi-level analysis investigating the effect of these knowledge proportions at both the firm and individual level. For this, we utilize patent data at the firm level from leading pharmaceutical companies, as well as individual level citation data among researchers.

Categories

Healthcare

35.

Disposition of Unused Medical Supplies

Submitted by

Rachel Baur

Biomedical Sciences

Oakland University

and

Gene Fliedner

School of Business Administration

Oakland University

Abstract

This research project investigates the waste associated with the disposal of unused medical supplies, the current methods used to repurpose and donate unused medical supplies, and the barriers associated with repurposing and donating unused medical supplies. A literature review and an interview with an expert in the field has been conducted. The research done for this project suggests that it is possible to decrease the disposal of unused medical supplies. One example of a means waste disposal can be decreased is by resterilizing open and unused medical supplies with hydrogen peroxide vapor. Another means the waste of disposal can be decreased is through the establishment of a collection, repurposing, and donation program. Examples of programs that currently exist include Recovered Medical Equipment for the Developing World (REMEDY), Recycling Unused Medical Supplies (RUMS), MedShare, and Project Commission on Urgent Relief and Equipment (CURE). Studies done with hydrogen peroxide vapor and the collection, repurposing, and donation programs show a decrease in money spent in healthcare systems and reduced waste in landfills. Along with the positive aspects of these methods, there are a handful of barriers that need to be addressed. These include deterioration of desterilized supplies, warranty issues of donated supplies, the responsibility of impoverished countries to desterilize the donated supplies, an overabundance of incorrect supplies donated to charities, the fallacy that individual action will not create a measurable difference, and little regulation of medical supplies that are in hospital supply rooms. Results of this project can help future investigations devise plans to effectively collect, repurpose, and donate medical supplies. Healthcare systems worldwide will benefit from this project because they will have a better understanding of the issue along with the barriers that still need to be overcome.

Introduction

It is a common practice in medicine to dispose of any medical supply that is or could have possibly been contaminated. Even if the packaging of the medical supply was never opened, many healthcare workers will opt to throw away the supply due to the idea that anything taken out of a supply room and not used, has to be thrown away because of possible contamination that can occur once the supply is in the patient's room. While medical supplies should be disposed of if they are directly and obviously contaminated, medical supplies that are unused or not directly contaminated should not be. The disposal of unused medical supplies is not a new problem and it is costly to not only hospitals, but also the environment.

The World Health Organization (WHO) defines medical supplies as any instrument, apparatus, implement, machine, appliance, implant, reagent for in vitro use, software, material or other similar or related article, intended by the manufacturer to be used, alone or in combination for a medical purpose (WHO, 2022). Some examples of supplies that are commonly disposed of include nasal cannulas, nebulizers, suction supplies, catheter needles, endotracheal tubes, masks, cleansing supplies, IV supplies, phlebotomy tubes, syringes, urology supplies, vital sign equipment, eye protection, gloves, gowns, towels, speculums, procedure kits, scissors, clamps, forceps, scalpels, sutures, tape, and gauze (Muldoon et al., 2019).

While unused medical supplies are disposed of in almost all medical facilities, the most common places that dispose of medical supplies include emergency departments, operating rooms, intensive care units, and isolation rooms. There is a large amount of unused medical supplies disposed of in emergency departments and operating rooms because both of these units have a fast patient turnover rate. Therefore, the emergency department and the operating room employees are rushed in between the arrival and discharge of patients leaving little time, if any, to collect and sort through unused supplies.

Results of a study showed there to be 87.9 pounds of recoverable medical supplies disposed of in an emergency department over the course of thirty days (Muldoon et al., 2019). In a similar study done to find the amount of medical supplies disposed of in operating rooms, it was discovered that there were approximately two million pounds of recoverable medical supplies disposed of each year from large non-rural operating rooms (Wan et al., 2014).

Intensive care units are another major contributor to the disposal of unused medical supplies because of the patient complexity and needs requiring extensive supplies, cleaning practices, and pre-emptive supplies (Ghersin et al., 2020). Results of a study done in an intensive care unit discovered that roughly 168 pounds of unused medical supplies were disposed of over a three-week period (Ghersin et al., 2020). Isolation rooms also have a high rate of unused medical supply disposal in order to prevent items from becoming pathways for transmission of the disease the patient was isolated for (Otter et al., 2020). In the same study done regarding the amount of unused medical supplies disposed of in an intensive care unit, it was estimated that greater than 75 percent of the disposed items in the intensive care unit over the course of the three weeks came from isolation rooms (Ghersin et al., 2020).

There are multiple reasons as to how unused medical supplies are disposed of and why the disposal of unused medical supplies is so common. First, medical supplies that are opened but unused are mainly disposed of because of federal regulations, hospital protocols, and procedural excess in the United States (Muldoon et al., 2019). Second, due to a sense of immediacy in the interest of optimal surgical management and protection from litigation, more than the requisite supplies are typically opened for a given operation (Rosenblatt & Silverman, 1992). Third, most hospitals make no effort to separate non-infectious waste from infectious waste for purposes of disposal mainly because of the time constraint put on healthcare workers to get rooms ready for a new patient immediately after a patient is discharged (Pennino et al., 1998). Fourth, supplies are unnecessarily readied in the surgical field only to be disposed of unused, which results from redundant procedures, outdated policies, and unsubstantiated practices (Rosenblatt & Silverman, 1994). Fifth, there is little incentive to monitor the issue and many people decide to just accept the inevitable loss (Rosenblatt & Silverman, 1994). In other words, many healthcare workers are indifferent to this issue because there is a fallacy that individual action will not create a measurable difference (Ghersin et al., 2020).

The disposal of unused medical supplies is wasteful for a variety of reasons. For example, it has considerable cost implications. In a study done to determine the possible savings from repurposing medical supplies, the results showed that the annual cost of supplies discarded at patient hospital discharge was $387,055 (Otter et al., 2013). Results of another study done in the United States estimated there to be $200 million worth of prepared materials discarded in operating rooms each year (Rosenblatt & Silverman, 1994). A third study found that as much as 25 percent of all United States healthcare spending is wasteful, with an accumulated debt between $760 million and $935 billion (Ghersin, 2020). However, only calculating the cost of medical supplies disposed of actually underestimates the financial burden of this practice. In other words, another major cost associated with discarding these supplies is the disposal cost of the waste generated. Research has shown that disposal fees may actually exceed the purchase cost for some products (Otter et al, 2013).

There is an additional opportunity in the avoidance of medical supply disposal. Medical facilities in much of the developing world are characterized by a lack of some of the basic supplies required for proper patient care (Miller, 2002). When medical settings in the United States dispose of their basic medical supplies, they are throwing away the chance to donate or sell these items to impoverished countries that are in need of them. For example, impoverished countries will actually reuse needles until they are too dull to pierce the skin. Disposable syringes, tubing, drains, catheters, and even gauze are recycled to be used again since there is such a low supply of them in these countries (Pennino et al, 1993). Unfortunately, medical facilities in these impoverished countries do not consistently sterilize these recycled medical supplies due to knowledge gaps and poor guideline adherence (Cuncannon et al., 2020). Having to reuse these unsterilized supplies not only limits the basic health care given to these patients, it also increases the stress of healthcare workers due to their knowledge of possible cross contamination between patients (Pennino et al, 1993).

The disposal of unused medical supplies is also dangerous for the environment. Since the greenhouse gas (GHG) emissions attributed to waste are not tangible, excessive waste has become deeply-rooted in the United States healthcare culture. Research has found that although the total amount of United States greenhouse gas production has decreased in the past ten years, the healthcare system continues to be a major contributor to overall GHG production. In a study comparing single-use versus reusable medical supplies, several medical products, such as scissors and laryngeal mask airways, were found to have lower life-cycle greenhouse gasses when used as a reusable as opposed to single-use disposable (Ghersin, 2020). Therefore, the avoidance of medical supply disposal presents another opportunity for healthcare systems.

Current research

Creating a method to combat the disposal of unused medical supplies represents an important opportunity for the environment and healthcare systems worldwide. One current method that is in place to help this issue includes collecting unused medical supplies and resterilizing them in order to reuse the supplies at the facility they are already located in. Another method that is currently in place to assist with this issue is collecting the unused medical supplies and donating them to countries and other medical facilities in need of them.

Sterilizing unused medical supplies instead of disposing of them is an available method that healthcare systems should consider. Not only does it generate financial benefits to medical facilities, this method also benefits the environment. While the manual application of disinfectants to unused medical supplies is an impractical idea due to the amount of time it would take for it to be completed and the risk of the disinfectant not ridding the medical supplies of contamination, the use of hydrogen peroxide vapor is an alternative approach that is much more efficient and effective against a wide range of pathogens. Hydrogen peroxide vapor is the vapor form of hydrogen peroxide, which itself is simply water with one added oxygen molecule. However, it is this extra oxygen molecule that has high-oxidizing powers making hydrogen peroxide an extremely powerful disinfectant. Normally, following a patient discharge, hydrogen peroxide vapor is used to sterilize the hospital room and large pieces of equipment in the room. Hydrogen peroxide vapor is distributed homogeneously in the gas phase to ensure adequate coverage in the hospital rooms and on the large pieces of equipment. This method has been associated with a reduction in the incidence of pathogen acquisition of patients. Therefore, hydrogen peroxide vapor technology may provide a creative option to sterilize the packaging of contaminated supplies and avoid their disposal (Otter et al., 2013).

To test the effectiveness of hydrogen peroxide vapor on unused medical supplies, a study was conducted evaluating the frequency of contamination on the packaging of unused medical supplies in twenty rooms of patients under precautions for multidrug resistant organisms and then the amount of multidrug resistant organisms remaining on the supplies after sterilization with hydrogen peroxide vapor (Otter et al., 2013). Five pairs of unused, packaged supplies were selected from the supply cart in each study room when the patients were discharged from the hospital. One item of the pair was sampled without exposure to hydrogen peroxide vapor, and the other item was sampled after hydrogen peroxide vapor exposure. The five selected items were laid out on a metal rack and placed in the patient rooms that were going to be sterilized with hydrogen peroxide vapor following the patient discharge as mentioned above. Results showed that none of the one-hundred paired supplies had multidrug resistant organisms after hydrogen peroxide vapor exposure. The data of this study demonstrates that treated supplies can be safely used on new patients when a patient under precautions is discharged.

There are multiple advantages of using hydrogen peroxide vapor to sterilize unused medical supplies. First, since hydrogen peroxide vapor is already used to sterilize environmental surfaces and medical equipment in a patient room after discharge, there is no additional work needed to sterilize the unused medical supplies. The operators of the hydrogen peroxide vapor technology simply have to keep the supplies in the intended room and continue with the normal process of hydrogen peroxide sterilization. Second, results showed there to be a direct savings associated with hydrogen peroxide vapor sterilization of $387,055 (Otter et al., 2013). Third, following the use of hydrogen peroxide vapor, the hydrogen peroxide breaks down rapidly into oxygen and water leaving no chemical residues making it environmentally friendly (Chemdaq, 2022). Fourth, hydrogen peroxide vapor is simple for the operator of the hydrogen peroxide vapor technology to use. The easy application of hydrogen peroxide vapor is advantageous because it ensures adequate coverage of the supplies, which means the supplies will be completely sterilized (Otter et al., 2013). Overall, this method is an efficient and effective way to prevent the disposal of unused medical supplies.

Another method that can be used to combat the disposal of unused medical supplies is to collect and repurpose, or donate the unused medical supplies. Some popular programs that collect and repurpose unused medical supplies include REMEDY and Recycling Unused

Medical Supplies (RUMS). REMEDY is a program designed to recover usable medical supplies that would otherwise be discarded from United States hospitals (Miller, 2002) while RUMS was created to collect unused medical supplies and outdated medical equipment from United State hospitals to reduce the environmental impact of hospital waste and provide direct support to Third World hospitals (Pennino et al., 1993).

REMEDY began in June of 1991 and went through an eleven-month trial at the Yale-New Haven hospital in Connecticut. Prior to the start of REMEDY, the nursing and technical staff at Yale-New Haven hospital were informed of program goals and guidelines. They were also provided lists of recoverable supplies as well as exclusion criteria. The list of recoverable supplies included surgical gloves, sutures, surgical staples, sterile laparotomy sponges, non sterile sponges, sterile 4x4 sponges, cautery pencils/pads, endotracheal tubes, syringes, drapes and gowns, surgical drains, suction hose, nursing/surgical miscellany, anesthesia miscellany, and urology misscellany. Next, a bag for collection of recovered supplies was placed on the case cart prepared for each operative procedure. The collection of supplies was accomplished with the volunteer efforts of the nursing and technical staff and the cooperation of the hospital administration and Department of Anesthesiology (Rosenblatt & Silverman, 1992).

The exact setup of REMEDY for the eleven-month trial included the following (Rosenblatt & Silverman, 1992). Throughout an operation, non-sharp and non-exposed sharp items whose sterile packaging had been violated, but had not been used in the surgical procedure or noticeably contaminated were placed in the bag on the case cart. Immediately following the procedure, the bag accompanied contaminated surgical instruments to the central decontamination area. At the end of each day, the bags were emptied by a designated individual who inspected the contents for unsuitable materials such as accidentally recovered opened sutures and spoiled sponges. The remaining contents underwent ethylene oxide sterilization and aeration, which is identical to the sterilization protocol used for reusable supplies at Yale-New Haven Hospital. This sterilization protocol meets or exceeds the sterilization procedures used by established charities that receive the supplies. In addition to these daily collections, the nursing staff and supply managers periodically reviewed the operating room inventory for items that were no longer used due to changes in surgical techniques or expiration.

Results of the REMEDY program were positive (Rosenblatt & Silverman, 1992). The operating room staff responded enthusiastically to this effort. A survey completed by fifty nurses and technical staff revealed no reports of injury, interference with patient care, or increased operating room time or labor. Along with the new sense of cooperative achievement among the operating room staff, REMEDY also identified areas of unnecessary waste and reduced biohazard bag disposal at Yale-New Haven Hospital at a savings greater than the minor costs of the program. The total direct cost for the paper bags and other items for the implementation and maintenance of the program was $420. The inspection and sorting of supplies by volunteers required one and a half to two hours per week once the program became more established. Throughout the eleven-month trial, 3,776.5 pounds of supplies were collected from operating rooms with an estimated value of $158,854. Recovered supplies were distributed to established charitable organizations with a history of medical supply relief work that had requested supplies. Highly sophisticated supplies such as cardiopulmonary bypass oxygenators were donated directly to colleagues abroad known to be experienced with the relevant technology.

The REMEDY program has also been in place at the Albert Einstein College of Medicine in New York since 1997 (Miller, 2002). REMEDY at the Albert Einstein College of Medicine uses several collection techniques, with the bulk of material originating from Jacobi Hospital surgical floors. Similar to the trial done at the Yale-New Haven hospital, training sessions with the operating room nursing staff at Albert Einstein College of Medicine are performed to familiarize employees with the purpose and scope of REMEDY. The staff is instructed to only recover clean items and to not recover sharps and body fluids. Following the instructions, unused and uncontaminated surgical supplies that would have been discarded were collected on a case-by-case basis by operating room nurses and placed in REMEDY storage containers. Common items collected were sutures, gloves, gauze, drapes, sterilization kits, catheters, intravenous tubing, airway tubing, sponges, staplers, syringes, and reusable surgical tools. These containers are then transported on a regular basis by medical students to the REMEDY supply room, where they are inspected and sorted by volunteer staff trained in universal precautions. Supplies to be donated are inventoried and stored for donation, and defective material is discarded. Donation and shipping of REMEDY supplies to their final destination relies on contacts made with institutions and established United States based charities.

The cost of the REMEDY program at the Albert Einstein College of Medicine is extremely low, consisting primarily of collection containers and use of storage space (Miller, 2002). REMEDY at Albert Einstein College of Medicine uses large recycling bins that can be brought into the operating room after a procedure during cleanup, minimizing the amount of time that operating room nurses separate usable items from medical waste. For an annual budget of approximately $200 spent mostly on boxes for storage and shipping, REMEDY is able to recover thousands of dollars of usable medical equipment. There is also no cost associated with the staff that collect and sort the supplies since volunteer labor is provided by hospital staff, medical students, and other interested parties. The resulting reduction in medical supply disposal actually saves the hospital six to ten times the cost of the program, as well as raising staff members' awareness of waste and recycling issues. Additionally, the data generated from the types of supplies recovered can be used to increase the efficiency of operating room supply management.

A common concern with the REMEDY program is how it is able to organize the supplies collected and distributed to the correct locations (Rosenblatt & Silverman, 1994). To monitor the nature and quantity of recovered supplies in any medical setting, REMEDY uses a computer database that inventories approximately 100 types of general operating room supplies and 320 types of sutures on a case-specific basis. Since recovery and inventory occur daily, detailed case-by-case assessments are applied periodically to a large sample and has been facilitated by a barcode system applied to both item stock and procedure numbers. Recovered inventory can be compared to a surgical pick list and use list. In addition to permitting REMEDY to project charitable donations from institutions adopting the REMEDY model, this database has evolved into a powerful resource utilization assessment tool. In other words, REMEDY has reduced the amount of excessively prepared supplies. Using the REMEDY database, the nursing administration and operating room supply managers at the locations that use REMEDY have initiated a campaign of heightened awareness and reevaluated operating room supply handling of all cases.

Another important question regarding the REMEDY program is the effect that these supplies have on the quality of care at the receiving institutions. In order to address this issue, primary medical centers in South America were surveyed regarding which supplies that they received through REMEDY would be clinically useful (Miller, 2002). In general, primary care clinics, particularly in rural settings of South America, stated that the most basic supplies such as gloves, gauze, sutures, syringes, sterilization solution for smaller rural clinics and advanced supplies such as ultrasound and electrocardiogram machines for larger clinics and hospitals would be the most useful to receive. Therefore, the supplies REMEDY donates are very effective in providing high quality care since most of the supplies donated are the ones in high demand in impoverished countries.

RUMS is a program similar to REMEDY. At its inception, RUMS involved twelve hospitals, two-free standing surgical centers, one pharmacy, and several nursing homes, health centers, and private doctors' offices (Pennino et al., 1993). RUMS is managed by a small paid staff to provide reliability and consistency in the collection of materials, maintenance of the database, grant writing, public relations, and day-to-day operations. Funding and supplies also have come from community and organizations in Rochester, New York. The Rochester Chapter of the Association of Operating Room Nurses has provided financial support as well as many nurses who not only salvage hospital supplies, but also volunteer to sort and inventory them. RUMS has received grants from two local foundations, and a local grocery chain has donated gift certificates to purchase recycling bins and regularly provides 40-gallon pressed cardboard drums from its bakery for shipping supplies. A local business donates 2000 square feet of warehouse space, and an automobile dealer provides a van to pick up large items. Local media have provided excellent publicity for the project encouraging individuals to contribute money and to become volunteers.

RUMS is divided into four phases: setup and education, collection, sorting and inventory, and distribution (Pennino et al., 1993). The setup phase is used to elicit support from hospitals and health-care facilities in the community. Similar to REMEDY, the RUMS program offers numerous slide presentations to nurses, administrators and allied health personnel to educate them about this program, which allows hospital personnel to recycle medical supplies overseas and locally reduce the environmental impact of disposing of them. Special emphasis is given to the supplies that could be collected which include anesthesia items, bassinets, catheters, cautery supplies, collection containers, drapes, dressings, endotracheal tubes, gloves, gowns, hospital beds, incubators, irrigation sets, needles, orthopedic items, packs and procedure sets, petroleum gauze, scrub brushes, specialty items, sponges, staples and removers, surgical instruments, sutures, syringes, tape, and tubing. Additionally, any outdated medical supplies that are still functioning but normally would be discarded are collected.

During the collection phase, paid collectors make regularly scheduled pickups at each facility and deliver the supplies to the warehouse (Pennino et al., 1993). The collector initially presorts materials into general categories, such as paper goods, tubing, instruments, sutures, and dressings. In the sorting and inventory phase, volunteers work under the guidance of volunteer nurses and technicians to sort the collected materials into designated categories. Those categories are divided into numerous subcategories for accurate record keeping. Supplies are packed in boxes, labeled, and weighed for shipping; an inventory sheet is completed for each box; and statistics are entered into a computer database to track the inventoried supplies along with their value, weight, and destination.

In the distribution phase, RUMS works with organizations in the United States that provide medical relief to underdeveloped countries (Pennino et al., 1993). The organizations RUMS works with are in close proximity to where RUMS is taking place. Therefore, the supplies that are distributed to these organizations can be sent via large vehicles driven by volunteers or via mail carriers.

RUMS collected more than 30,000 pounds of supplies, with a retail value greater than $500,000 during its first ten months of operation (Pennino et al., 1993). These numbers represent only the supplies and equipment that were inventoried and packed in boxes. The amount of supplies collected was so great that it was not possible to inventory all of it.

The success of RUMS depended on community involvement. Volunteers, mainly nurses and other healthcare personnel, sort and inventory the collected materials and equipment, because familiarity with the supplies is key to tracking them (Pennino et al., 1993). As a result of RUMS, three shipments of medical supplies were sent to a hospital in Novgorod in Russia, a Somali refugee hospital in Kenya, a hospital in Uganda, and hospitals in Poland, Jamaica, Haiti, El Salvador, and Nicaragua.

While these programs collect, sort, and repurpose the unused medical supplies, they are not the ones that directly give the supplies to the countries in need. There are separate organizations that programs such as REMEDY and RUMS work with directly in order to get these supplies to the correct destinations. Two of the most popular organizations that specialize in the donation of unused medical supplies include MedShare and Project CURE.

MedShare is a 501c(3) humanitarian aid organization, which is a non profit organization established for charitable purposes, dedicated to improving the quality of life of people, communities, and the planet by sourcing and directly delivering surplus medical supplies to communities in need around the world (MedShare, 2022). In 1998, A.B. Short and Bob Freeman were concerned with the large amount of discarded medical supplies and the critical healthcare needs of medically underserved populations. In an effort to address these disparities, MedShare was born. MedShare works with hospitals, distributors, and manufacturers to collect and redistribute medical supplies to qualified healthcare facilities in medically underserved communities. MedShare also helps increase health system capacity and drives sustainability by providing biomedical equipment training and service to healthcare organizations and medical professionals serving populations in need. MedShare’s deliveries of vital medical supplies has decreased the nation's carbon footprint by decreasing the amount of unused medical supplies found in landfills and it has brought health, healing and promise of better lives to 100 countries and countless patients.

MedShare works specifically with hospitals in Metro Atlanta, Northern California, and the New York Metropolitan area to set up recovery programs. From these hospitals, MedShare accepts donations of unused, unexpired medical surplus supplies and used biomedical equipment. Common supplies that MedShare often recovers includes supplies that have cosmetically-damaged packaging, production overages, usable returned products, and new products that a company wishes to donate to charity. Some of the most frequently visited areas that MedShare donates to include but are not limited to Haiti, Kenya, Nicaragua, and Guatemala (MedShare, 2022).

MedShare has collected over $240 million worth of life saving medical supplies from hundreds of hospitals and corporations. It has also diverted over 17.5 million pounds of quality and unused medical supplies from local landfills. MedShare supplies nearly 1,000 under-resourced free and safety net clinics in marginalized communities. Additionally, it has equipped over 930,000 healthcare professionals in underserved communities around the world and it has built an annual volunteer corps of 16,000 humanitarians. MedShare helps strengthen local health systems and serves over 32 million people in medically underserved areas around the world. It has delivered lifesaving medical supplies and equipment to 117 countries and territories worldwide. MedShare has provided biomedical training and support to over 7,000 healthcare professionals in 18 countries on equipment delivered to healthcare facilities around the world (MedShare, 2022).

Project CURE is another donation organization (Project CURE, 2022). It was founded in 1987 in Colorado to address the staggering shortage of medical resources around the world. Since then, Project CURE has become the world's largest distributor of donated medical supplies serving the sick and dying in more than 135 countries. Each week Project CURE delivers approximately three to five semi-truck-sized ocean containers packed with the medical supplies desperately needed to save lives in hospitals and clinics in resource-limited countries. Additionally, each year hundreds of healthcare professionals travel with Project CURE to provide medical treatment to communities in need and training to those dedicated to serving them. Project CURE is supported by over 30,000 volunteers annually and operates distribution warehouses in seven U.S. cities. In addition to its global work, Project CURE pivoted in response to the COVID-19 pandemic to support the need for personal protective equipment (PPE) and medical supplies in the United States. Since March 2020, Project CURE has provided more than 15 semi-truck loads of medical aid, which equates to 4.5+ million pieces and nearly $5 million of PPE for healthcare workers and first responders.

Project CURE is funded by grants from the United States government, non-governmental organizations, and by the generous support of philanthropic foundations, corporate partners and individual donors from all walks of life. It is regularly ranked as one of the top charities in America by Charity Navigator, The Chronicle of Philanthropy, Guidestar, Top Nonprofits, and Forbes Magazine, and it was recently given a Top COVID-Relief distinction by Charity Navigator (Project CURE, 2022).

Overall, the collection and repurposing programs REMEDY and RUMS and the donation programs MedShare and Project CURE have each demonstrated positive results. If more healthcare systems were to adopt these methods, there may be a dramatic decrease in the disposal of unused medical supplies. This has the potential to save medical facilities money as well as decrease the amount of waste in landfills productively helping the environment.

Methods

In an attempt to further the issue of the disposal of unused medical supplies and to investigate some additional remedies, the Vice President of Supply Chain Management at

Henry Ford Hospital in Detroit, Michigan, Mr. Joe Pettinato was interviewed. Mr. Pettinato previously worked in supply chain management at other hospitals, including Ascension Health, Trinity Hospital, and Beaumont Hospital. Mr. Pettinato is responsible for implementing procurement policies, developing operations playbooks, and developing outlined standards for the supply chain team at Henry Ford Hospital. Mr. Pettinato was asked the following questions and provided the noted answers:

1. What unused supplies are disposed of the most? Mr. Pettinato responded by stating that the most impactful unused supplies that have a long history of being disposed of are general medical supplies used in inpatient settings.
2. What items that are not allowed to be disposed of? Mr. Pettinato responded by saying that many items cannot be disposed of because of what their regulations are for disposal. He gave examples of chemicals, fluids such as saline, and drugs. Supplies such as these have to be disposed of in certain containers no matter the state of them.
3. What are the most costly items that are disposed of? Mr. Pettinato responded with physician preference items.
4. What are the cost implications of disposing of unused medical supplies? Mr. Pettinato didn't know the exact numbers, but he mentioned that along with the costs of the supplies that are disposed of, the storage costs of these supplies also contribute to the issue. Mr. Pettinato also said that supplies purchased during Covid were purchased in large quantities due to usage which created storage issues.
5. What units of the hospital dispose of the most unused medical supplies? Mr. Pettinato stated that a large percentage of the waste is surgical waste from the physician preference items and inaccurate surgical cards, which state the supplies needed for certain surgical procedures. He said items asked for by a surgeon are not always used and result in the product being returned to the shelves, when the product is handled so many times the packaging breaks down and the supplies must be disposed. Additionally, he stated that the waste that comes from the nursing units are usually all low cost products and not largely impactful as compared to items used in surgery.
6. Are there any methods that are currently in place at the hospital to help with the issue of unused medical supplies? Mr. Pettinato responded by saying that expired supplies are collected at each of the sites. However, he noted that the hospital recently donated disinfectant wipes to other hospitals and companies in need of them. The hospital also sells and/or donates its expired products to facilities that accept them. He pointed out that many places do not want to accept expired products so the hospital ends up disposing of them or selling them to WestMed, Centargin, and live auctions likely for education purposes. When asked about the advantages of donating or selling some medical supplies, he stated that it is mainly just a simple solution for disposal and the hospital does not generate much revenue from it.

Repurposing and Donating Barriers

It is apparent that while there are alternative methods in place to help address the issue of medical supply disposal, there are barriers that need to be addressed in order to improve existing methods to be more sustainable and effective. First, while the sterilization of unused medical supplies with hydrogen peroxide vapor is cost effective and environmentally friendly, research has found that multiple desterilizations might compromise device integrity (Rosenblatt & Silverman, 1992). For example, most surgical supplies are made of metal and metal is known to be susceptible to electrochemical corrosion. Microscopic pictures of desterilized metal medical supplies show wear and fatigue on the item (Alfa, 2012). Therefore, if a facility decides to desterilize its unused medical supplies, it runs the risk of the supply not functioning properly, which can put the patient at risk of injury or illness.

Collection and repurposing programs, such as REMEDY and RUMS, and donation programs, such as MedShare and Project CURE, run the risk of having issues with interference from implied warranty. For example, a manufacturer's warranty may no longer be valid upon reuse of single-use products (Rosenblatt & Silverman, 1992). Therefore, collection and donation programs must take the extra time to make note of this to the recipients of the medical supplies. For instance, REMEDY tells its recipients that the hospital offers these goods for donation "as is," and with no express or implied warranties, including no implied warranty of merchantability (Rosenblatt & Silverman, 1992). REMEDY also states that these goods are being disposed of by the hospital as a casual sale or donation and not as part of the hospital's normal business of providing medical services (Rosenblatt & Silverman, 1992). Additionally, many of the unused medical supplies that are disposed of in the so-called wasteful procedures are based on concerns about patient care and medicolegal issues (Rosenblatt & Silverman, 1992). While saving costs in healthcare systems and improving the environment are important, the health and safety of patients is even more important and therefore, guidelines concerning the safety of patients must be followed even if it is wasteful.

Another barrier with collection and donation programs is the fact that the sterilization of supplies being donated happens before they are sorted by volunteers in order to ensure the safety of the volunteers. Following the sorting of supplies, they are packaged to be shipped to their destinations and they are not desterilized. Therefore, the recipient of the donation has to be responsible enough to desterilize the supplies upon receipt. The recipients also need to be responsible enough to establish standards and determine whether a given product meets such standards (Rosenblatt & Silverman, 1992). This can be an issue because many of the medical facilities receiving these supplies do not have the money, resources, or education necessary to desterilize the donated supplies. If the donated supplies are not desterilized, medical facilities run the risk of cross contamination between the donated supplies and patients the supplies are used on, which puts the patient at risk of illness or injury.

An additional barrier includes how overburdened charities such as MedShare and Project CURE are with unrequested, unusable, and unsortable donations. This can be attributed to the fact that most guidelines for donation of unused medical supplies are lacking (Rosenblatt & Silverman, 1992). Therefore, programs such as REMEDY and RUMS will send their collected supplies without considering what the charities they are donating to are seeking. This causes charities to have an excess of supplies that cannot be donated to impoverished countries due to regulations, liability issues and safety concerns, which leads to charities disposing of those excess supplies. This completely counteracts the goal of trying to redirect the disposal of unused medical supplies.

Medical supplies are easily accessible and sparsely regulated in hospitals, which also serves as a barrier to disposal (Ghersin et al., 2020). For instance, some units of hospitals do not have a padlock on its supply room door. Therefore, anyone can go into the room at any time and take any supply. While most units do have a padlock on its supply room door, the staff that have access to the room will often take more than the necessary amount of supplies needed. These supplies end up adding up in a patient's room and then the supplies are typically disposed of unused once the patient is discharged because of potential contamination.

There is also a fallacy that individual action will not create a measurable difference (Ghersin et al., 2020). Therefore, trying to convince healthcare workers to only take the necessary supplies out of the supply room is difficult because no one is there to regulate how many supplies they take and most healthcare workers do not think that what they do as a singular person will have enough of an effect to change anything. This is a barrier because it proves that it will be extremely difficult for healthcare systems to implement a plan to combat the issue of unused medical supply disposal since they can not force healthcare workers to do work outside of their job description. In other words, no one wants to spend the extra time it would take to communicate with their peers to figure out methods that would help to solve this issue.

Discussion

The barriers to the existing methods used to combat the disposal of unused medical supplies are a good starting point to figuring out what needs to be done next. First, the observation that metal medical supplies become more fatigued with multiple sterilizations suggests that a study should be completed comparing and contrasting different disinfectants and the effects they have on the metal medical supplies that undergo multiple desterilizations. For instance, on one metal medical supply hydrogen peroxide vapor could be used to sterilize with subsequent sterilizations using ethylene oxide. These are both common disinfectants following standardized safety procedures developed to prevent metal fatigue from appearing.

There is not much that can be done to get around the warranty and regulation issues of the disposal of unused medical supplies since procedures and regulations have the safety of the patients in their best interest. However, if healthcare workers receive training regarding the disposal of unused medical supplies and the costs incurred by the hospitals and the environment, it might make workers more apt to only take out what they need from the supply room. For example, upon hiring a healthcare worker could be required to watch a video that explains how harmful the disposal of unused medical supplies can be. This should ensure awareness of the issue and possibly encourage employees to equip facilities with only necessary supplies. By doing this, the unused medical supplies that have to be disposed of due to liability and regulations will decrease since only the needed supplies are being used.

Once the unused medical supplies are donated and received by the medical facilities in need, one might believe that the programs in place, such as REMEDY, RUMS, MedShare, and Project CURE, will be successful. However, if the supplies received are not sterile, it will actually put these medical facilities at greater risk. Therefore, once the unused medical supplies are sorted by volunteers, they should go through an additional sterilization process. This could be done by placing the supplies in an already sterile room and then using disinfectants such as hydrogen peroxide vapor or ethylene oxide. As previously mentioned, sterilization can be a low cost procedure and it will ensure that the supplies are able to be safely used once they reach their final destination. Putting the responsibility of desterilization on medical facilities that barely have the basic supplies needed to provide healthcare is improbable. However, adding this step will reduce the risk for contamination between donated supplies and patients.

The overabundance of unused medical supplies provided to charities such as MedShare and Project CURE from collection and repurposing programs such as REMEDY and RUMS is counteractive. Improved communication between the programs offers a potential solution to this problem. For example, if MedShare were to receive medical supplies that it cannot donate, it could directly inform the programs of REMEDY and RUMS so they can make note of it and make sure to not send them until the need arises. Eventually, this should lessen the amount of incorrect supplies sent to charities preventing the accumulation of unused medical supplies and their disposal at the charities.

A possible solution to help regulate the amount of supplies taken out of hospital supply rooms is to have a volunteer in place near the supply room. Since it would be a volunteer, there would be no cost to this method. The volunteer could be given a standardized list of supplies needed for certain procedures. Then, when a healthcare worker needs supplies, the volunteer can provide only the necessary amount of supplies. This would not only reduce the disposal of unused medical supplies, but it could also increase efficiency. For example, it can be difficult to find supplies in the supply room since there are so many options available. Therefore, if one volunteer becomes familiar with the supply room, they will be able to provide the necessary supplies in a shorter amount of time compared to the healthcare worker.

Healthcare systems cannot force healthcare workers to do things outside of their job description. Therefore, getting healthcare workers on board with collecting unused medical supplies or decreasing the amount of supplies taken out of the supply room can be challenging. A solution to this issue could be to provide an incentive. For instance, the healthcare system could set a goal of the amount of unused medical supplies it hopes to collect and donate. Or they could set a goal of the amount of money it hopes to save that month. If the goal is reached, an incentive for healthcare workers at that hospital could be provided. While there is some cost associated with possible incentives, the costs associated with the disposal of unused medical supplies may be greater and it affects the hospital and environment in a detrimental way.

Conclusion

The disposal of unused medical supplies is an issue that should no longer be ignored. While the desterilization of unused medical supplies with hydrogen peroxide vapor and collection, repurposing, and donation programs such as REMEDY, RUMS, MedShare, and Project CURE are all good starts to address this issue, there are still many barriers that need to be resolved to completely fix the issue. By looking at each of these barriers, more effective and efficient solutions may be created. Through the creation of new methods and improvements to the existing ones, hospitals in the United States can save money, hospitals in impoverished countries can be provided with urgently needed supplies to provide proper healthcare, and there will be less waste built-up in landfills positively impacting the environment.

References

Project CURE. (2022). *About project C.U.R.E: Donate Medical Supplies & Equipment*. Retrieved October 19, 2022, from https://projectcure.org/about-us/

MedShare. (2022). *About Us*. Retrieved October 19, 2022, from https://www.medshare.org/about-us/

Alfa, M. J. (2012). *the 'pandora's box' dilemma:* reprocessing of implantable screws and plates in orthopedic tray sets. *Biomedical Instrumentation & Technology*, *46*(sp12), 55–59. https://doi.org/10.2345/0899-8205-12.1.55

Cuncannon, A., Dosani, A., & Fast, O. (2020). Sterile processing in low- and middle-income countries: An integrative review. *Journal of Infection Prevention*, *22*(1), 28–38. https://doi.org/10.1177/1757177420947468

Ghersin, Z. J., Flaherty, M. R., Yager, P., & Cummings, B. M. (2020). Going green: Decreasing medical waste in a pediatric intensive care unit in the United States. *The New Bioethics*, *26*(2), 98–110. https://doi.org/10.1080/20502877.2020.1767916

Miller, Steve. (2002). The REMEDY Program: Recovery and Donation of Unused Surgical Supplies to Developing Countries. *Einstein Q J Biol Med*. 19, 121-124.

Muldoon, L. B., Chan, W. W., Sabbagh, S. H., Rodriguez, R. M., & Kanzaria, H. K. (2019). Collecting unused medical supplies in emergency departments for responsible redistribution. *The Journal of Emergency Medicine*, *57*(1), 29–35. https://doi.org/10.1016/j.jemermed.2019.02.009

Otter, J. A., Nowakowski, E., Salkeld, J. A., Duclos, M., Passaretti, C. L., Yezli, S., Ross, T., Carroll, K. C., & Perl, T. M. (2013). Saving costs through the decontamination of the packaging of unused medical supplies using hydrogen peroxide vapor. *Infection Control & Hospital Epidemiology*, *34*(5), 472–478. https://doi.org/10.1086/670210

Otter , J., Perl, T., Carroll, K., Ross, T., Yezli, S., Passaretti, C., Duclos, M., Salkeld, J., & Nowakowski, E. (2020, November 14). Decontamination of unused medical supplies reduces healthcare costs. *Infection Control Today.* Retrieved October 17, 2022, from https://www.infectioncontroltoday.com/view/decontamination-unused-medical-supplies-reduces-healthcare-costs

Pennino, R., Mayer, A. M., Dahn, A. T., & Husser, W. (1994). Recycling unused medical supplies. *Plastic And Reconstructive Surgery*, *94*(2), 397–399. https://doi.org/10.1097/00006534-199408000-00031

Rosenblatt, W. H. (1992). Recovery, resterilization, and donation of unused surgical supplies. *JAMA: The Journal of the American Medical Association*, *268*(11), 1441. https://doi.org/10.1001/jama.1992.03490110079033

Rosenblatt, W. H., & Silverman, D. G. (1994). Cost-effective use of operating room supplies based on the remedy database of recovered unused materials. *Journal of Clinical Anesthesia*, *6*(5), 400–404. https://doi.org/10.1016/s0952-8180(05)80010-8

Wan, E. L., ScM, L. X., Barrett, M. D., Baltodano, P. A., Rivadeneira, A. F., Noboa, J., Gentry, C., Palacios, J. H., Rosson, G. D., & Redett, R. J. (2014). Global Public Health Impact of recovered supplies from Operating Rooms: A critical analysis with national implications. *Journal of the American College of Surgeons*, *219*(3). https://doi.org/10.1016/j.jamcollsurg.2014.07.130

World Health Organization. (2022). *Medical supplies*. World Health Organization. Retrieved November 18, 2021, from https://www.who.int/health-topics/medical-devices#tab=tab_1.

What is hydrogen peroxide: Chemdaq. ChemDAQ Incorporated. (2022). Retrieved October 18, 2022, from https://www.chemdaq.com/hydrogen-peroxide/what-is-hydrogen-peroxide

39.

The effect of the cybersecurity implementation on hospital's efficiency and financial performance

Jeong Hoon Choi [1], Christopher Lee [2]

[1]Youngstown State University, Youngstown, USA. [2]Central Connecticut State University, New Britain, USA

Abstract

This study examines the impact that cybersecurity implementation has on efficiency and financial performance (ROA) of a hospital. The U.S. hospitals are classified into two groups – cybersecurity implemented group and no cybersecurity implemented group. A multifactor authentication (MFA) serves as a proxy variable to measure cybersecurity implementation. This study employs data envelopment analysis (DEA) models and the DEA model's objective function value serves as a proxy to measure hospital efficiency. Using the data compiled from different sources, this study finds that the implementation of cybersecurity has a positive impact on both efficiency and ROA. This study also measures the impact of some control variables such as hospital size and location and finds that larger hospitals in the metropolitan area show positive impacts on efficiency and ROA.

Categories

Healthcare

41.

Linking government technology incentives and operational risks and performance

Rajib Dutta [1], David Dobrzykowski [1], Randy Bradley [2]

[1]University of Arkansas, Fayetteville, USA. [2]University of Tennessee, Knoxville, USA

Abstract

While most hospitals have adopted Health Information Technology (HITs) that comply with recent government incentives, how these programs have impacted hospitals' investments across different technology categories remains unclear. This is important because different types of HIT influence different operational performance measures. Additionally, different HITs may be vulnerable to different levels of operational risks. Data breaches represent a major operational risk stemming from HIT that can have a costly negative impact on hospitals. We adopt a complexity perspective to investigate how HIT incentives influence hospitals' HIT adoption (Clinical and Augmented HIT), and ultimately their impact on Total Performance Score (TPS) and healthcare data breaches. We analyzed archival data from four publicly available sources using 2SLS, seemingly unrelated regression, and logit regression to examine 141 data breaches transpiring during 497 hospital years. Results show that incentive payments lead hospitals to invest more heavily in Clinical HIT which helps to reduce the magnitude of healthcare data breaches and also improves TPS. However, while Augmented HIT increases the magnitude of data breaches, it improves TPS. Post-hoc analyses results provide additional insights on the longer-term effects of these HIT Bundles on hospital performance. We provide insights into how hospitals react strategically to regulatory pressure and financial incentives, and how these incentives impact hospitals' IT infrastructural decisions. Next, we show how these HIT bundles differentially impact operational performance and data breaches. The overall implications and policy recommendations are discussed.

Categories

Healthcare

Information Systems

30.

Reviewing Social Engineering Theories through Contrasting Three Case Studies

Te-Wei Wang [1], Xin Chen [2]

[1]University of Illinois Springfield, Springfield IL, USA. [2]Columbus State University, Columbus GA, USA

Abstract

Online scam is a wildly studied topic. Many researchers and organizations have proposed frameworks and strategies. Nonetheless, the number of online scams keeps increasing. Online scammers change their methods and stories so that older warning signs do not work. This presentation compares and contrasts three recent online scams using the combination of Instagram, WhatsApp, WeChat, and LINE in the last 12 months. We first review current literature document types and prevention strategies for online scams. Then, we verify these strategies through case studies. Finally, we will discuss future research directions with the audience.

Categories

Information Systems

36.

Cybersecurity measures from new technology users' perspective

Hongjiang Xu [1], Mark Hwang [2]

[1]Butler University, Indianapolis, USA. [2]Central Michigan University, Mt. Pleasant, USA

Abstract

In the cloud computing environment, cyber vulnerability is a critical issue that poses a threat to organizations worldwide. Modern economies and businesses rely heavily on new technology for daily operations, making the need for effective cybersecurity measures more critical than ever before.

To protect information and systems from cyber-attacks, businesses allocate a significant portion of their annual technology budgets to security measures. An example is measures that deal with software vulnerability, flaws within a software product that can lead to work contrary to its documented design, which can be exploited to cause the system to violate its documented security policy. Studies have shown that software vulnerabilities can have a widespread impact, resulting in significant economic and noneconomic damage to firms.

The annual costs of all malicious code attacks to businesses worldwide run into billions of dollars and continue to increase each year. These costs include additional elements such as lost business revenues, cleanup costs, productivity losses, overtime costs, and damage to reputation and trust.

In addition to software, users play a critical role in ensuring cybersecurity. As more users become more dependent on network-enabled services for communications, social networking, and web-based transactions, the number of cyber-attacks and the ease with which they occur have increased, putting more systems at greater risk. Therefore, prioritizing cybersecurity measures to help users navigate the new technological landscape is crucial for securing today's assets. This paper aims to examine the human aspects of cybersecurity, including user characteristics and user perception of cybersecurity measures.

Categories

Information Systems

55.

The Pioneer of Modern Social Media: AOL Instant Messenger (AIM) - Insights Into How AIM Changed The Way People Communicate

Mark Donner, Hongjiang Xu

Butler University, Indianapolis, USA

Abstract

Social media has a big impact on today's society. Nearly everybody and all the organizations use social media for communication, marketing, promotion etc. Social media provides the means of interaction among people and organizations to create, share, exchange or promote information and ideas through social networks and the Internet. Many of the functions of social media originated from AOL Instant Messenger (AIM) that started 1997.

AIM revolutionized the way people communicated online. Launched without any fanfare or advertising, AIM quickly gained popularity and transformed internet communication. Despite the existence of other messaging products at the time, AIM became the leading consumer internet instant messaging product within a year of its launch.

In this paper, we are going to discuss the development of AIM, and how AIM changed the way people communicate through the Internet, which lead to many of the functions and features that are used in modern social media.

In addition to its instant messaging, chat and friends-list, AIM was the first to introduce innovative communication features such as video chat, stock and news tickers, personal alerting, away messaging, bots, and file transfer, which kept it ahead of the competition for years. Perhaps even more significant was AIM's persona feature set which transformed mere texting and made the experience personal by enabling users to express their personality for the first time online through persistence presences, away notification, and unique client customization. All of which made AIM a true pioneer of modern social media.

Categories

Information Systems

63.

A Systematic Literature Review for Adoption of Big data in Firms: An Exploratory Study Using Text Mining

Narges Mashhadi Nejad, Marcelo Alvarado-Vargas

University of Toledo, Toledo, USA

Abstract

The adoption of big data in firms has emerged as a critical area of research and practice in recent years. In this exploratory study, by using text clustering techniques, we aim to identify the top 10 topics in the literature related to this topic. For that purpose, we tabulated the information from more than 1000 academic articles and conference proceedings and applied the latest text-mining techniques. Furthermore, results for the most salient topics per journal category (i.e., journal quality) are presented. This study's contribution is to provide a framework for performing a systematic literature review for the specific current academic topics combined with the latest technology available.

Keywords: Big Data Adoption, Text Mining, Systematic Literature review, Text Clustering

Categories

Information Systems

67.

The Learning from the Data Breaches Involving 30,000 or More Records

Xiaodong Deng

Oakland University, Rochester, USA

Abstract

Organizations have increasingly paid attentions to securing their valuable digital assets However, security incidents such as data breaches continuously occur or have been identified, jeopardizing the organizations' reputation in their data collection or utilization. The website at https://en.wikipedia.org/wiki/List_of_data_breaches maintains a list of data breaches over more than 15 years involving the theft or compromise of 30,000 or more records, based on various press reports, government news releases, and mainstream news articles. The website has also summarized the breaches by the involving entity, the year of the breach occurred, the number of records breached, organization type, the method that the records were breached, and the related referent sources.

By following the related referent resources listed at the website or some possible additional reports, this study moves one step further by exploring the nature of the breaches, where the breaches occurred along an organization's value chain, the key attributes of the breached records, how the breaches were identified, and what were the consequences of the breaches.

The summary may help organizations better develop their cybersecurity strategies and allocate resources to their cybersecurity initiatives. The summary may also help highlight the areas that warrant more research efforts to identify factors that could help alleviate organizations' data breaches.

Categories

Information Systems

71.

Exploring the Landscape of Cloud Security: A Systematic Literature Review, Content Analysis, and Prospective Research Questions

Sayed Abu Sayeed, Omar Faruq Osama, Samiul Alam

Emporia State University, Emporia, USA

Abstract

The advent of cloud computing has marked a significant shift throughout the IT world, bringing numerous benefits including dynamic resource allocation, elasticity, global network connectivity, rapid adaptability, and on-demand self-service. Due to its scalability, flexibility, and cost-effectiveness, cloud computing has become an essential module of the current information technology infrastructure. While cloud computing has expanded precipitously, it has also caused several new studies. This study uses Web of Science-indexed journals to find pertinent papers and conducts content analysis to comprehensively examine the space. This study intends to contribute to the knowledge of cloud security by recognizing literature gaps, potential research questions, existing trends, and repercussions. This study examines the recurring topics and academic debates in the field of cloud security. Many cloud computing security frameworks, such as ISO 27017 and NIST SP 800-210, have been developed. These frameworks were created to provide guidelines and best practices for protecting cloud computing systems, and various businesses and industries have adopted them. This research also compares ongoing research trends with various cloud security frameworks.

This paper contributes to the subject of cloud security by offering a complete assessment of the existing state of knowledge, identifying literature gaps, and proposing future research issues. The findings of this study have relevance for both practitioners and researchers, as they shed light on the existing and future issues cloud security faces.

Categories

Information Systems

72.

Deep Learning Classification Modeling To Measure Online Classroom Engagement

Jason Davidson

Butler University, Indianapolis, USA

ABSTRACT

Student enrollment in online courses has nearly tripled over the last decade, with 72% of college students participating in at least one online course. There are many advantages to online education such as increased classroom diversity, the reduction of geographical limitations, and overall convenience. However, studies have shown students participating in online courses underperform when compared to their counterparts in traditional face-to-face classroom settings. Reasons for these discrepancies include limited supervision, lack of resources, feelings of isolation, and low engagement. The challenge of student engagement is exacerbated by educators' reduced ability to visually assess students' nonverbal cues of the affective states (engagement, boredom, confusion, frustration). This research seeks to address the challenge of reduced nonverbal communication by automating student engagement recognition with computer vision. Modern deep learning techniques will be used to create a convolutional neural network to measure the affective states of students in a synchronous online environment. Additionally, this project will explore the correlation, if any, between the affective states and seven universal emotional states of anger, contempt, disgust, fear, joy, sadness, and surprise. The ultimate goal is to develop a real-time monitor that indicates to online educators the level of each student's engagement during a synchronous online course.

Categories

Information Systems

Innovative Teaching & Education

52.

Multi-Modalities and Instructional Strategies for Teaching Information Systems Courses Leveraging SAP HANA and DbSchema

Dr. Kevin Barrons, Grand Valley State University

Dr. Asli Yagmur Akbulut, Grand Valley State University

Christopher Gillespie, MA, Grand Valley State Univeristy

Abstract

Enterprise Resource Planning (ERP) systems in general and SAP systems in particular are widely used by organizations of all sizes, across a variety of industries. Business students graduating with SAP skills are highly in demand because of the widespread use of SAP systems. Therefore, both conceptual and hands-on practical experience with SAP sytems provides students with numerous benefits in the job market. Currently, SAP education mostly focuses on system configuration and processing business transactions in the system. Our research discusses the use of the most recent version of the software, SAP S/4 HANA system, in an innovative and quite uncommon way: to teach advanced database management concepts and tools to students in an introductory level management information systems course.

Teaching students database management using SAP systems empowers them with both the SAP skills and database skills at the same time. Students understand and experience the use of SAP systems beyond conducting business transactions as they deep dive into the powerful and unique multi-model SAP HANA database. Moreover, our study combines the SAP HANA database with another powerful software tool, DbSchema, which is used for database design. Working with the DbSchema and SAP HANA, students gain experience of working with real-time data with access to advanced search and analytics capabilities. Using multiple cutting edge software tools, supported by a multi-modality approach as part of the instructional strategy promotes active learning, higher levels of engagement, and higher student success. Multi-modality approaches, including kinesthetic, auditory, and visual stimuli, help students process and retain the information learned more effectively.

Keywords: Information systems education, database management, database design, SAP HANA, DbSchema, multi modalities, innovation, kinesthetic, visual, and auditory learning, andragogy, instructional design

Introduction:

This paper is organized into sections to thoroughly address the complexity of using the SAP Hana Database and DbSchemas in a traditional class and hybrid learning model for MIS competencies.

The Zone of Proximal Development (ZPD) and experiential learning theory, Vygotsky (1978), is applied to teaching the SAP Hana Database and DbSchemas in a Hybrid Learning Model for MIS Competencies. The first part introduces the topic and selected definitions of the terms used in the discourse. In this discourse, two questions will be proposed in the second part. 1) How does the ZPD create value for the learner and the organization? 2) How does the shared experience enable learning? Finally, in the last part, a conclusion is drawn, mentioning the limitations of the arguments provided in the debate. A suggestion is also made for future research in "hybrid learning" instructional pedagogy.

An Overview of Selected Learning in the classroom:

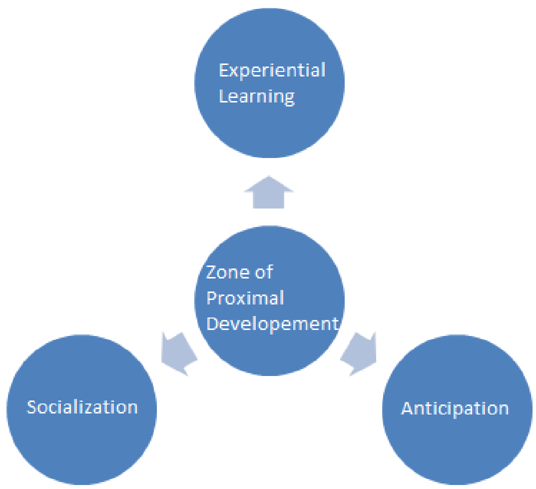


Figure 1: Zone of Proximal Development (Vygotsky, 1978)

The "Hybrid" approach is applied during the MGT 268 Introduction To Management Information Systems instructional delivery at GVSU located at the Pew Campus in Grand Rapids, Michigan. Much success with the course in the past has applied a traditional instructional method of a class with separate lecture and computer lab time. Courses in management information systems are still taught using the traditional computer lab and lecture model. In addition, the "Hybrid" classroom" model allows more time for students to problem solve in teams, review class lectures outside of class with pre-recorded slides, sound, and videos, and use valuable class time completing class exercises. Seeing more than 50% of the focus of the course requires weekly hands-on exercises applying the integrated business processes with enterprise resource planning (ERP/SAP) software, students are allowed more in-class time for problem-based learning (PBL) activities related to each weekly activity. In addition, students had required reading assignments, quizzes, and instructor-prepared recordings that were reviewed before class and applied to each weekly exercise. This "hybrid classroom" model allows students to work independently and in teams to complete class assignments to enable them to achieve mastery at their learning pace. Further, when students come to class, they spend more time working with the instructor (subject matter expert), applying their outside class activities to higher-level learning activities. Therefore, the "hybrid classroom" model can effectively use the theoretical approach to "The Zone of Proximal Development."

*The Z*one of Proximal Development (ZPD) inside the classroom setting or cyberspace continuous learning environments has borrowed practices and techniques from Vygotsky (1978), Bandura (1977), Senge (1990), Lave & Wenger (1991), Brown & Duguid (1991), Jenkins (2009), & SAP.Com(2015), among many others.

Vygotsky (1978) suggests that the Zone of Proximal Development (ZPD) can be adopted to develop the skillfulness of learners in different learning environments such as MIS. In the ZPD, Vygotsky propounds that competencies can be enhanced by the expert(s) demonstrating the "how-to" in developing competencies. In the ZPD, Vygotsky explains the novice's three stages of skills development. The lower limit of ZPD is the level of skill reached by the learner working independently. The upper limit is the level of potential talent that the learner can achieve with the assistance of a subject matter expert.

In Figure 1, the centricity of the ZPD is illustrated to be surrounded by three equal circles. If the circles are misaligned, the ZPD will not be genuine, and the potential will not be achieved. Vygotsky also refers to a community of practice. The ZPD develops concepts that can be applied to understand job specifications and the competencies required to do specific jobs.

Bandura (1977) posits that the socialization process must be understood for learning. The socialization process constitutes the learner's environment at work, at home, and in social contexts. Bandura argues this environment is critical for successful skills and language acquisition. According to Bandura, he discusses the importance of social learning theory in skills acquisition, where people learn through observing others. As MIS students imitate the teacher, this can be applied to the concept of socialization and learning. In the MIS area of study, the students develop the MIS confidence and increase self-efficacy to use their skills to execute a set of courses of action.

Roehl, A., Reddy, S.L. & Shannon, G.J. (2013) suggest various ways to use technology to free class time from the lecture. Indeed, increasing learning activities during class time improves teacher-to-student mentoring, peer-to-peer collaboration, and cross-disciplinary engagement. Most recently, the Flipped Classroom has been approached in various formats. This may include self-learning packages, problem-based learning, project-based learning, pre-recorded video lectures, and demonstrations for personal review. It is essential to focus on the level of knowledge transfer when one identifies the ZPD in this discourse. The Flipped Classroom is a step ahead of the ZPD model because of the emphasis on experiential learning, problem-based learning, and the relationship to solving problems in social contexts.

The 4 Step Method of Program Development for SAP Hana Database and DbSchemas in a Hybrid Learning Model for MIS Competencies.

The process for creating this exercise would include the following:

Step 1

Create test fields and tables in DbSchema

Step 2

Connect DbSchema to the SAP Hana Database

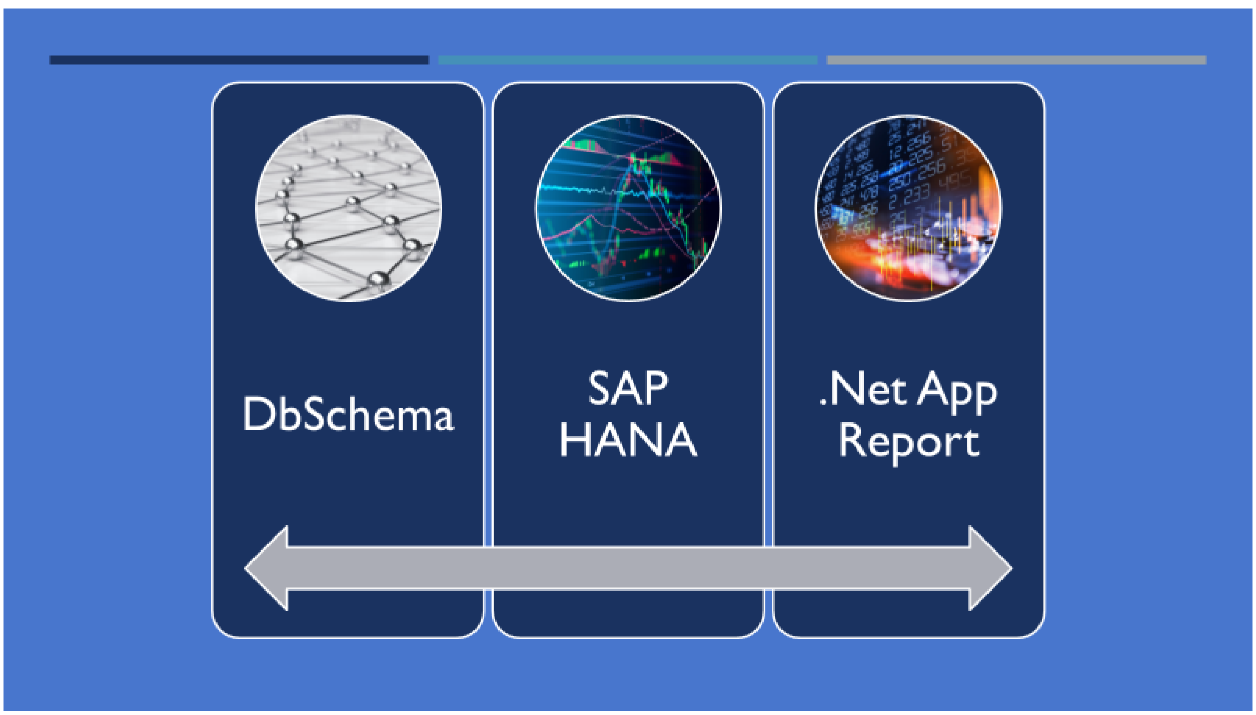
Step 3

Develop Schema Structure for students

Step 4

Execute the .Net app to input and review the data. (Run report to view the data)

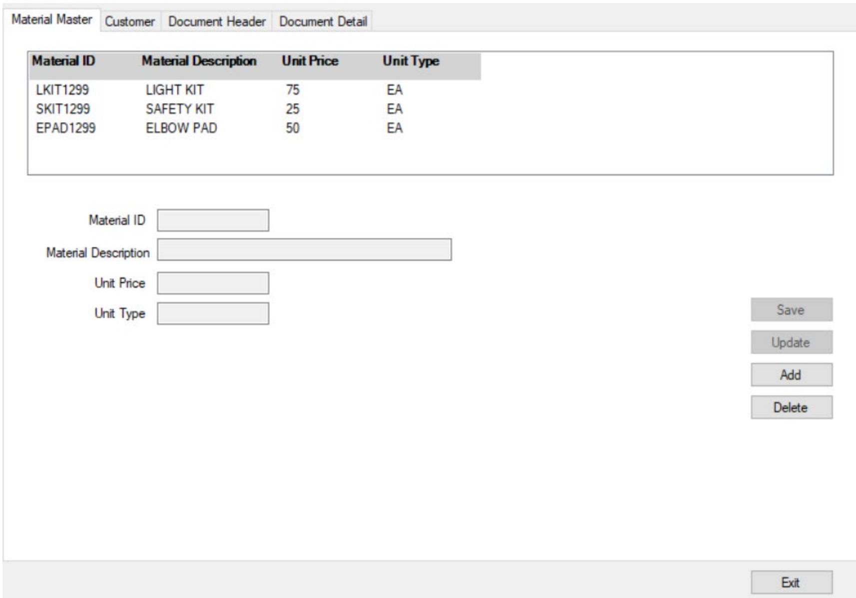
The following illustrates the connection between the DbSchema Program, SAP HANA, and MS Excel. Student activities include a two-parts storing the data in SAP HANA, then using DbSchema to create the tables, fields, and data, and running an MS Excel Report.



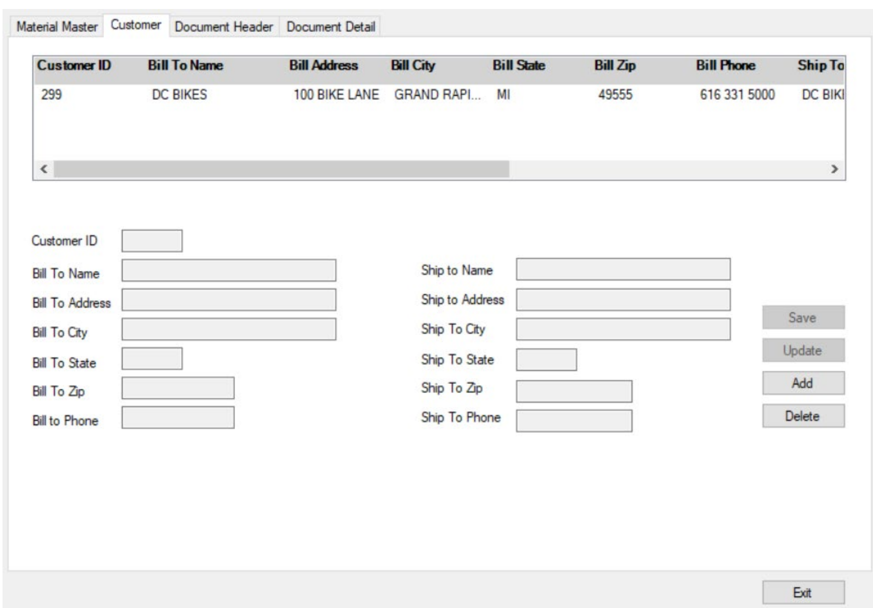
DbSchema Process Overview

After the first trial, modifications were made to the DbSchema project to provide efficiency to the exercise. Given the complexity of the assignment for introductory students, templates were created for the various tables. The four templates were included as follows:

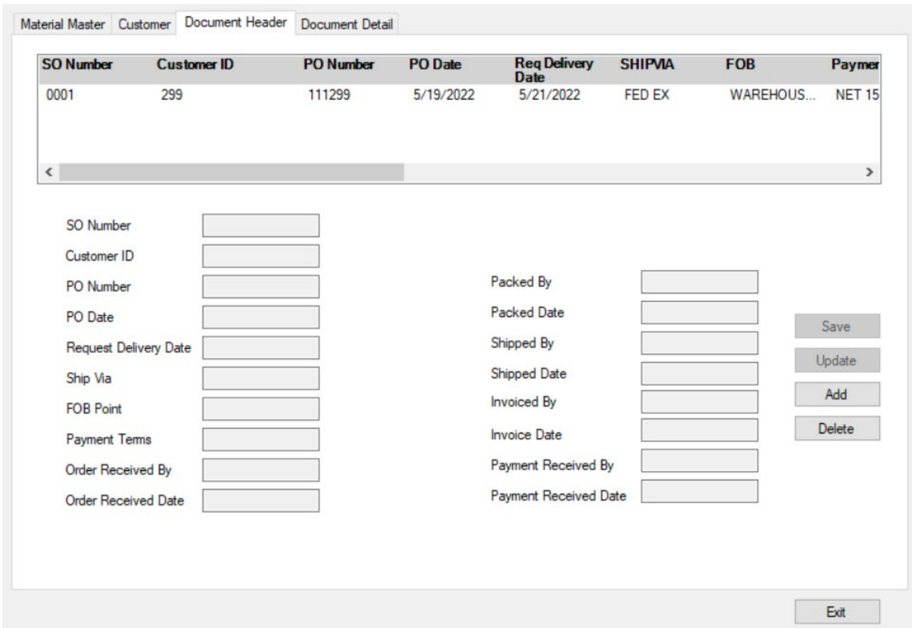
Template 1 – Material Master



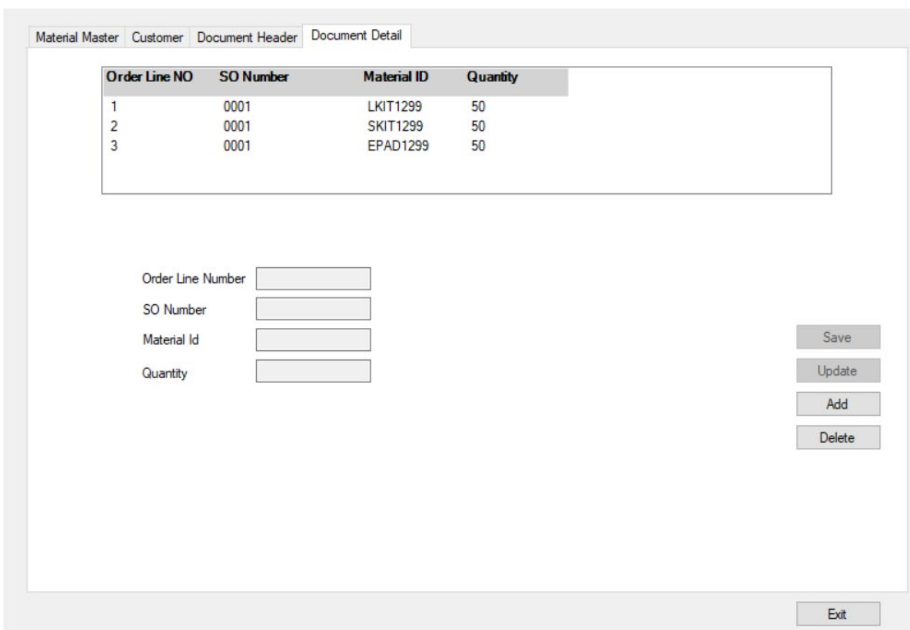
Template 2 – Customer Master



Template 3 – Document Header



Template 4 – Document Detail



An Overview of Pedagogical Perspectives With SAP Hana Database and DbSchemas in a Hybrid Learning Model for MIS Competencies:

An overview of current pedagogical perspectives has been drawn from a careful selection of literature to answer the questions: 1) How did the SAP software development start? And 2) How does the shared experience enable learning?
In 1972, creating a corporate software company was unheard of. However, SAP's five founders refused to let that stop them. Instead, they set out on a path that would transform the world of information technology and forever alter the way companies do business. Since then, SAP has moved on to what it is today to deliver, cloud solutions, enterprise resource planning, banking, customer relationship management, human resources, retail, and mobile solutions. According to SAP (2015 a), the company is the world leader in enterprise applications in software and software-related service revenue. Furthermore, based on market capitalization, the company asserts that it is the world's third-largest independent software manufacturer, preceded by Oracle and Microsoft at the helm. SAP has more than 291,000 customers in 190 countries, more significant than 74,500 employees globally, a 43-year history of improvement and evolution as a true industry leader, and a yearly income exceeding 18 billion eurodollars (SAP 2015 b).

The concern of researchers Bandura (1977), Vygotsky (1978) & Senge (1991) has been focused on organizational learning & competence development. The research on skills transfer and knowledge acquisition in socially constructed learning environments has been expanded very well in Bandura, Vygotsky & Senge. In the case of the MIS approach, the skills transfer occurs in defined places such as; the workstation, the software, application usage, computer labs, and real business problems. Vygotsky (1978) argues that the ZPD is the space where the learner acquires desirable skills after consultation with the subject matter experts. Long after the experts have left, the novice continues to practice to make perfect their skillfulness. Bandura (1977) suggests that the environmental factors influence the learner to acquire knowledge and skills necessary to function in the ways the community desires. Bandura introduces a community of practice concepts positioned in the environment. This community can be the family, the school, and the workplace. Thus the socialization of the learner is critical to the acquisition of explicit practices.

Both Bandura and Vygotsky used children in their research. We are now applying the business model approach as the focal point of an older age group. Bandura and Vygotsky used their research with children and did not have a business organization. The ZPD context can now be focused on the context of students' learning in a business simulation in a computer lab in this situation. Bandura defined the socialization process as a continuum of learning from home to school and work. Thus, a student appreciates skills, knowledge, and attitudes from this experience. The same can be said where the child can develop attitudes at home that can be shared at school or in many social contexts.

Senge, Kleiner, Roberts, Ross & Smith. (1994) propound about learning techniques. They identify five learning modules that can be applied to progress organization. The first learning model is "personal mastery," which constitutes the skill sets an individual must possess in doing a given job. Personal mastery allows the learner to build the capacity to use their skills to produce the goods and services demanded by the customers. In the MIS area of study, students develop their confidence to use software such as SAP by constantly conversing and consulting with experts and sharing knowledge with other students. The second is "the mental model." This model includes reflection, clarification, and the ability to improve a shared window of the world. This model allows a community of practice to create, diffuse and distribute a co-constructed approach. Finally, Senge suggests that a "shared vision is co-constructed as well. A shared problem is a problem solved.

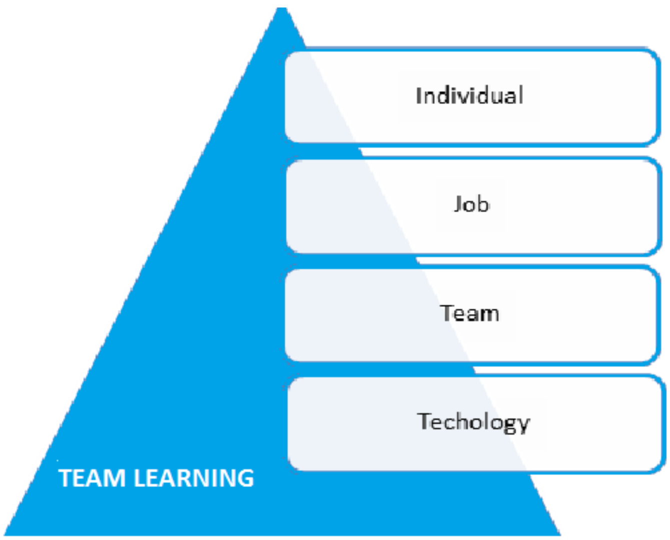


Figure 2: TEAM LEARNING (Senge et al.; 1994)

The fourth model is about "team learning." The shared learning concept leads us to Lave & Wenger (1991) about knowledge creation in distributed cognitive environments. They suggest that an organization can build a community of practice based on the notion of highly cohesive teams. They propound that a community of practice can be composed of a group of people preferably from the same organization that meets in defined places at agreed times or in cyberspace. A community of practice becomes a tool that can be deployed to enhance learning in the value creation process. The Hybrid Classroom embraces social interaction in the community of practice by having students problem-solve in groups. In the MIS course, students learning the SAP software use the exchange of business problems to increase knowledge and skillfulness in solving problems. Using the SAP Hana Database and DbSchemas in a Hybrid Learning Model for MIS Competencies is very practical in a team learning environment. This process continues to build as the course progresses until the assignment is complete. This problem-solving experience amongst the team members builds confidence in the learning process.

Senge et al. (1994) allude to the fifth learning discipline, "systems thinking," enabling the entire organization to operate on the same page. Because a community of practice shares the same language, it implies they are aware of the organizational mission. Each individual deduces his job description that resonates with the mission statement.

According to Scida, E. E. & Saury, R. E. (2006) on hybrid courses and their impact on student and classroom performance, the median grade for students was higher in the hybrid course compared to the traditional course in a case study at the University of Virginia. For example, in the hybrid course by comparison to the regular traditional course. In a collection of student performance data, 11 of 19 students (58%) earned As in the hybrid course versus 7 out of 22 (32%) in a traditional course. Three students in the hybrid course (16%) earned a C or below compared to 6 students (27%) in the traditional course. Overall, 84% of the hybrid course students earned a B or above compared to 73% in the traditional course.

An essential consideration for teaching the SAP Hana Database and DbSchemas in a Hybrid Learning Model for MIS Competencies is to apply "Active Learning Strategies" to maintain a high level of engagement. In the educational experience, the following ideas for strategies that benefit hybrid courses were used when providing instruction for SAP Hana Database and DbSchemas in a Hybrid Learning Model for MIS Competencies. These strategies provide meaningful learning activities and give students ways to think about what they are doing.

The centricity of Active Learning in a course:

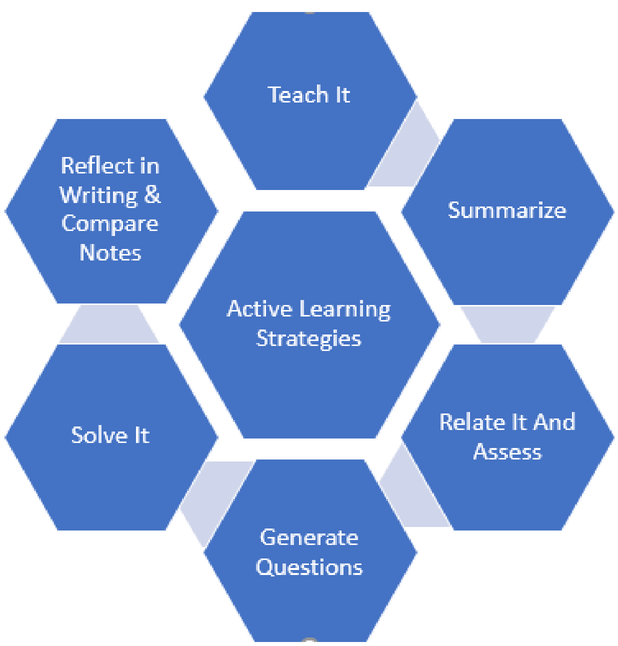


Figure 3: Active Learning Strategies

Appreciating models:

The classroom instructor must identify a pedagogical model that stresses the learning. This may include concepts such as project-based learning (PBL), problem-based learning (PBL), game-based learning (GML), understanding by design (UbD), and genuine literacy. Then the Hybrid Classroom model can be assimilated into any of these concepts. Lynch (2018) alludes to the fact that the above models could be effectively implemented in the Hybrid Classroom by a community of practice. The ZPD, on the other hand, has been described distinctly by Vygotsky (1978) to be the area where the novice acquires competence.

Simple Tools:

The community of practice needs to know how to use the simple tools available in the classrooms or within the learning community. Twenty-first-century students are highly engaged with the educational technology tools in the ZPD. The tools allow for higher motivation, stimulation, and comprehension. In addition, the technology enables students to interact among themselves, the virtual libraries, software, and instructors, which promotes higher-level thinking skills.

Replication of Competencies:

The use of technology and visual aids such as YouTube, pre-recorded videos, text, and PowerPoint presentations allow for reflection and enable students to benchmark competence in areas in which proficiency is sought. In the flipped model, students can review information materials before attending an in-classroom class which composes their prior knowledge before coming to class. In addition, some activities can be assigned and completed as they pertain to the instructional objectives. Flavell (1979) suggests shared experiences are any conscious cognitive or affective familiarities that accompany and pertain to intellectual exchange, which happens in flipped classrooms.

Time, Stage, and Location:

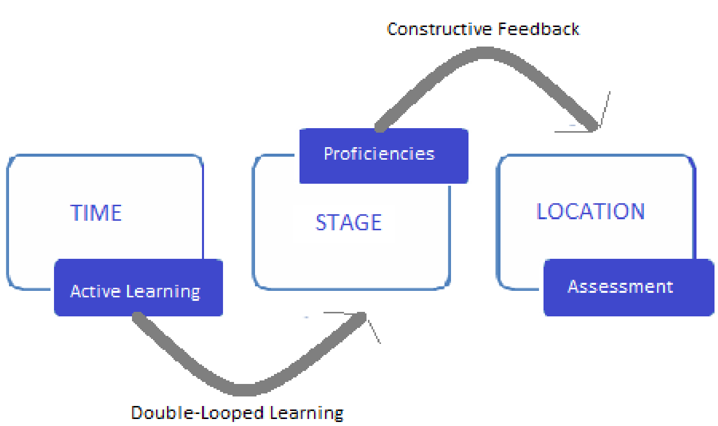


Figure 4: Double-Looped Learning (Argyris, 1991)

In Argyris (1991), double-looped learning is introduced into the organization as imperative to a successful business. We can apply this concept to the Flipped Classroom strategy. Both instructor and the learner receive constructive feedback (see Figure 4). Given the Flipped Classroom requires structure, the learning environment is essential. The instructor will need to incorporate proficiencies to keep the students engaged. Most learners reviewing information before coming to class will need accountability. Allowing active learning to occur outside of class, similar to hybrid or online learning techniques, will require a means of formative assessment to evaluate student performance. These assessments will assist the educational leader in addressing the needs of the learners. Meyers & Jones (1993) discuss the active learning process and instructional tools that keep the learning environment "active," including problem-solving exercises, cooperative student projects, informal group work, simulations, case studies, and role-playing. Argyris (1991) propounds that feedback is critical in business environments which would apply to classrooms. He refers to this as double-looped learning.

Epistemic Communities:

The epistemic communities or knowledge communities have been described in Foulcault (2002) as a group of people who meet regularly to create priorities in organizations. Foucault leads to the thought that an epistemic community co-constructs those priorities. Mupepi (2010) develops an organizational strategy hinged on an epistemic community. For example, a knowledge community can advance organizational efficiency and effectiveness. Mupepi gives the analogy of the Adam Smith pin production factory showing the division of labor. Adam Smith proposes that the division of labor into specialization units can yield increased productivity. By determining the desired capacity and skill set that captures what the organization does best using appreciative inquiry (AI), the entire organization can advance to the next level. Adam Smith envisioned an increased output ratio as each man perfected his act. AI, a change management method, can be deployed to determine the skillfulness required to avoid wasted time and effort. The AI methodology makes it possible for the entire organization to participate in deciding what needs to be changed, making it possible to distinguish between efficiency and effectiveness. It is essential to make this differentiation because successful corporations are both efficient and effective. By drawing from the Adam Smith pin-making enterprise, learning the best methods used to perfect the acts of different specialists can be documented and communicated to all epistemic community members. Feedback in all the practices is imperative to developing specialists. The Flipped Classroom offers the opportunity to co-construct curricula by sharing experiences and drawing lessons from the instructors' syllabus.

The Impact of Communications:

In Barrons (1993), organizational communications are critical to effective learning. In later research, Mupepi, Mupepi & Motwani (2015) posit that the centricity of organizational understanding is clear and concise communications. There should be an all-around communication among students and the instructor in the classroom as they share information and experiences. Barrons suggests that the environment nowadays includes information from the media, social networks, and prevailing cultural trends. For example, if students attended a concert, they would discuss the performance in class. This reflects that learning continued outside the classroom. The motivation from this experience influenced students' understanding. Mupepi et al. argue that skills acquisition and motivation are highly correlated.

Instructors Role:

Barrons (1993) considers the instructor central to all the (flipped) classroom activities. The role of the instructor is still paramount in the reduction of dropouts and in increasing the skillfulness of higher-order learning. Barrons argues that the most critical problem of all times is to address the problem-solving skills of each student. Chall (1983), the scripts of Sesame Street, which serves as an informal television classroom for preschool children up to age 5 (Vygotsky & Bandura), were changed to incorporate problem-solving skill development. In much later research Jacobs & Chall (2003) argue that in the developmental stages of reading development, reading is conceptualized not as a process that is the same from the beginning stages through mature, skilled reading but as one that changes as the reader becomes more able and proficient. Therefore, problem-solving is a developmental skill. In the flipped classroom, group exercises constitute problem-solving and the conceptualization of explicit practices with co-constructive epistemology.

One of the instructor's responsibilities in the flipped classroom is to dictate notes to students. This exercise enables the student to continue to develop their language arts skills. Lave (1988) suggests a correlation between writing and memory. In the flipped classroom, it is essential as students share information, they write down what they want to remember. Lave argues that apprentices in electrical engineering took notes as they saw the instructor provide demonstrations. In their final assessment, they were required to repeat those demonstrations. From their notes, they remembered the demonstration were learned in the classroom during the final exams.

The Way Forward

Whether using the regular or flipped instructional concept as an approach in the classroom strategy as applied to pedagogical models will continue to open alternatives to the learning community. Effective instructors are looking for ways to provide the highest levels of achievement to their students. In business alike, learning continues similarly when it comes to new tasks. Learning that has high-level relevance to the participants provides a high level of engagement. Using the flipped strategy allows one to apply real-world problem-solving experiences by adding a high significance level to developing explicit practices from tacit knowledge. In Figure 5, a practical model shows how learning will move forward the pedagogical process. However, this can only be done if one considers all the variables in this discourse and believes that the overall goal is to improve instruction. The SAP HANA and DbSchema program allows students to learn at the highest level of critical thinking and problem solving.

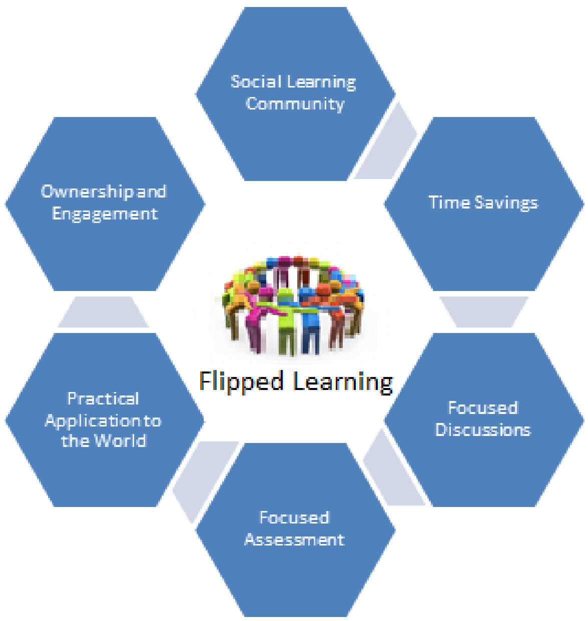


Figure 5: The Making of the Explicit Practices (Barrons, 2015)

Conclusion:

One can continue to observe achievements as organizational learning continues as a process of creating, retaining, and transferring knowledge within an organization as applied to the Hybrid Classroom to develop explicit practices. Thus, an organization improves as it gains experience. These experiences allow one to create knowledge and increase motivation for learning. The creation of the information systems classroom is an example by which learners can experience learning in a different approach from past standards, where traditional models are included in-class lectures and assignments after class sessions. One can look back at the pedagogical standards that have historically led pedagogy to today. A Hybrid Classroom for example, does not have a standard framework. However, there are many ways a classroom can be "Hybrid" by implementing traditional instructional practices. This article has offered only one of the many approaches to teaching SAP Hana Database and DbSchemas in a Hybrid Learning Model for MIS Competencies in higher education organizations.

References

Argyris, Chris (1991). Teaching Smart People How to Learn. *Harvard Business Review, May-June pp. 99-109.*

Bandura, A. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice-Hall.

Barrons, Kevin P. (1993) *Whittle Communications Channel One: The Effects of News and Advertising on Secondary Students.* Unpublished Ph.D. Dissertation. Detroit: Wayne State University.

Barrons, Kevin P. (2016) *Flipped Instructional Technology: Developing MIS Competencies Applying Enterprise Resource Planning Software*. IGI Global ISBN 9781522508243 Hard Cover Publication 2017.

Chall, J. S., & Jacobs, V. A. (2003). Poor children's fourth-grade slump. *American educator*, *27(1), 14-17.*

Chall, J.S. (1983) *Stages of Reading Development.* New York: McGraw-Hill.

Flavell, John H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive-Developmental Inquiry. *American Psychologist. Vol. 34, No. 10, 906-911.*

Foucault, Michel (2002). The Order of Things: An Archaeology of the Human Sciences. London, U.K.; Routledge classics. Psychology Press.

Fraser, Kelly (Dec. 7, 2006). Why doesn't The University of Michigan have snow days? *The Michigan Daily News.*

Jenkins, Chris (Feb. 02, 2009). The Australian Financial Review - ABIX via COMTEX -- SAP will supply software for a large revamp of Woolworths' merchandising systems.

Lave, Jean; Wenger, Etienne (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press. ISBN 0-521-42374-0.; first published in 1990 as Institute for Research on Learning report 90-0013

Lave, J. (1988). *The culture of acquisition and the practice of understanding* (pp. 259-286). Palo Alto, CA: Institute for Research on Learning.

Meyers, Chet; Jones, Thomas B. (1993) Promoting Active Learning, Strategies for the College Classroom. San Francisco: Josey-Bass.

Mupepi, Mambo (2010). Appreciating social structures: a strategy for advancing the efficiency and effectiveness in an organization. *International Journal of Education Economics and Development 2010 - Vol. 1, No.4 pp. 306 - 319.*

Mupepi, Mambo; Mupepi, Sylvia; Motwani, Jaideep (2015). Social Construct: Methods, Tools, and Applications in Re-Engineering Organizations. *The Proceedings of the MBAA International Conference March 25-27. Chicago.*

Roehl, A., Reddy, S. L., & Shannon, G. J. (2013). The flipped classroom: An opportunity to engage millennial students through active learning strategies. *Journal of Family & Consumer Sciences*, *105*(2), 44-49.

SAP.COM (2015 a). SAP grows up, but never grows old. Accessed 05/29/2015.https://www.sap.com/careers/about-us/history.html

SAP.COM (2015 b). SAP at a glance: Company Information. Accessed 06.02.2015. http://www.sap.com/corporate-en/about/our-company/index.html

Scida, E. E., & Saury, R. E. (2006). Hybrid Courses and Their Impact on Student and Classroom Performance: A Case Study at the University of Virginia. *CALICO Journal*, *23*(3), 517–531. http://www.jstor.org/stable/24156352

Senge, Peter, M., Art Kleiner; Charlotte Roberts; Richard Ross; Bryan Smith. (1994). *The Fifth Edition Fieldbook.* New York: Doubleday Books.

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children*, *23*(3), 34-41.

65. Efficacy of Collaborative Exams

Garret Marshall

Purdue Fort Wayne, Fort Wayne, USA

Abstract

Traditional college exams are often underutilized as a learning tool. The design of this research was based on the cognitive learning theory of **reciprocal determinism. In breaking the exam paradigm, personal and environmental factors that would normally inhibit engagement, theoretically are replaced by a new reinforcing stimulus that could lead to increased understanding. This study is intended to apply these ideas through a** structured approach to collaborative exams. These exams were administered in introductory statistics courses over the course of seven semesters. The scores of collaborative test-takers have been found to be consistently higher than those of a control group in the same semester. Protocols for group assignment and assembly are discussed, along with measures to ensure individual accountability.

Research Question

Do collaborative exams lead students to a better understanding and retention of course material compared to traditional exams?

Preliminary Lit Review

Jang, Hyewon et al. "Collaborative Exams: Cheating? Or Learning?" American journal of physics 85.3 (2017): 223–227. Web.

Scott T Miller, and C Renée James. "What Is The Impact Of Collaborative Exams On Learning And Attitudes In Introductory Astronomy Classes?" Journal of astronomy and Earth sciences education 6.1 (2019): 17–. Print.

Shindler, John V. "'Greater Than the Sum of the Parts?' Examining the Soundness of Collaborative Exams in Teacher Education Courses." Innovative higher education 28.4 (2004): 273–283. Web.

Cooke, James E, Laura Weir, and Bridgette Clarkston. "Retention Following Two-Stage Collaborative Exams Depends on Timing and Student Performance." CBE life sciences education 18.2 (2019): ar12–ar12. Web.

Arteaga, Martin. "Delayed Feedback Might, but Not Collaborative Exams, Improve Long-Term Student Retention of Course Content." eScholarship, University of California, 2021. Print.

Eaton, Timothy T. "Engaging Students and Evaluating Learning Progress Using Collaborative Exams in Introductory Courses." *Journal of geoscience education* 57.2 (2009): 113–120. Web.

Ives, Joss. "Measuring the Learning from Two-Stage Collaborative Group Exams." (2014): n. pag. Print.

Stearns, Susan A. "Collaborative Exams as Learning Tools." *College teaching* 44.3 (1996): 111–112. Web.

Theoretical Framework

- Active Learning
- Cognitive Learning Theory - Reciprocal determinism
- Free-rider problem / Moral hazard
- Game Theory - Signaling Game
- Item Response Theory

Overview of Procedure

In seven semesters (PFW and UT), my teaching assignments have included multiple sections of the same class. One class was selected as experimental, and the other(s) served as control groups. All classes were designed with two or three midterm exams and a final exam. The grades from the first midterm were used to assemble **balanced groups** in which the grade percentile rankings were approximately equal. Care was taken to ensure break up the students' preferred seating with seating assignments by group. As much as possible, groups were limited to four members. Once assigned, the students were made aware that the second exam would be collaborative and were "advised" to prepare in their assigned groups and that measures would be taken to ensure **individual accountability** – though the specifics were not discussed.

On the day of the second midterm exam, a **mandatory seating chart** was shown on the class projector. Five versions of the second exam were prepared so that each group member would have a **different version**. The problems were the same on each version, but the calculations varied – often leading to different conclusions. For example, the conclusion of one version of a hypothesis test might reject the null hypothesis, where another version would not. Student could collaborate on strategy, but were has to complete their own exam.

The final exam from each course included a number of questions related to the material covered on the second midterm exam. Typically, there were five or six weeks between the second midterm and the final exam. The final exam was not collaborative. For the purposes of the study, the results of each question were recorded as either correct or incorrect. The performance on the final exam was used to test if students in the experimental classes performed at a higher level compared to the control class.

Test of Efficacy

Chi-squared test for independence using collaborative or control as the row variable, and correct or incorrect as the column variable. Tests for dependence was performed in three ways:

- Overall performance – The counts of correct (and incorrect) questions on all material found a dependence on the experimental group. Conclusion, there was a statistically significant difference in scores.
- Performance on material from Exam 2 – The counts of correct questions on material directly related to that of the second midterm found a strong dependence on the experimental group.
- Performance on material ***not*** related to Exam 2 – The counts of correct questions found no association between performance on this material and the groups. Inference: the students performed better when they collaborated.

To demonstrate the direction of the difference found in the chi-squared tests, **two-sample confidence intervals for the difference in proportions** were calculated for each of the above. Using the proportion of correct answers for *overall performance* and that for *Exam 2 material*, the intervals of the differences found that the experimental groups were significantly higher. The interval for differences in *performance on material not related to Exam 2* found no significant difference.

Accountability measures

- Different versions of the exam
- Dominant strategy seeding (game theoretic)
- Maximize distance from original seating (max composite distance scores for group)

Group selection protocol

- Objective: to break up "friends" and even the social dynamics
- Ideal groups 3-4 students
- Use cumulative points after the first exam to rank students.
- Divide class into number of groups according to ranks
- Assumption: leaders will tend to be the students in the highest rank.
- Populate each group with random selection from each ranking group
- Use cumulative scores from all member to generate performance weight
- Adjust until groups are as even as possible.
- Break apart students that normally sit together.

Seating protocol

- Create a seating chart before the exam (where they normally sit)
- Assumption: students that normally sit together are often friends
- Transfer the pre-exam chart to a square grid
- Once groups are assigned, place the groups onto the grid such that students are as far from their "normal" seats as possible
- With each group position has an aggregate displacement score
- Maximize the aggregate score for the groups

Conclusions

The **odds ratio** of correct answers vs experimental groups showed that students that completed a collaborative exam are more than 4 times more likely to answer these questions correctly compare to the control groups.

Project Management

38.

Prefabrication for Successful Mechanical, Electrical and Plumbing Scope of Work in Construction Projects

Sherif Attallah

Fishers, Fishers, USA

Abstract

Shortage of qualified skilled manpower in the construction industry is an ongoing issue for many years now and this is expected to get worse with time. The mechanical, electrical and plumbing (MEP) scope of work is no exception, to the contrary, those areas are affected more as they require more skills and, in some cases, licensed technicians. To address this problem, mechanical and electrical contractors are pursuing prefabrication as a strategy to reduce the manhours needed to complete jobs on site. Prefabrication is not only an efficient strategy for saving manpower on construction size, but it helps contractors with achieving higher quality and safety objectives. Working in the workshops is by far safer than site conditions, some of which are very rough and the quality of products are higher and more consistent.
The objective of this presentation is to explain how mechanical, electrical and plumbing (MEP) contractors utilizes prefabrication to improve their performance and achieve project objectives in terms of time schedules, quality and safety

Categories

Project Management

42.

Strategies to Effectively Manage an Internship Program

Jennifer Warrner, Natalie Sanchez, Carlie Stalnecker

Ball State University, Muncie, USA

Abstract

This presentation highlights strategies for how to effectively manage an internship program for an academic major. Information about how to assist students with conducting an internship search, how to recruit and work cooperatively with employers and industry partners, how to develop an academic internship course, and how to supervise students while completing their work experiences will be discussed. Information from the faculty, employer, and student perspectives will be shared. The information in this session can be applied to academic programs at other institutions.

Categories

Project Management

45.

The Importance of Community: Retaining Female Students in a Male-Dominated Industry

Sarah Strong, Jennifer Warrner

Ball State University, Muncie, USA

Abstract

Although it's been found that there is an increase in women working in the construction industry, they are still considered distinct minorities in the field. It's important to have gender diversity represented in construction and other male-dominated sectors to benefit from having different perspectives and better problem-solving. Recruiting and retaining students from underrepresented populations is challenging. Outreach activities are one strategy to help build a sense of community and to maintain one underrepresented population, female students. Activities from one outreach program, best practices for scheduling, and funding ideas will be shared that could be incorporated into other academic programs and organizations.

Categories

Project Management

46.

Teaching the Art and Science of Construction Scheduling: Activity Development and Logic

James Jones, Natalie Sanchez

Ball State University, Muncie, USA

Abstract

Scheduling, along with cost, quality, and safety is typically considered to be one of the critical attributes for success in construction project management. While some aspects of the scheduling process are straightforward and can be approached objectively, some students seem to struggle with the subjective determination of how to break down a construction project into distinct work activities, and how to properly sequence them.

The Precedence Diagramming Method (PDM) is now the most commonly taught approach for graphically showing the logical relationships between activities and calculating key dates, such as the earliest date an activity may be started, as well as its total and free float. However, this method is reliant on the proper determination of its component activities and their logical relationship with other activities. Poor activity development and sequencing will result in poor management of the project; as the saying goes, "garbage in, garbage out."

This presentation provides attendees the presenters' perspective on teaching the "art" of activity determination to students in construction curricula. These include both contractual aspects as well as industry preferences, with the intent of preparing students to better meet these challenges when they begin their careers in industry.

Categories

Project Management

48.

Is Artificial Intelligence an Opportunity or a Threat for Construction Project Management?

Tamer Breakah

Ball State University, Muncie, USA

Abstract

The recent launch of ChatGPT developed by OpenAI raised several questions in many industries. The major question is how would this artificial intelligence breakthrough, which will be followed by other artificial intelligence competitors, would affect specific industries? The main goal of this research is to investigate the effect of this breakthrough on construction project management. Each development in technology can be considered as an opportunity or a threat. In this research, an analysis of the possibilities of using artificial intelligence as a tool in construction project management is presented. This analysis includes a critical review of these possibilities and whether they are opportunities or threats. As a case study, some trial runs were done using ChatGPT, to study its capabilities and limitations as a tool to be used by construction project managers. The runs showed that the software has a big potential, but still has lots of limitations. It can be used as a support tool to gather information, but not for decision making. Finally, this research includes guidelines on how to best implement newly developed artificial intelligence tools in construction project management.

Categories

Project Management

49.

The impact of modality on pedagogy and students in construction management

Tarek Mahfouz

Ball State University, Muncie, USA

Abstract

At the time institutions of higher education are facing multitude of challenges including, but not limited to, lower enrollment and admissions numbers, changed demographics, increased pressure to substantiate the value proposition, new modality and delivery methods, non-credit bearing knowledge opportunities, badging and stackable credentials, and higher competition, among many others, the pandemic has forced educators to re-evaluate the pedagogical premise of the course materials and assessment instruments. The current research endeavor presents a case study at which the impact of different delivery methods (Face-to-Face, online, and hybrid) on students' performance, engagement, and attitude in a Construction Finance and Law class taught in a Construction Management program at a four-year mid-west Institution of higher education are being evaluated and assessed. To that end, construction finance and law classes (multiple sections) taught over a period of three years are evaluated. While the faculty member and the course content are maintained constant, the student population is different in each year. The study compares accreditation assessment data of four (4) student learning outcomes (SLO) associated with the class, students' performance in Exams and Assignments, and level of engagement and participation to assess the impact of modality variation. The outcomes indicate that providing the flexibility for students related to crucial milestones, like exams, enhances students' performance overall.

Categories

Project Management

59.

C, D & E in Ethics for Construction Management – Conformity, Diversity, & Ethics

Gary Birk

Ball State University, Muncie, USA

Abstract

Beyond high school, many young people move directly into the workforce, or advance their education with training geared directly toward their specific goals and skillset. Those students who choose to move toward a more advanced degree are often able to encounter more diversity in the higher education system. Construction management is no exception. For example, stereotypically, construction has been a male-dominated field. Fortunately, diversity is becoming more commonplace currently in construction. Still, even students entering into construction management, or more broadly a diverse college environment in general, may show up with a fairly narrowly defined set of tools in their toolbox. The values that students may bring with them when entering college may be confining when it comes to where the business community is more moving to more broadly. Ethical decisions made in the professional environment can have impacts on all parties and may be linked to an individual's upbringing and their exposure to diversity. A course at the college level dealing with ethics and construction management is a start.

Categories

Project Management

Supply Chain Management & Logistics

15.

The Effects of Supply Chain Disruptions Post COVID-19 era

Ceyhun Ozgur

Valparaiso University, Valparaiso, Indiana, USA

Abstract

This paper analyzers the effect of supply chain disruptions on a firm. For example, how would an increase or decrease in inventory would disrupt the firm's operations. How would change in climate effect the firm's operations? How would supply chain would be affected by a change int the quality of the product manufactured by the company. How would the company be impacted by the environmental changes and supply chain disruptions? How would the change in forecast effect the supply chain? How would a change in inventories effect the supply chain of the company? How would the change in the increase or decrease in the ability to purchase raw materials, work- in--process or finished goods inventory effect the supply chain of the product.

Categories

Supply Chain Management and Logistics

16.

DATA MODEL FOR MULTI-BIOMASS SUPPLY CHAIN

Nasim Zandi Atashbar [1], Christian Prins [2], Nacima Labadie [2], Kuntal Bhattacharyya [1]

[1]Indiana State University, Terre Haute, USA. [2]University of Technology of Troyes, Troyes, France

ABSTRACT

Biomass plays an important role as one of the main renewable energy sources. Given the high logistical cost associated with biomass adoption, having an efficient supply chain that provides bio-refineries with adequate quantities of biomass at a reasonable price and in a timely manner is critical. To this end, this paper proposes a flexible data model for a multi-biomass supply chain. The role of this data model is to list, analyze and structure a large amount of data for a multi-biomass supply chain in a logical way. The result is a set of tables that can be recorded in a database management system. These tables contain input data that can then be loaded into a mathematical programming environment to support biomass supply chain modeling and optimization.

Keywords: Conceptual data model, database design, logistics, and multi-biomass supply chain.

INTRODUCTION

Unavoidable effects of climate change caused by greenhouse gas emissions, together with a growing global concern for energy demand, motivated a recent stream of research for alternatives to fossil fuels (Bauer et al. 2016; Liu et al. 2007). Biofuel obtained from biomass, as a renewable and clean energy source, is one of the few potential replacements for fossil fuels and can play an important role in transitioning from traditional sources of energy (Zandi Atashbar et al. 2018a). Although biomass is cheap relative to other energy sources, the delivery cost at refinery gates varies significantly due to the complicated logistics of biomass. Unlike many other industrial logistics with a constant flow of resources, the biomass supply chain is highly seasonal and dynamic. The raw materials (oilseed and lignocellulosic crops) are produced slowly, seasonally, and with a limited yield over a vast territory. A refinery must use successive crops during the year, e.g., miscanthus in spring, rape in July, cereal straws in August, camelina in October, and short-rotation trees like willows in winter.

From a field to fuel, the biomass supply chain includes various activities such as cultivation, harvesting, handling, storage, transportation, and biofuel conversion. In practice, a biomass supply chain is a complex network of enterprises. The role of biomass supply chain management is to make all the complex decisions involved in designing and operating such chains. To date, most models focus on one or two steps at a time, perhaps because of the system's complicity. However, if we are to design an agile biomass supply chain, it is essential to model and optimize a multi-biomass supply chain as an integrated entity.

A feasible way to design a straightforward descriptive tool can be a layered graph; nodes are associated with entities (e.g., production activities, end-customers, storage facilities, preprocessing plants, transshipment stations, etc.), and arcs are associated with product flows (Gumus & Guneri 2009). Figure 1 shows a sample layered graphical framework for multi-biomass biomass supply chains.

This research stream combines operations research techniques with biomass supply chain data models (Zandi Atashbar et al. 2018a; Sharma et al. 2013; Sun et al. 2020; Nunes et al. 2020), wherein these models reference a database (Frombo et al. 2009; Freppaz et al. 2004). The model is required to be independent as possible from its data to be able to handle a large amount of data. Any logistic network can be defined by the user and stored in the database instead of imposing a frozen structure in the mathematical model. The goal is to come to a "data-driven" mathematical model which can be easily modified and extended. Very few studies provide a comprehensive data model to address such issues (De Meyer et al. 2016). Here, we aim to address this gap by developing a simple and unified framework of the biomass supply chain, which deals with complexity and adaptiveness specific to this industry.

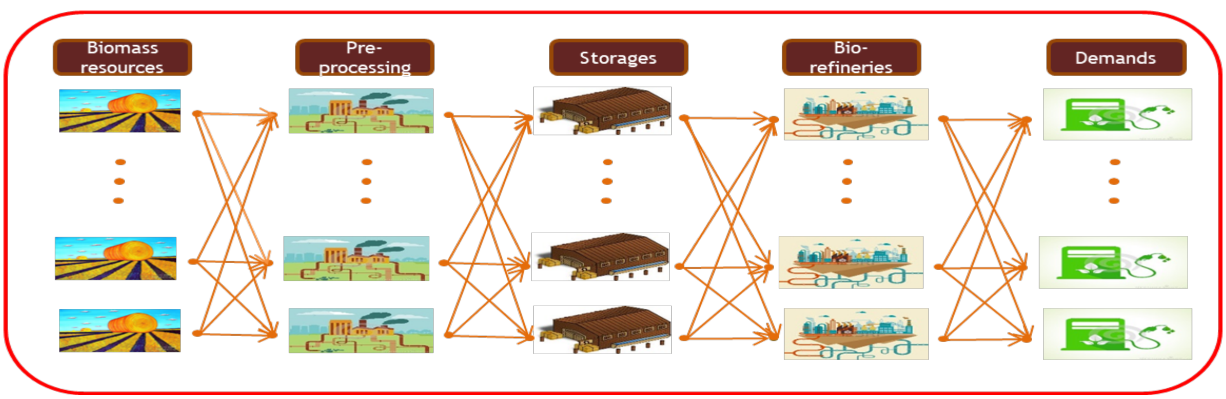


Figure 1. A Biomass Supply Chain Framework

METHODOLOGY

Data Description, Assumptions, and Set-up

The activities in the biomass supply chain are often regrouped into three major segments shown in Figure 2. The **upstream segment** takes place before the refinery. It includes biomass production, harvest, collection, preprocessing, and centralized storage. The **midstream segment** corresponds to the conversion processes at the refinery. Finally, the **downstream segment** covers the output storage of the refinery and the distribution to customers.

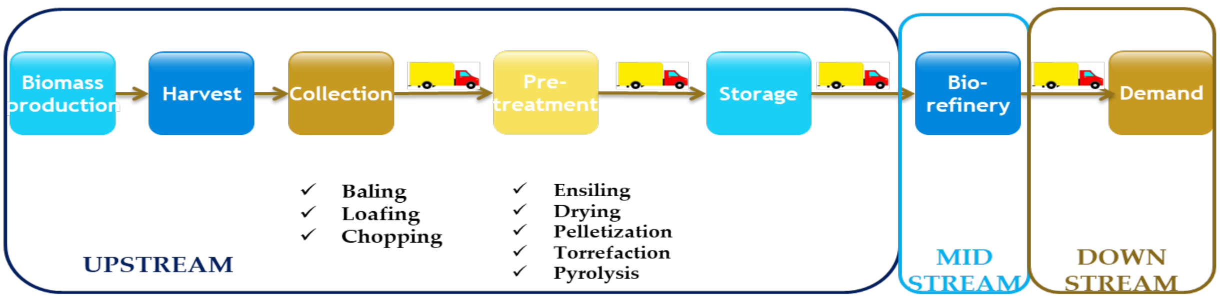


Figure 2. Major segments and activities in biomass supply chain.

Very few authors like (Eksioglu 2009) have tried to model the three segments together. There are multiple reasons for this void in the literature. First, very different actors and contract types are involved in the upstream and downstream. Secondly, the midstream and downstream are not really unique to the biomass supply chain since they are similar to the production and distribution in the petroleum industry. This explains why most research articles focus on the upstream segment which raises the most interesting problems. Biomass supply chain planners face complex and challenging problems at different decision levels such as biorefineries location, transportation mode selection, preprocessing technology selection, and vehicle fleet size.

This paper focuses on the biomass supply chain over a multi-period planning horizon, with different biomass types called "products", centralized storage facilities, and several biorefineries of which some are actual locations and others potential locations designated as nodes on the network. The supply chain is analyzed begins with harvested products, which are ready to ship and then to refinery storage. The aim is to prepare a database for this multi-biomass supply chain, at the tactical and strategic decision levels. The planning horizon is divided into discrete time slots ("periods"), currently 52 periods of 7 days. The number and duration of periods are designed in a way that future modifications to the model can be easy and streamlined. The strategic decisions concern involves the identification and location of refineries to minimize logistical costs to and from the refineries. Tactical decisions deal with the amounts collected in the farms, stored and transported, in each period. The area studied corresponds to Picardie and Champagne-Ardenne in France. It is partitioned into discrete territorial units called "zones" (currently 279 cantons). The cantons are those of the 2010 agricultural census, used to prepare biomass production data. Refineries are either already placed or need to be located, and there is at most one refinery per zone. Each refinery defines its needs per product and per period over the planning horizon. Biomass production data is computed by one partner of this project (AgroTransfert et Territoires – AGT-RT) and include cultivation and harvesting. The density and humidity of a product are the same, whatever the zone is. Moreover, the humidity and density of stable products do not change along the chain, but storage loss is handled. Products are transported by road, but other transport modes can be added. The transportation network is defined by a graph. The database design needs to be flexible despite current choices (e.g., new products, preprocessing sites, new locations for refineries, etc.).

Data Model Design

In the biomass supply chain, decision support systems are generally described as three modules 1) a database module 2) a query module 3) a decision module shown in Figure 3. The database module stores the input data of the optimization model. These input data include non-georeferenced data such as the different biomass operations and georeferenced data such as the characteristics of different facilities. To organize, visualize and process input data, the database module is linked to a GIS-based query. The decision module includes the tool to optimize the biomass supply chain (De Meyer et al. 2016). This paper focuses on the database module, and it is flexible enough to easily add, delete, or change types, attributes, and attribute types.

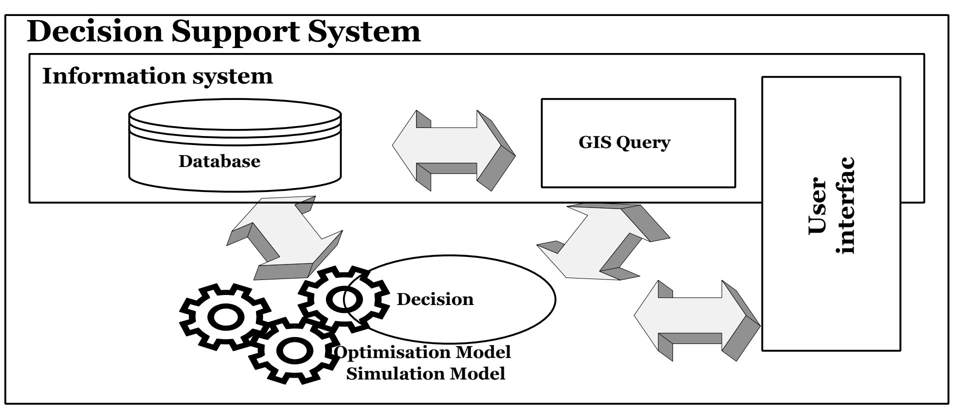


Figure 3. General architecture decision support system for biomass supply chains (De Meyer et al. 2013)

THEORETICAL UNDERPINNING & CONCEPTUAL MODEL

Conceptual Data Model (CDM)

A good way to logically structure a big set of data is to design a database, a set of linked files offering better protection against common errors like duplicate or missing records. The structuration of a database, independently of data values, is called a **data model**. Most databases today comply with the **relational data model**, in which data must be decomposed into a set of linked tables. To reduce complexity when designing a relational data model, it is recommended as a first step to prepare a graphical representation called **Conceptual Data Model (CDM).** Our conceptual data model is based on only two concepts - entities and relationships. Given the interconnectivity of relationships between constituents of a multi-biomass supply chain (amount of produce collected in farms, storage, shipment, and location of refineries to optimize logistics), our conceptual data model is best grounded in the entity-relationship model as prescribed by (Chen 1976). The significance of the entity-relationship model lies in its capability to merge best practices from a family of predecessor models, namely the network model (Bachman 1975), the relational model (Codd 2002), and the entity-set model (Senko et al. 1973). The entity-relationship model, by virtue of its high degree of data independence, adopts a more natural view of the real-world (Chen 1976). An **entity** is a set of persons or things (real or abstract) of the same nature, called members of the entity. The **identifier (ID) or key** of an entity is one attribute to identify unambiguously each member. Entities are linked by **relationships or associations**. These processes and flows are translated into entities and relationships to design an entityrelationship data model (Chen 1976; 1988; Storey 1991), In this way, the supply chain can be modeled by linking the components of the chain (the entities) with each other by user-defined relationships and “if-thenelse” rules that determine the dependencies and requirements between the components (Chen 1976; Chen 1988; Storey 1991; Kelly et al. 2013). In addition, these entities are characterized by properties, also called **attributes** (Chen 1976; Chen 1988; Storey 1991). Table 1 lists the relationships and entities of the CDM for this study.

Name	Type	Contents
Parameters	Entity	General parameters: number of periods, period duration, etc.
Zones	Entity	Territorial units considered, currently cantons
Distances	Relationship	Inter-zone distances precomputed using Map Point
Crops	Entity	Crops harvested to get products
Products	Entity	Products: crop of origin, density, humidity, etc.
Local	Relationship	Local product variations: yields, harvesting windows, etc.
Node types	Entity	Node types, currently BP, FS, CS and RL
Nodes	Entity	Nodes (zone + node type + product) with storage capacity, etc.
Arc types	Entity	Allowed arc types, at least (BP,FS), (BP,CS), (FS,CS) and (CS,RL)

Vehicles	Entity	Vehicle types
Refineries	Entity	Refinery types which can be created on RS nodes
Demands	Relationship	Demands per refinery type, product, and period
Flows (X)	Relationship	Flow variables per arc and per period (results)
Stocks (S)	Relationship	Stock per node and per period (results)
Locations (Y)	Relationship	Binary location variable per refinery type and per zone (results)

Table 1. Entities and relationships of the CDM giving a table in EXCEL

Figure 4 illustrates the CDM for the biomass supply chain. Our initial idea was to define the vertices of the supply chain graph as territorial units like cantons. The problem is that they contain various sites with different properties, at least fields where crops are harvested, on-farm storage facilities, centralized storage facilities, and refinery storage. Other kinds of nodes such as intermodal transshipment hubs, preprocessing facilities, importations via harbors, and even stocks of by-products from other refineries can be added in the future.

Moreover, we can notice that the activities on each site are separated per product. For instance, as we need to know the amount of each product in centralized storage, this kind of site can be viewed as one stock per product. In fact, all sites can be viewed as sets of stocks. This is obvious for on-farm and centralized storage units. Concerning cultivated fields, the amount harvestable for each product is also a kind of stock. Finally, as the supply chain does not cover the conversion processes of refineries, a refinery site can also be modeled as a set of input stocks, one for each required product.

The main design choices can be derived from these remarks:

§ We call zones the territorial units considered and nodes the different sites they contain. They have gathered respectively in two entities "Zones" and "Nodes". The zones selected are currently cantons because it is difficult to compute potential productions at lower levels.

§ Nodes are partitioned according to their role, called node type. Currently, four node types are defined in an entity "Node types": BP (biomass production node), FS (farm storage), CS (centralized storage), and RS (refinery storage). Other node types could be added in the future.

§ Each node is viewed as a stock, defined by a zone, a node type, and a product. As biomass data from AGTRT are consolidated by zone (canton), we do the same for all nodes. So, for each zone and each product, there can be at most one node of each type, e.g., the CS node for rape seeds aggregates all centralized stocks of rape seeds in a canton. Note that this also implies at most one refinery per zone. All nodes in a zone are assumed to be in a city chosen as the geographic center for the canton. This location is required to compute distances between any two cantons.

§ Some product data like the amount available are associated with the nodes of the zones, but the yield, harvest window, and cost depend on the region, while density and humidity are constant, whatever the zone is. So, these three types of product data are respectively stored in three different tables: "Nodes", "Local", and "Products".

§ There are many potential refinery locations but only a few sizes of refineries. We think the demands of a refinery of a certain size will be determined by its conversion processes but not by its exact location. So, to enter fewer data, we separate refinery locations and refinery types (sizes). An attribute in entity "Zones" indicates if a zone has already one refinery, if it can receive one, or if creation is forbidden. The few refinery types and the demands for each type are respectively defined in one entity "Refineries" and one relationship "Demands". For instance, if we have two small and one large refinery to build among 200 possible zones, we have two records only in "Refineries", instead of copying the same data into all possible zones.

§ We use a classical trick for the general parameters which cannot be considered as attributes of a particular entity. They are stored in a table "Parameters" which contains a single record, with one parameter per column: number of periods, number of days in one period, etc.

§ All product flows, storage and vehicle capacities, and refinery needs are expressed in metric tons, i.e., the tons used in Europe and not the US tons. "tRM" means "ton of raw material".

Other design choices aim at limiting model size. The largest number of variables comes from product flows, which are indexed by one arc (pair of nodes) and one period. Each arc can be defined by a zone of origin, its node type, a zone of destination, its node type, and a product. It is quite likely that the model will be too big if we allow flows between any two nodes, for all products and all periods. So, the number of arcs and flow variables must be as small as possible. This can be done at 5 levels:

§ **First level**. To avoid flows $F(p, i, t, j, t')$ like in Shastri et al. 2013 (amount of product p shipped from zone i in period t and arriving in zone j in period t'), we assume that any product shipped in one period arrives in the same period. This is possible due to the territory considered (two regions) and the rather long period duration (one week). As the product is known for each node and as i and j store the same product, we need no product index and use flow variables $F(i, j, t)$.

§ **Second level.** The allowed arc types can be restricted using one entity "Arc types". For instance, (BP,FS), (BP,CS), (FS,CS) and (CS,RS) are always allowed. If the user permits direct deliveries from farms to refineries, arc types (BP,RS) and (FS,RS) may be added. One advantage is that each arc can now be defined by one pair of zones, one arc type, and one product. Another advantage is that a vehicle can be specified for each arc type. The set of arcs of the logistic network can be generated automatically, instead of being typed, but only for the arc types allowed in "Arc types".

§ **Third level**. The arcs with a BP or FS origin can be limited to the products harvested in the zone.

§ **Fourth level**. For each arc type, we can specify a default distance when nodes *i* and *j* are in the same zone, and minimum and maximum distances otherwise. For instance, (BP,FS) arcs must not leave a zone, so we can specify an interval [0,0]. As farmers bring their products to centralized storage facilities at a 10 km maximum, we can specify an interval [0,10] for (BP,CS) and (FS,CS) arc types linking two distinct zones. Long-range arcs should be exceptional, e.g., between one external port where biomass is imported (Le Havre, Rotterdam…) and cantons where a refinery may be located.

§ **Fifth level**. We restrict the periods of flow variables, e.g., (BP,CS) and (BP,FS) flows for a product are limited to the harvest period. Due to insurance problems, on-farm storage is limited to 1 to 2 months after the end of the harvest window, so (FS, CS) flows can be defined only in these time intervals. More generally, a flow from node *i* to node *j* may exist only when both nodes are open.

Finally, an effort has been made to have logical names for all tables and attributes. A letter used in attribute prefixes or codes like arc types has always the same meaning. Here is the logic:

A = arc D = distance N = node R = refinery type
AT = arc type Dur = duration NT = node type S = storage/stock
B = biomass F = farm Num = number T = type
C = center/centralized L = local product data P = product/production
V = vehicle

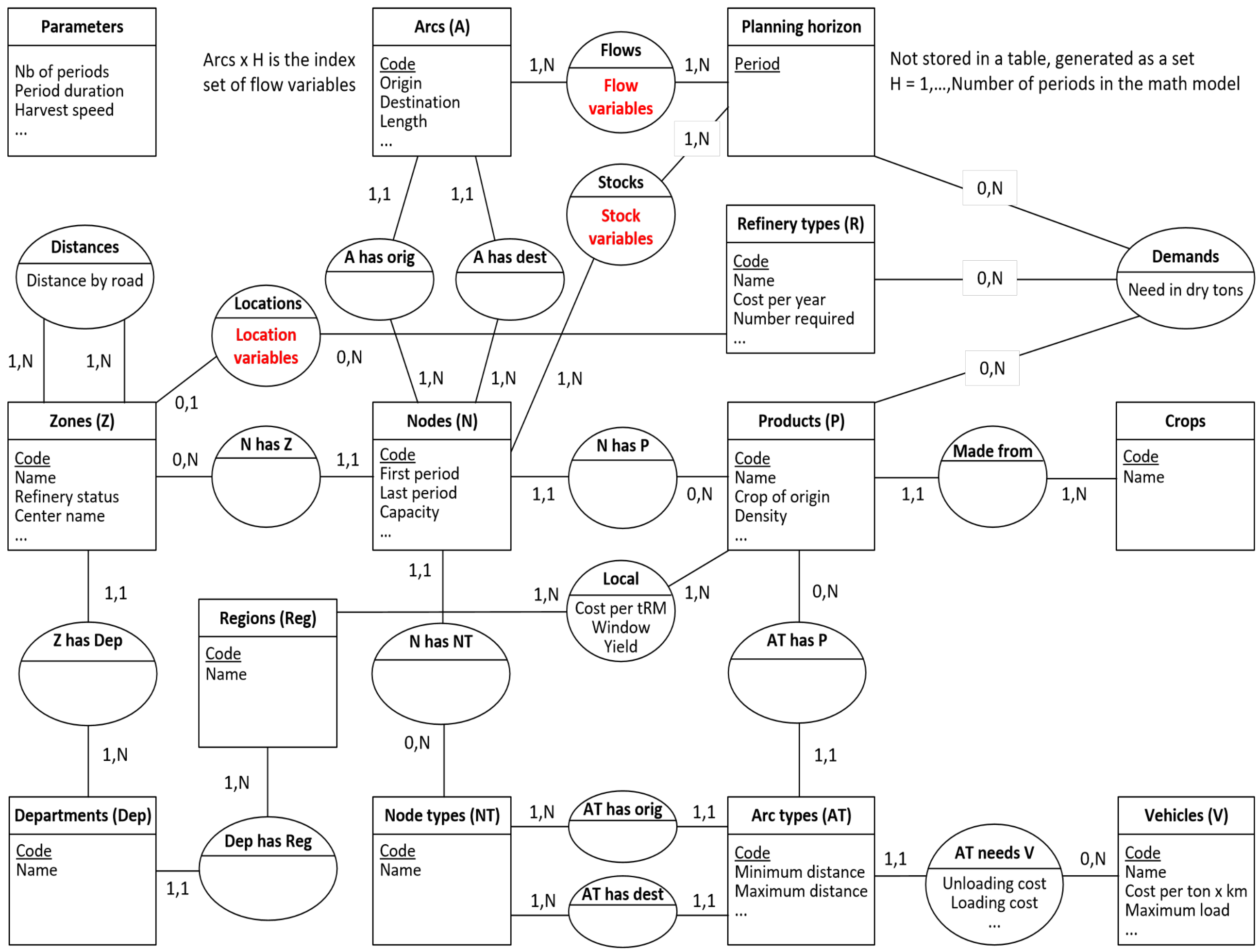


Figure 4. CDM for biomass supply chains

Conversion of the CDM into a set of tables

In a second step, this CDM can be converted into a set of tables which are then implemented in EXCEL or a true database management system (DBMS). The CDM must be translated into a set of rectangular tables which specify the set of members of each entity and the values of their attributes. Each table can be implemented as one worksheet in EXCEL. In the first step, each entity of the CDM must be converted into one table. The name of the EXCEL worksheet is often the same as the entity name. Its rows correspond to the members of the entity and its columns to the attributes. At this stage, rows and columns are also called records and fields. The first row (column headers) must give attribute names. Table 1 is the list of entities and relationships that we decided to implement as tables (worksheets) in our EXCEL database.

APPLICATION OF DATA MODEL

Before coming to a complete database, a lot of upstream work is required to obtain the set of zones considered (cantons), inter-canton distances by road, centralized storage data, and biomass production data. We have compiled an auxiliary workbook titled “all zones and distances” that contains administrative data for the 279 cantons of Picardie and Champagne-Ardenne: canton code and name, a city chosen as a geographic center to compute distances, number of farms, area cultivated, the storage capacity of silos, area of platforms and intercanton distances. The workflow to prepare this workbook is illustrated in Figure 5. The process is rather complex because it has been decomposed in successive steps to be more reliable. Note that rectangles are EXCEL workbooks while ellipses represent procedures. A workbook may contain several worksheets (inner rectangles) and macros written in VBA (Visual Basic for Applications, inner ellipses). Arrows indicate data flows.

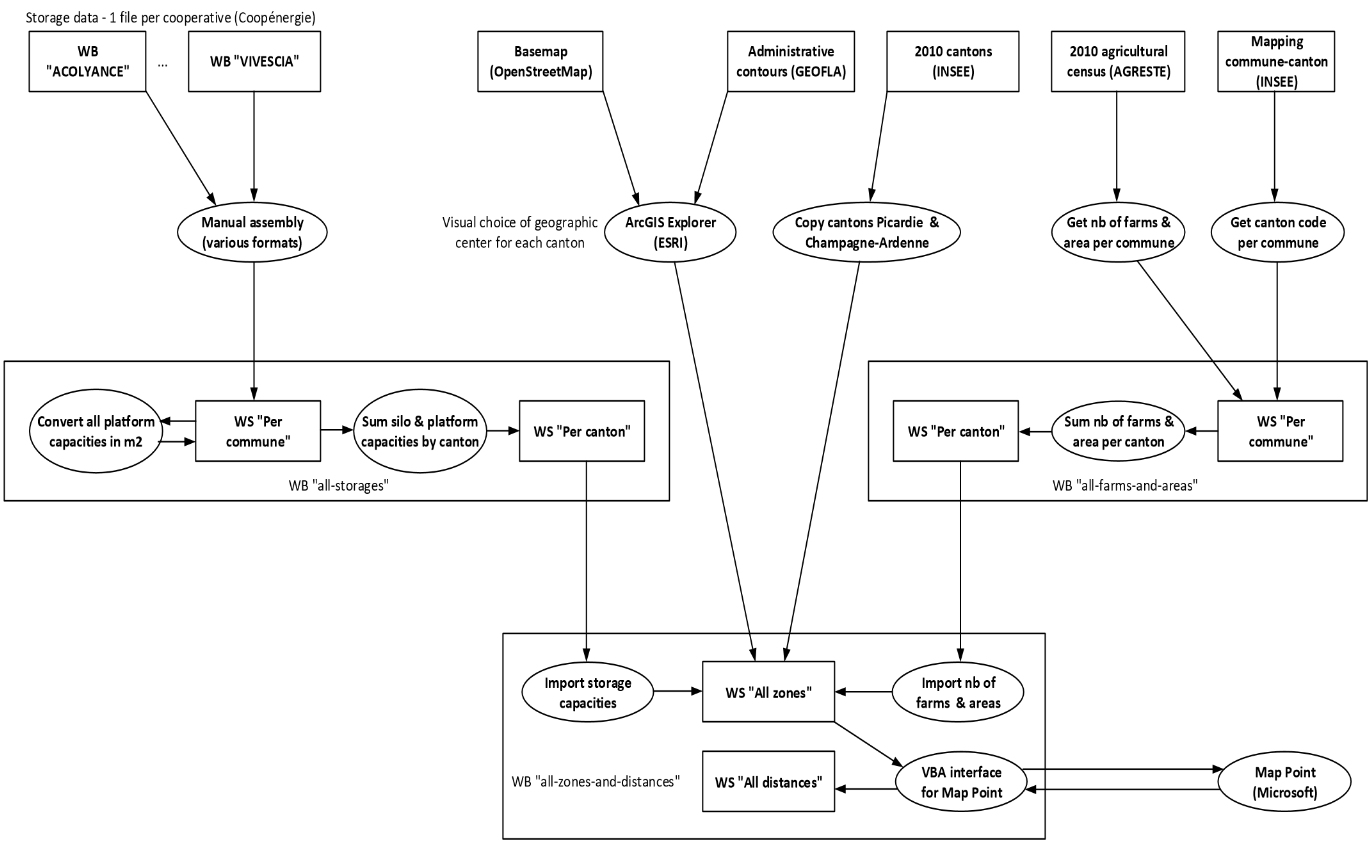


Figure 5. Workflow to prepare workbook "all-zones-and-distances"

The other auxiliary workbook “all-production-data” was built by our industry partner AGT-RT. It gathers biomass production data per canton and per product, e.g., current and potential amount available, production cost, loading cost, GHG emission, and fuel/energy consumption. It contains one worksheet per department. Table “all-production-data” contains one row with the following fields for each pair (canton, product harvested in this canton):

- Code and name of the canton
- Product, defined by the crop and the part concerned (seeds, straw, chaff, whole plant)
- The current production, null or very small for camelina, Ethiopian mustard, miscanthus, and willow
- Potential production for camelina, Ethiopian mustard, miscanthus, and willow
- Harvest window (first and last week)
- Percentage of dry matter
- Conditioning at the end of the technical itinerary (bulk or bales) • Product density

Indicators for biomass production, loading on the tractor, loading on truck, tractor unloading, and truck unloading: cost per tRM, diesel consumption, energy consumption, and GHG emission. The links between a scenario and the two auxiliary workbooks are shown in Figure 6.

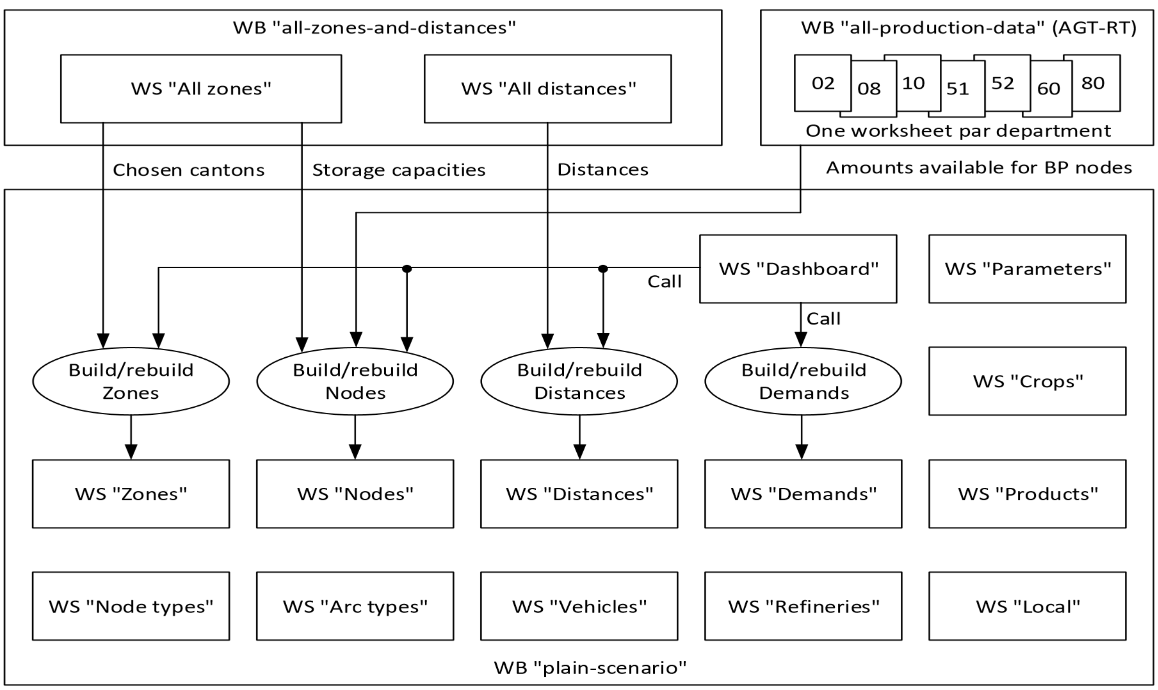


Figure 6. Scenario file and links with auxiliary workbooks

Dashboard for scenario generation

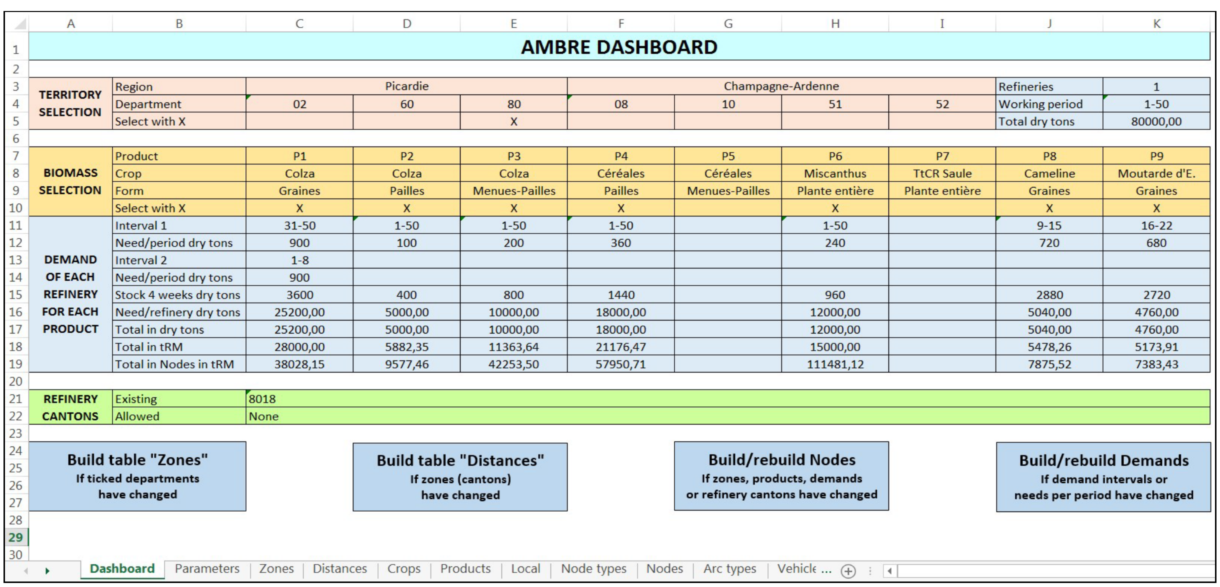
We have now a set of structured tables allowing all data for an optimization scenario to be stored in an EXCEL workbook. The data model is flexible enough to add new node types, new products, etc., and it becomes easier to write the equations of the math model.

To generate easily various test scenarios, we designed a workbook. It contains one worksheet for each table of the data model. All are prefilled except the largest ones which contain column names only: Zones, Distances, Nodes, and Demands. These large tables are going to be filled automatically from a set of directives entered by the user in a new worksheet “Dashboard” in Figure 7, inserted before the real tables of the database.

This control panel allows specifying:

- The departments of Picardie and Champagne-Ardenne are to consider in the scenario (pink cells).
- The number of refineries, existing or to create (cell K3) and their opening periods (cell K4).
- The products asked for by refineries, are among the ones for which we have production data (orange cells).
- For each product, up to two demand intervals (first period, last period, and constant demand per period in dry tons) can be specified in blue rows 11-14. The other blue cells are computed in EXCEL. For each product, we have the storage capacity computed for 4 weeks of demand (row 15), the total need in dry tons for one and all refineries (rows 16-17), and the total need of all refineries in tRM (row 18).
  The amount available in table "Nodes" in tRM is also displayed but only if this worksheet has been already built. Finally, the total demand of refineries for all products is computed in dry tons in cell K5.
- The list of cantons where one refinery already exists and the one where a refinery may be created (green cells). Creation in all other zones is forbidden.

Then four buttons call VBA macros to fill automatically the largest tables: Zones, Distances, Nodes, and “Demands”. The idea is to generate data only for selected departments and products. This is necessary to reduce as much as possible the volume of data.



	A	B	C	D	E	F	G	H	I	J	K
1	AMBRE DASHBOARD										
2											
3	TERRITORY SELECTION	Region	Picardie			Champagne-Ardenne				Refineries	1
4		Department	02	60	80	08	10	51	52	Working period	1-50
5		Select with X			X					Total dry tons	80000,00
6											
7	BIOMASS SELECTION	Product	P1	P2	P3	P4	P5	P6	P7	P8	P9
8		Crop	Colza	Colza	Colza	Céréales	Céréales	Miscanthus	TtCR Saule	Cameline	Moutarde d'E.
9		Form	Graines	Pailles	Menues-Pailles	Pailles	Menues-Pailles	Plante entière	Plante entière	Graines	Graines
10		Select with X	X	X	X	X		X		X	X
11	DEMAND OF EACH REFINERY FOR EACH PRODUCT	Interval 1	31-50	1-50	1-50	1-50		1-50		9-15	16-22
12		Need/period dry tons	900	100	200	360		240		720	680
13		Interval 2	1-8								
14		Need/period dry tons	900								
15		Stock 4 weeks dry tons	3600	400	800	1440		960		2880	2720
16		Need/refinery dry tons	25200,00	5000,00	10000,00	18000,00		12000,00		5040,00	4760,00
17		Total in dry tons	25200,00	5000,00	10000,00	18000,00		12000,00		5040,00	4760,00
18		Total in tRM	28000,00	5882,35	11363,64	21176,47		15000,00		5478,26	5173,91
19		Total in Nodes in tRM	38028,15	9577,46	42253,50	57950,71		111481,12		7875,52	7383,43
20											
21	REFINERY CANTONS	Existing	8018								
22		Allowed	None								

Figure 7. Example of a screenshot of worksheet "Dashboard"

CONCLUSION

This paper proposes a data model which describes and structures the real-world data required for solving the biomass supply chain problem at hand. This data model is meant to serve as a template for the database component of information and decision support systems related to the upstream segment of biomass supply chains. It covers biomass feedstocks, storage facilities (farm storage and centralized storage), biomass production zones, and the demands of bio-refineries. It is flexible enough to add new facilities and new biomass feedstocks. While the dashboard offers a front-end tool for decision-makers, the programming logic needs to be implemented in the back end. The next step in this research is to generate scenarios specific to logistics costs within the multi-biomass supply chain and optimize the scenarios for efficient decision-making. Separating the data model from the mathematical model allows equations to be written more easily and modifications in the data do not necessarily require reciprocating changes in the programming model, therefore rendering a more efficient interface.

ACKNOWLEDGMENTS

This work was performed, in partnership with the SAS PIVERT, within the frame of the French Institute for the Energy Transition (Institut pour la Transition Energétique (ITE) PIVERT, http://www.institut-pivert.com) selected as an Investment for the Future ("Investissements d'Avenir"). This work was supported, as part of the Investments for the Future, by the French Government under the reference ANR-001-01.

REFERENCES

Bachman, C.W., 1975, May. “Trends in database management: 1975,” In Proceedings of the May 19-22, 1975, *National computer conference and exposition*, pp. 569-576.

Bauer, N., Hilaire, J., Brecha, R.J., Edmonds, J., Jiang, K., Kriegler, E., Rogner, H.H. and Sferra, F., 2016. “Assessing global fossil fuel availability in a scenario framework,” *Energy*, 111, pp. 580-592

Chen, P.P.S., 1976. “The entity-relationship model—toward a unified view of data,” *ACM transactions on database systems (TODS)*, 1(1), pp. 9-36.

Chen, Y.S., 1988. “An entity-relationship approach to decision support and expert systems,” *Decision Support Systems*, 4(2), pp.225-234.

Codd, E.F., 2002. “A relational model of data for large shared data banks,” *In Software pioneers*, Springer, Berlin, Heidelberg, pp. 263-294.

De Meyer, A., Almeida, J., Achten, W., Muys, B., Cattrysse, D. and Van Orshoven, J., 2013. “Incorporating life cycle impact assessment in a mathematical model to optimize strategic decisions in biomass-forbioenergy supply chains,” In Fulfilling LCA’s Promise, *Proceedings from the LCA XIII International Conference*, American Center for Life Cycle Assessment: United States, pp. 24-33.

De Meyer, A., Snoeck, M., Cattrysse, D. and Van Orshoven, J., 2016. “A reference data model to support biomass supply chain modelling and optimization,” *Environmental Modelling & Software*, 83, pp.1-11.

Eksioglu, S.D., Zhang, S. and Li, S., 2009. “Analyzing the Impact of Intermodal Facilities to the Design of Supply Chains for Biorefineries,” *In IIE Annual Conference. Proceedings*, Institute of Industrial and Systems Engineers (IISE), pp. 1209.

Freppaz, D., Minciardi, R., Robba, M., Rovatti, M., Sacile, R. and Taramasso, A., 2004. “Optimizing forest biomass exploitation for energy supply at a regional level,” *Biomass and Bioenergy*, 26(1), pp. 15-25.

Frombo, F., Minciardi, R., Robba, M., Rosso, F. and Sacile, R., 2009. “Planning woody biomass logistics for energy production: A strategic decision model,” *Biomass and bioenergy*, 33(3), pp. 372-383.

Gumus, A.T. and Guneri, A.F., 2009. “A multi-echelon inventory management framework for stochastic and fuzzy supply chains,” *Expert Systems with Applications*, 36(3), pp. 5565-5575.

Liu, L., Cheng, S.Y., Li, J.B. and Huang, Y.F., 2007. “Mitigating environmental pollution and impacts from fossil fuels: The role of alternative fuels,” *Energy Sources*, Part A, 29(12), pp. 1069-1080.

Nunes, L.J.R., Causer, T.P. and Ciolkosz, D., 2020. “Biomass for energy: A review on supply chain management models,” *Renewable and Sustainable Energy Reviews*, 120, p. 109658.

Kelly, R.A., Jakeman, A.J., Barreteau, O., Borsuk, M.E., ElSawah, S., Hamilton, S.H., Henriksen, H.J., Kuikka, S., Maier, H.R., Rizzoli, A.E. and Van Delden, H., 2013. “Selecting among five common modelling approaches for integrated environmental assessment and management,” *Environmental modelling & software*, 47, pp. 159-181.

Senko, M.E., Altman, E.B., Astrahan, M.M. and Fehder, P.L., 1973. “Data structures and accessing in database systems, I: Evolution of information systems,” *IBM Systems Journal*, 12(1), pp. 30-44.

Sharma, B., Ingalls, R.G., Jones, C.L. and Khanchi, A., 2013. “Biomass supply chain design and analysis: Basis, overview, modeling, challenges, and future,”.*Renewable and Sustainable Energy Reviews*, 24, pp. 608-627.

Shastri, Y., Hansen, A., Rodríguez, L. and Ting, K.C., 2011. “Development and application of BioFeed model for optimization of herbaceous biomass feedstock production,” *Biomass and Bioenergy*, 35(7), pp. 29612974

Storey, V.C., 1991. “Relational database design based on the Entity-Relationship model,” *Data & knowledge engineering*, 7(1), pp. 47-83.

Sun, O. and Fan, N., 2020. “A review on optimization methods for biomass supply chain: models and algorithms, sustainable issues, and challenges and opportunities,” *Process Integration and Optimization for Sustainability*, 4(3), pp. 203-226.

Zandi Atashbar, N., Labadie, N. and Prins, C., 2018a. “Modelling and optimisation of biomass supply chains:

a review,” *International Journal of Production Research*, 56(10), pp. 3482-3506.

Zandi Atashbar, N., Labadie, N. and Prins, C., 2018b. “A Model to Locate and Supply Bio-refineries in LargeScale Multi-biomass Supply Chains,” *In Operations Research Proceedings 2016*, Springer, Cham, pp. 357-363.

22.

Analysis of Semiconductor Global Shortages and its Effects on US Automotive Industry

Rajkamal Kesharwani

Mercyhurst University, Erie, PA, USA

Abstract

The COVID-19 pandemic and subsequent variants greatly impact the regular supply chain operations of global production for semiconductor and integrated circuit companies. The automotive industry was negatively impacted because of the pandemic and affiliated chip shortages. Most importantly, there is a greater need for a more comprehensive industrial policy that supports the domestic production of semiconductors within the United States of America. It is likely that American companies need to implement and create innovative semiconductor fabrication sites and isolate the market from supply chain issues from abroad (e.g., recessions, wars, or another pandemic). Furthermore, there is a higher demand for materials such as silicon and other rare earth elements to execute successful production of automobiles. Semiconductor improvements and partnerships will likely affect auto-manufacturer production and efficiency. Therefore, the consequences of the semiconductor shortage resulted in brands halting production plants, temporary unemployment, and increased profit loss. The author performed a competitive intelligence analysis based on the data collected from media sources, industry associations, and enterprises, the structure of global integrated circuit market, and the supply/demand of the automotive market. The results showed that:

1) The pandemic control measures, and the re-allocation of market resources are the short-term factors unbalancing the supply and demand for automotive chips.

2) The global shortage of chips of the automotive industry will continue to hinder automotive production and supply chain management.

3) Given the lack of R&D within the auto industry, the current optimal range for chips for automobiles is between 28-68 NM.

Categories

Supply Chain Management and Logistics

28. SUPPLY CHAIN DIVERSITY: FROM PLEDGE TO ACTION

INTEGRATING RACIAL DIVERSITY INTO THE SUPPLY CHAIN: THE UNFORESEEN CHALLENGES

Yahzmine Rodriguez

Graduate Student
Lacy School of Business
Butler University
4600 Sunset Avenue
Indianapolis, IN 46208 yclayton@butler.edu

Janaina Siegler

Assistant Professor of Supply Chain Analytics

Lacy School of Business
Butler University
4600 Sunset Avenue
Indianapolis, IN 46208 jsiegler@butler.edu
317/940-3049

Stephanie A. Fernhaber

Professor of Entrepreneurship & Innovation

Lacy School of Business
Butler University
4600 Sunset Avenue
Indianapolis, IN 46208
sfernhab@butler.edu
317/940-8977

February 2023

INTEGRATING RACIAL DIVERSITY INTO THE SUPPLY CHAIN: THE UNFORESEEN CHALLENGES

ABSTRACT

Organizations are increasingly making commitments and equity pledges to racially diversify their supply chain. While such commitments are commendable, historical attempts at reform have experienced limited success. This suggests that there is a need for organizations to understand the hurdles and complexity that might limit impact, and then be intentional about developing strategies that take these into account. In this study, we conduct an inductive, qualitative investigation involving five organizations in Indianapolis USA. Three are corporations that are actively trying to racially diversify their supply chain (the organization has either made a public commitment to diversity and inclusion in procurement or has a committed department or a member leading diversity and inclusion efforts in procurement). The other two are support organizations that serve as a link between suppliers and corporations. The results identify three levels of interrelated management challenges associated with racially diversifying the supply chain at the industry, strategy, and operation levels. Strategies for overcoming the challenges when integrating racial diversity into the supply chain are discussed, and a set of directional questions are provided to help organizations with their quest to racially diversify their supply chains.

Key Words: Equity pledges, supply chain, challenges, racial diversity, inclusion.

THE RISE OF RACIAL EQUITY PLEDGES

"Lilly will double the national annualized spend with African American suppliers and vendors over the next two years." –Eli Lilly and Company, Indy Racial Equity Pledge

As part of the Indy Racial Equity Pledge launched in October 2020[1], Eli Lilly and Company was one of more than 20 Indianapolis-based organizations in the United States that pledged to advance racial equity in Indianapolis and the neighboring communities. One of the key components of Eli Lilly's pledge involved diversifying the supply chain in a more equitable way and, more specifically, doubling the annualized spend on African American suppliers and vendors. A focus on integrating racial diversity into the supply chain is not unique to Eli Lilly or Indianapolis USA. In 2020, Brother Vellies founder Aurora James challenged American retailers

[1] https://www.indyracialequitypledge.com/

to devote 15% of their shelf space to Black-owned brands given that approximately 15% of the United States population is Black (CBS News, 2021). Many major retailers including Ulta

Beauty, Macy's, and Nordstrom have since responded by committing to this Fifteen Percent Pledge (Penrose & Hall, 2021). Launched in 2021, the Diverse Procurement Collaborative involves 100 of the largest employers in Philadelphia USA, committing to increasing the racial diversity within their contract networks (Allen, 2021). Corporate employers such as Comcast as well as large nonprofit employers including the University of Pennsylvania are making the commitment. Other examples include McDonald's committing 25% of its domestic supply chain to minority and women suppliers by 2025 (Patton, 2021), Kohls's pledge to triple its spending with diverse suppliers by 2025 (Kohls, 2022), and Coca-Cola pledging to double its spend on Black-owned businesses (Atlanta Daily World, 2021). A Gartner Survey in 2020 of 298 supply chain professionals across the United States, Canada, and Europe found that over half of the supply chains had some sort of diversity, equity, and inclusion initiative. While the United States appears to be leading the efforts, the rest of the world is not far behind (Gartner, 2021).

Despite the many pledges and commitments to racially diversify the supply chain, there is some skepticism that the organizations will follow through. Indeed, some firms are already reported to have missed their targets and others criticized for not even developing targets (McGirt, 2022). Historically, this is not the first attempt to create change. There were similar initiatives that have been attempted in "set-aside" programs in the United States dating back to the 1970s to address the gaps in unemployment, enterprise, and racial disparities (Chatterji et al., 2014). Although progress has been made since that time, some contend that the attempts have fallen short, as change has been slower than desired. This suggests that racially diversifying the supply chain may be more complex than some business leaders may have initially thought when making their pledge. Indeed, the decision to make a commitment is influenced by mounting pressure from both internal and external stakeholders (U.S. Chamber of Commerce, 2020).

In this paper, we posit that if business leaders want to successfully move the needle when it comes to integrating racial diversity into the supply chain, there is a need to first understand all the inherent challenges and potential hurdles that stand in its way. When dealing with complex and multifaceted topics such as racial inequity, Gras and colleagues (2020) note that individuals have a reductive tendency that causes them to oversimplify both the causes and proposed solutions. In this case, it can translate to viewing the diversification of the supply chain as a simple solution to issues of racial inequity and the incorrect assumption that implementation is straightforward. Increased knowledge and understanding of the challenges and potential hurdles can balance one's reductive tendency so that the necessary resources can be put forth, thereby increasing the odds of success. This leads to our research question: *What challenges do organizations face when trying to diversify their supply chain in a more equitable manner?*

To explore this question, we conducted an inductive qualitative study of five organizations in Indianapolis, USA. Three organizations are corporations that have publicly committed to addressing racial inequities through diversifying their supply chain. To ensure multiple perspectives, we considered both large publicly held corporations and a nonprofit corporation that has made such a commitment. The other two cases are supporting organizations. They serve as liaisons between smaller, diverse businesses and supply chain professionals. We collected primary data using semi-structured interviews and drew on secondary data from articles and other archival sources. Our findings suggest that there are three interrelated sets of management challenges at the industry, strategy, and operations levels when trying to racially diversify the supply chain. Strategies for how to integrate racial diversity into the supply chain more successfully are discussed.

PREEXISTING DISPARATIES AND A HISTORY

While racially diversifying the supply chain is a topic of interest to companies around the world, the movement appears to be led by firms headquartered in the United States (Gartner, 2021). Thus, it is important to consider the preexisting disparities and unique historical context that are behind it. As of 2019, approximately 60.1% of the American population is reported to be white and non-Hispanic. The remaining minority is split between 18.5% Hispanic, 12.2% Black, 5.6% Asian, and multiple/other races. While all minorities are impacted by initiatives to racially diversify the supply chain, in recent years, there has been higher attention in the media towards the challenges specific to Black business owners (Lesonsky, 2022). Access to financial capital is a major hurdle for all small businesses; however, according to the Federal Reserve (2017), Black owners are turned down for loans at twice the rate of white owners. This financial inaccessibility is attributed to bias and the traditional emphasis on the size of credit needed, capacity, collateral, and conditions of the loans (Frank, 2021). Due to systemic issues, these criteria can be more difficult for Black owners to meet. A study by Jackson and Sanyal (2019) found that Black women who start businesses specifically face negative stereotypes held about them and experience difficulties interacting with Black customers. The additional time needed to navigate these identity issues takes away from their business. Other challenges often faced by Black business owners include a limited social network, lack of support, discrimination, and access to resources (e.g., Jackson, 2021).

Historical Attempts to Support Minority Businesses

Looking back, there have been numerous attempts to “level the playing field” with Black and other minority business owners in the United States. The 1960s is considered to be the most transformational decade for American minority businesses. The first minority business opportunity fair in the country was hosted by Western Electric, now known as AT&T, in Chicago in 1968, sparking a series of civil rights victories such as Richard Nixon's Black Capitalism Program. This program would mark the establishment of the U.S. Small Business Administration, the Minority Enterprise Small Business Investment Company, “minority purchasing” or diverse supplier programs, and the impetus for initial funding under the National Minority Supplier Development program (Moore, 2019). Corporations like AT&T, IBM, Ford, General Motors, and Chrysler would go on to become the leaders of minority purchasing programs; however, the perception was that these programs were more like charity initiatives rather than overall good business practice; therefore, investment with diverse suppliers never really took off as hoped (Moore, 2019).

By the 1970s, cities had begun to experiment with contracting set-aside programs. These programs allocated set percentage goals for government contracts and minority-owned businesses in order to "develop minority enterprise, counter the effects of past discrimination, and reduce the high unemployment rates among urban minorities" (Chatterji et al., 2014, p. 508). These programs were faced with legal and political controversy. The staff director of US

Commission on Civil Rights, J. Al Latham Jr., stated in a 1986 letter to the *New York Times* that

"Set-aside programs are unjust because they pit Americans against one another in a scramble for

Government business based on race and national origin." The dismantling of programs began in 1983, with the US Supreme Court concluding in 1989 that 'generalized assertions" of past racial discrimination could not justify 'rigid' racial quotas for the awarding of public contracts.

Recent Surge in Attention

In recent years, there has again been a surge of attention in the United States toward diversifying the supply chain in a more equitable manner. Unlike the earlier attempts that came through policy making and government contracts, these initiatives are being initiated by corporations through a series of public pledges or commitments to shift their purchasing power to include more Black-owned and minority businesses. Some of the many examples include the Indy Racial Equity Pledge (n.d.), the 15% Pledge by retailers (CBS News, 2021), and the Diverse Procurement Collaborative in Philadelphia (Allen, 2021). These public pledges can be attributed to multiple colliding factors.

As the COVID-19 pandemic spread in 2020, it became apparent that Black Americans were hit harder than others in the nation, in terms of both economic impact and health (Hardy & Logan, 2021). Closures were brought on by drops in consumer demand, government mandates, and concerns for the health and safety of employees. Despite government relief assistance like the Paycheck Protection Program (PPP), businesses owned by African Americans dropped by 41%. Latinx owners followed closely behind, with a drop of 36%, and women-owned businesses, which fell by 25%. According to the Bureau of Labor Statistics, 27% of Black owners operate in industries that have seen the most significant job losses during the pandemic. These include food services, retail, and social assistance (Gould & Wilson, 2020). While every industry has continued to be affected by government mandates, shutdown, and other healthrelated risks, Black entrepreneurs have dealt with the brunt of it. Black entrepreneurs are 90.7% more likely to know someone who has contracted COVID-19 when compared to their white counterparts (Score, 2020). In 2020, there was a wave of protests in the United States, and eventually across the globe, standing against systemic racism and police brutality. This stand against social injustice prompted business leaders to reevaluate their role and to subsequently find ways to increase their engagement with other underrepresented groups (Zimmerman, 2021).

Moreover, COVID-19 exposed the vulnerabilities of operations within global supply chains. With inventories low, worker shortages high, and the coordination and visibility of at-risk supplier relationships still a challenge, the global supply chain found itself in a very reactive scenario where all stakeholders were affected, including consumers. While the long-term effects of the virus are still unknown, corporations and manufacturers are reassessing risks in critical suppliers and business functions in order to build resiliency (Beninger & Francis, 2022; Kouvelis, 2022). Much like the global supply chain, local economies are only as successful as their most vulnerable members. It is evident that the cost of multi-sourcing and maintaining alternate suppliers should be considered in the cost of doing business, with racially diverse suppliers being an untapped resource. The attitudes of procurement managers toward supplier diversity programs have a significant influence over contracting decisions and thus can either help or hinder closing the gaps of inequalities (Blount & Li, 2021).

In sum, business leaders have determined that diversifying the supply chain in a more equitable way is important for society and for business. While the pressure from both internal and external stakeholders has led to numerous pledges to racially diversify the supply chain, we posit that there is a need to understand the associated challenges from a corporate perspective in order to be intentional about making the pledges a reality.

SUPPLIER DIVERSITY PLEDGES IN INDIANAPOLIS: A MULTIPLE-CASE STUDY

The goal of this study is to explore what challenges organizations face when trying to integrate racial diversity into the supply chain. An inductive qualitative approach with cross-case and content analysis was selected due to gaps in data concerning diversity and inclusion in procurement (Benbasat et al., 1987). Data was collected using semi-structured interviews in five selected cases. We followed the field-based data collection methods described by Eisenhardt (1989), Pagell (2004), and others, whose goals are to "not only identify constructs, but also develop an understanding of why identified constructs might be important" (Pagell, 2004, p. 464). Suggestions from Krippendorff (2013), Miles et al. (2014), and Yin (2013) were also taken into consideration in selecting and analyzing the cases. Cross-case analysis was used to compare the experiences of the participants due to the exploratory nature of this study in order to build theories around the findings (Meredith, 1998). The use of content analysis was justified to develop focus around challenges surfacing among organizations undertaking diverse supplier initiatives. This creates better consistency in the data collection process and analysis (Benbasat et al., 1987; Krippendorff, 2013).

Organizations Studied

The five participating organizations in this study are designated as Company A, B, and C, and Support A and B. To maintain a level of consistency, all five organizations are headquartered in Indianapolis, Indiana USA. The state of Indiana has a diverse representation of industries, including advanced manufacturing, aviation, agriculture, life sciences, technology, and logistics. It has made a concerted effort to attract businesses, highlighting rankings such as first among the states in manufacturing, fifth best for business, and second best for property tax rates in the nation (Indiana Economic Development Corporation, 2022). There are two types of organizations represented in this study: those that directly hire and manage suppliers (corporations/companies) and those that support the relationships between suppliers and corporations (support organizations). The cases were chosen considering the importance of gathering more than one perspective concerning diverse supplier programs, and the recognized importance of support organizations (Adobor & McMullen, 2007). In addition, the selection of the organizations for this study was based on two main criteria: (a) the organization has made a public commitment to diversity and inclusion in procurement, and (b) the organization has either a committed department or a member leading diversity and inclusion efforts in procurement.

Company A and B were chosen because they had made public pledges in 2020 to address equity challenges in Indianapolis USA, particularly in their contracting efforts for suppliers. Company C had been publicly recognized for its efforts in setting aggressive goals for hiring diverse and inclusive suppliers for its development program. Support A and Support B were chosen because they help facilitate relationships between suppliers and corporations through training, education, grants, and research. The individuals interviewed for this study were all directors of departments representing equity, operations, strategic initiatives, or partnerships.

As shown in Table 1, data collected from these five cases includes type, industry, revenue, size, founding, interview duration, and transcript length. Company A is a publicly traded pharmaceutical corporation with both national and international locations and gross revenue of $24 billion. This corporation employs more than 32,000 employees worldwide. Company B is a publicly traded diesel and natural-gas engine manufacturer with national and international locations and gross revenue of $19 billion. This manufacturer has more than 52,000 employees. Company C is an Indiana social enterprise with more than 10 employees. Support A and Support B are Indiana nonprofits with more than 50 and 10 employees.

[Insert Table 1 about here]

Data Collection and Organization

The data used in the analysis of this study came primarily from interviews with these five organizations in addition to media articles and databases used to provide financial and organizational profile data. To maintain consistency, a set of seven questions was prepared in a semi-structured interview format. These can be found in Appendix A. The questions were developed based on the theoretical background and updated after a round of expert consultations with supporting organizations. Each interview was conducted through Zoom calls and recorded. The interviews lasted between 30 and 60 minutes. The recordings were later transcribed with Microsoft tools and added as individual cases. Finally, the transcripts were organized and coded using QDA Miner software, along with their respective variable fields.

Data Analysis & Results

An inductive approach was adopted in the data analysis to understand what challenges organizations face when trying to diversify their supply chain in a more equitable manner. Inductive analyses are more complex and require extra effort in the coding process to guarantee consistency and data reliability. However, they are richer processes in allowing meaning-making, developing findings, and generating evidence using representative quotes (Bingham &

Witkowsky, 2021; Miles et al., 2014).

The first step in the inductive analysis was to perform some initial open coding, searching for early patterns. This phase was performed using iterative sessions. First, each of the authors did analysis separately, and then the entire group met to discuss possible divergences and complementariness in the search for emerging patterns. Multiple iterative sections were held until we agreed with the patterns, themes, and findings. This process is in line with Bingham and

Witkowsky (2021), as illustrated in Figure 1.

[Insert Figure 1 here]

Three general dimensions were built in QDA Miner as placeholders for the codes that would later emerge from the iterative and interactive coding sessions. In the first round of building the codes, the focus was to identify all the themes and their occurrences. In the second round, the codes were then consolidated. In the third round, subthemes were identified, and the categories were reevaluated to see if they were still representative of the transcripts. In the first round, three categories, three subcategories, and 34 codes were identified from the transcripts. The final analysis includes three categories, seven subcategories, and 18 codes, as illustrated in

Figure 2.

[Insert Figure 2 here]

Following guidance from Krippendorff (2013), several analytical and representational techniques were used in the data analysis. The first involved word count and frequency analysis to understand the frequency of code occurrences in each of the cases, codes, and categories. Figure 3 illustrates the coding frequency tree with information related to the count of how many times a certain code appeared, its overall percentage, in how many of the cases it appeared, and the total number of words it involved.

[Insert Figure 3 here]

When analyzing the distribution of the codes by frequency, it is possible to note that of the top 10 most-cited challenges, eight are related to operational challenges to search, onboard, and manage talent, as illustrated in Figure 4.

[Insert Figure 4 here]

CHALLENGES WHEN PURSUING SUPPLIER DIVERSITY PLEDGES

Our findings identified three dimensions of challenges experienced by organizations when taking on procurement initiatives to racially diversify their supply chain. Those dimensions include challenges at three levels: Industry, Strategy, and Operations. As summarized in Table 2, those main challenges are also broken down into themes: supporting collective growth and addressing barriers for racially diverse suppliers (Industry); measuring impact and making the commitment (Strategy); and searching for, onboarding, and managing talent (Operations). When analyzing the distribution of the codes by frequency, we noted that of the top 10 most-cited challenges, eight are related to operations-level challenges to search, onboard, and manage talent.

[Insert Table 2 about here]

#1: Industry-Level Challenges

At the broadest level, challenges in pursuing pledges to racially diversify the supply chain exist at the industry level. On the one hand, there are barriers for corporations. The lack of historically documented and proven models makes it less of a straightforward strategy for corporations to implement racially diverse supplier programs, creating situations that require them to start from scratch. Likewise, it is a challenge to find enough qualified racially diverse suppliers despite the various nonprofit supplier advocacy associations that serve as directories and certification resources for suppliers. As noted by Support A, the question faced by organizations within a highly regulated industry could be, "so how many Black-owned developers or engineering firms do we have? ... that's what you're running into, it's not that there aren't that many [racially diverse suppliers]. It just depends on the area or the field." In addition, there is a higher level of systemic complexity faced by racially diverse suppliers that corporations do not fully understand, including a history of economic exclusion, gentrification, discrimination and racial violence faced by this group.

On the other hand, there are also barriers from the perspective of racially diverse suppliers. The minority certification processes that are commonly required for identifying diverse suppliers can be viewed as an extra hurdle to go through. As observed by Company C,
"we have businesses that are owned by women or people of color who for whatever reason have elected not to go through the city certification process." Another challenge is the lack of infrastructure. Many large corporations are global in nature, and many racially diverse suppliers that are small in scale may not have the experience or infrastructure to service the corporations at a global level.

#2: Strategy-Level Challenges

The second level identified in this study are strategy-level challenges which are those related to the controls and decisions made by the upper management. At this level, there are challenges associated with both making the initial commitment to pursue a racial diverse supplier initiative and in measuring the impact once an initiative is underway. When a corporation first makes a commitment, it needs to start by defining impact. This can be a challenge because there are many different definitions and assumptions about what is meant by "diverse." An example was noted by Company B: "It's because it's global.

What qualifies as diverse is contingent upon the region in the world." In order to get employees on board with such a commitment, there is also a need to shift from data and statistics to a humanization of the issue. Support A reflected, "I don't want to say [we were] oblivious to the problem because I think they've known where their numbers have been for a long time. So, they knew there was a problem, but I think the human element was completely not in it."

Once a racial diverse supplier initiative is being pursued, the other major challenge faced by management is associated with the measurement of impact. This is especially important for larger and more visible companies making such a pledge. There is a need to develop a system of reporting, in which racial supplier diversity is tracked and measured. Of course, this also relates back to how racial supplier diversity is measured upfront with the commitment. In addition to reporting, there is a level of internal accountability that needs to be addressed. Company A noted that "procurement reports are now at the procurement director's level with additional goals for each director, holding procurement more accountable."

#3: Operations-Level Challenges

There are many additional challenges associated with a racial diverse supplier initiative that exist at the operations-level. Foremost, participants of this study expressed challenges associated with searching for talent and the time, effort and coordination needed to leverage partnerships. As noted by Support B, supplier diversity "isn't something that companies can solve on their own." Expanding the network of those working in procurement to include more diversity is also a necessity, as otherwise it can be challenging to find those racially diverse suppliers. In addition, searching for talent requires additional communication with potential suppliers. This is exemplified by a director from Company A, who shared that the company provides procurement opportunity reports to better communicate and showcase where diverse suppliers might be a good fit.

Once potential racially diverse suppliers are identified, a second set of challenges exists with onboarding the talent. The bidding process commonly needs to be revisited, sometimes breaking it down into smaller pieces. Company C explained, "sometimes it's a capacity issue." The diverse suppliers can sometimes also require coaching to help prepare them for working with a larger corporation. As noted by Company A, "we have a nine-month program that they'll go through to help them anywhere from, you know, financial workshops, marketing workshops,

HR workshops, strategy, and then one-on-one coaching with our executives." Finally, engagement cannot always be assumed. Even though resources and opportunities can be put out there, the diverse suppliers cannot attend to everything.

A third and final set of challenges at the operational level is associated with managing the talent. Bringing on additional suppliers requires more time and effort by the procurement team. Moreover, if the suppliers are smaller in size, this can also equate to a larger number of suppliers to manage. There is also a challenge associated with the existing suppliers, as maintaining the status quo is commonly seen as easier and more efficient. Company A reflected that "the initial thought is to keep doing business with suppliers that

already know and already have a better reputation at [company] and who are already under contract. This is the easiest and quickest way
... we have to do things quickly." Thus, there is also a need to create new norms. These new norms should not only be ones in which racially diversifying the supply chain is encouraged but also supported with adjusted expectations regarding short-term efficiencies.

IMPROVING THE SUCCESS OF SUPPLIER DIVERSITY PLEDGES

The number of organizations pledging to increase racial diversity in their supply chain has increased considerably over the last few years. Nonetheless, commitments alone are not a guarantee that we will see major evolution in the diversification of the supply chain as they relate to working with minority-owned or small businesses. The desire to increase racial diversity and the willingness to make public commitments are commendable (Bateman et al., 2020), but there is a need to understand the systemic challenges that exist at different levels and from different perspectives. To sustain their commitments, organizations should seek out a full-time Diversity, Equity and Inclusion (DEI) leader, focus on providing additional recruitment resources and support to drive pipeline and scale a task force or internal champions committed to progress (Harvard Business Review, 2021). Building a culture of diversity and inclusion within an organization will put good intentions into action (Seijts & Milani, 2022).

As summarized in Figure 5, we identified three main levels of management challenges when implementing racially diverse supply chain initiatives. The triangular shape reflects the magnitude of the effort required at this operational level, but the success of the efforts is also dependent also on the strategy and industry levels. At the top is the industry level. Some industries are naturally more challenging than others to be more diverse due to their intrinsic nature. For example, it might be easier and more straightforward to hire janitorial companies owned by minority groups, but it may not be as easy to find the same availability from an engineering or medical perspective. Highly regulated industries, such as pharmaceuticals and aerospace and defense, will naturally face more challenges in their efforts to diversify their supply chains. These are structural limitations and illustrate systemic complexities that increase the challenges for both the companies trying to diversify their supply chains and suppliers trying to enter or be successful in that industry.

[Insert Figure 5 here]

After being able to understand and address industry challenges, the next step is to focus on the organization's strategy-level challenges. Strategy-level challenges relate to making the commitment to racial diversity and then measuring its impact. One of the first difficulties to emerge is the definition of impact, as it can mean different things for different organizations. In addition to the definition, it is important that the organizations understand that this diversification only really makes a societal impact if there is a recognition and commitment to humanizing the efforts. Moreover, it is also management's responsibility to commit to the process, collect data relevant to the diversity efforts, and report on the developments and impacts achieved over time.

Finally, the operations-level challenges are where we see things coming into place and enabling (or preventing) an impact. These efforts are related to searching for, onboarding, and managing talented suppliers. In searching for racially diverse talent, there is a value in leveraging, communicating, and expanding different network opportunities that may have the potential to help identify these potential talented suppliers. Once talented suppliers are identified, the process to onboard them needs to receive significant attention. Often, onerous paperwork tasks and certifications need to be prepared before any actual business is completed. There are several different examples of diverse suppliers who start certification programs but give up before becoming fully certified due to the hurdles related to time, cost, and bureaucracy. It is important that companies share the responsibility and take actions to coach the pool of diverse suppliers, engage with them, and even help with the bidding process, as these areas are going to be crucial for supplier diversity success. The key relevance of these issues is also demonstrated in the data from our research, as 80% of the most-cited challenges are related to operational challenges connected to searching for, onboarding, and managing the talent pool of diverse suppliers. This seems to suggest that to improve diversity, a "top-to-bottom" approach may not be the most appropriate one. To be successful, publicly committing to supplier diversity is a good beginning, but making sure that all internal challenges are recognized and addressed will eventually be the make-or-break part of the deal for a diverse supply chain base.

As displayed in Appendix B, a checklist for overcoming common management challenges when integrating racial diversity into the supply chain has been developed based on our research. This checklist is an organizational tool designed to address and bring awareness to the challenges in each of the three dimensions identified in this study. Foremost, it is important that organizations invest the necessary time and effort to understand the opportunities and specific challenges within their industry. They also need to make sure they have clarity on what they want to accomplish and have commitment from management. Most important, diverse supplier initiatives must take the proper actions on the processes related to searching for, onboarding, and managing a pool of talented suppliers. Ignoring any one of the challenges attached to these dimensions can create future barriers to efficiency and effectiveness as expressed by the participant interviews.

The checklist can also be used to improve or scale programs that already exist. This tool considers that the success of a racially diverse supplier program is contingent upon the investment of all levels of management within an organization and its suppliers. It involves both an internal and external focus. Foremost, it should be used by DEI leaders, teams, committees, executives, and management to uncover or better understand the challenges of running diverse supplier initiatives with the intent of building actionable strategies to overcome them. For those corporations that already have a diverse supplier program, it can be used as an exercise to evaluate and improve upon their efforts. Support organizations can use this checklist as a tool to not only evaluate their own programs, but also start the dialogue between organizations and diverse suppliers. Understanding the challenges and developing actionable solutions can both help inform and enable them to develop strategies that better support stronger relationships. In this way, it can solidify better forms of collaboration as well.

While relationships between organizations, supporting organizations, and racially diverse suppliers grow, leadership should resist cutting diversity and inclusion programs especially during economic uncertainty. According to a Conference Board survey, as of August 2022, 41% of adult respondents believe the US has already entered into a recession. Thirty-three percent believe a recession is inevitable and not yet here (Colvin, 2022). Organizations looking to start or scale a program despite this fear should view their diverse supplier program as a long-term strategy that involves “mutual learning, smaller-scale pilots, and, eventually, broader partnerships” (Prilepok, 2022). Supplier diversity initiatives should be evaluated with a financial and equitable lens as leadership focus moves towards driving revenue more than ever.

Taking on Procurement Diversification Commitments

Our findings suggest that integrating racial diversity into the supply chain is a complex task. Even very large and resourceful companies that have been investing in similar actions for over a decade still face many challenges. The commitment is important but not enough to guarantee the success of a supply chain diversity program. Diverse supplier programs should be strategic initiatives that drive value for communities and businesses. They require a major shift from leadership that trickles down into the everyday culture of an organization, therefore making it a collaborative effort. In understanding the complexity of the task, we offer some actionable steps, such as the checklist for bringing awareness and overcoming common management challenges in the industry, strategy, and operations levels when integrating racial diversity into the supply chain.

REFERENCES

Adobor, H. & McMullen, R. (2007). Supplier diversity and supply chain management: A strategic approach. *Business Horizons,* 50: 219-229.

Allen, T. (2021). 100 Philly companies—including Comcast and Wawa—vow to diversify supply chains. Whyy.org. https://whyy.org/articles/100-philly-companies-including-comcast-and-wawa-vow-tohttps://whyy.org/articles/100-philly-companies-including-comcast-and-wawa-vow-to-diversify-supply-chains/diversify-supply-chains

Atlanta Daily World. (2021, July 5). Coca-Cola is latest major company to pledge more ad spending with Black-owned media. AtlantaDailyWorld.com. https://atlantadailyworld.com/2021/07/05/coca-cola-ishttps://atlantadailyworld.com/2021/07/05/coca-cola-is-latest-major-company-to-pledge-more-ad-spending-with-black-owned-media/latest-major-company-to-pledge-more-ad-spending-with-black-owned-media

Bateman, A., Barrington, A., & Date, K. (2020). Why you need a supplier-diversity program. *HBR*.org.

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information systems. *MIS Quarterly*, 11(3), 369–386.

Beninger, S. & Francis, J.N.P. (2022). Resources for business resilience in a COVID-19 world: A communitycentric approach. *Business Horizons*, 65: 227-238.

Bingham, A. J., & Witkowsky, P. (2021). Deductive and inductive approaches to qualitative data analysis. In "*Analyzing and interpreting qualitative data: After the interview*" Edited by Vanover, Mihas and Saldana. Sage Publications: 133–146.

Blount, I. and Li, M. (2021), How Buyers' Attitudes Toward Supplier Diversity Affect Their Expenditures with Ethnic Minority Businesses. J Supply Chain Manag, 57: 3-24. https://doi.org/10.1111/jscm.12237

CBS News. (2021). Meet the woman on a mission to have retailers commit 15% of their shelf space to Blackowned brands. CBSNews.com. https://www.cbsnews.com/news/15-percent-pledge-aurora-james

Chatterji, A. K, Chay, K. Y., & Fairlie, R. W. (2014). The impact of city contracting set-asides on black selfemployment and employment. *Journal of Labor Economics*, *32*(3), 507–561. https://doi.org/10.1086/675228

City of Richmond v. J. A. Croson Company. 488 US 469 (1989). Oyez. https://www.oyez.org/cases/1988/87https://www.oyez.org/cases/1988/87-998998

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532–550.

Fain, K. (2017). The devastation of Black Wall Street. JSTOR Daily. https://daily.jstor.org/the-devastation-ofblack-wall-street

Fairlie, R. W. (2020, July 6). *The impact of COVID-19 on small business owners: The first three months after social-distancing restrictions*. NBER. https://www.nber.org/papers/w27462

Federal Reserve. (2017). Report to the congress on the availability of credit to small businesses.

https://www.federalreserve.gov/publications/2017-september-availability-of-credit-to-smallbusinesses.htm

Frank, A. (2021). 'There's work to do': Black-owned businesses need connections, infrastructure to access capital, other resources. *Crain's Detroit Business, 37*(13), 8. https://ezproxy.butler.edu/login?url=https://www.proquest.com/trade-journals/theres-workhttps://ezproxy.butler.edu/login?url=https://www.proquest.com/trade-journals/theres-work-do/docview/2509704268/se-2?accountid=9807do/docview/2509704268/se-2?accountid=9807

Gartner. (2021). Gartner survey finds over half of supply chain organizations have DEI initiatives, but only 23% commit to formal targets. Gartner.com.

Gould, E., & Wilson, V. (2020, June 1). *Black workers face two of the most lethal preexisting conditions for coronavirus-racism and economic inequality*. Economic Policy Institute. https://www.epi.org/publication/black-workers-covid

Gras, D., Conger, M., Jenkins, A., & Gras, M. (2020). Wicked problems, reductive tendency, and the formation of (non-)opportunity beliefs. *Journal of Business Venturing*, *35*(3). https://doi.org/10.1016/j.jbusvent.2019.105966

Harvard Business Review. (2021). *Creating a culture of diversity, equity, and inclusion*. Harvard Business Review Analytic Services.

Indiana Economic Development Corporation. (2022). Good for business. IEDC.

https://www.iedc.in.gov/indiana-advantages/good-for-business

Indy Racial Equity Pledge. (n.d.). Retrieved January 20, 2021, from https://www.indyracialequitypledge.com Jackson, T. M. (2021). We have to leverage those relationships: How Black women business owners respond to limited social capital. *Sociological Spectrum, 41*(2), 137–153.

Jackson, T. M., & Sanyal, P. (2019). Struggles and strategies of black women business owners in the U.S. *Journal of Business Anthropology*, *8*(2), 228–249.

Kohls. (2022, May 4). Kohl's expands supplier diversity initiative with pledge to triple spending with diverse suppliers by 2025. https://corporate.kohls.com/news/archive-/2022/may/kohl-s-expands-supplier-diversityhttps://corporate.kohls.com/news/archive-/2022/may/kohl-s-expands-supplier-diversity-initiative-with-pledge-to-tripinitiative-with-pledge-to-trip

Kouvelis, P. (2022). Paradoxes and mysteries in virus-infected supply chains: Hidden bottlenecks, changing consumer behavior, and other non-usual suspects. *Business Horizons*, 65: 469-479.

Krippendorff, K. H. (2013). *Content analysis: An introduction to its methodology*. Sage.

Latham, J. Al. *The New York Times*. (1986, May 2). Set-aside system for minority businesses is flawed.

https://www.nytimes.com/1986/05/02/opinion/l-set-aside-system-for-minority-businesses-is-flawedhttps://www.nytimes.com/1986/05/02/opinion/l-set-aside-system-for-minority-businesses-is-flawed-689086.html689086.html

Lesonsky, R. (2022, May 24). Black Entrepreneurs Face Their Challenges and Feel Optimistic.

https://www.score.org/resource/blog-post/black-entrepreneurs-face-their-challenges-and-feel-optimistic

McGirt, E. (2022, January 25). Why your corporate supplier diversity efforts are not going as well as you think. Fortune.com. https://fortune.com/2022/01/25/your-corporate-supplier-diversity-efforts-are-notgoing-as-well-as-you-think

Meredith, J. (1998). Building operations management theory through case and field research. *Journal of*

Operations Management, *16*(4), 441–454. doi:10.1016/S0272-6963(98)00023-0

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.

Moore, R. G. (2019). Supplier diversity: Yesterday, today, tomorrow. *Minority Business Entrepreneur, 36*(Summer), 24–32.

Pagell, M. (2004). Understanding the factors that enable and inhibit the integration of operations, purchasing and logistics. *Journal of Operations Management, 22*(5), 459–487. doi:10.1016/j.jom.2004.05.008

Patton, L. (2021). McDonald's pledges to boost spending at minority-owned suppliers. Bloomberg.com.

Penrose, N., & Hall, C. (2021). Ulta, Macy's, and Bloomingdales, join Aurora James's 15 percent pledge. Elle.com.

Perry, A. M., See, R., Barr, A., Romer, C., & Broady, K. (2022). Black-owned businesses in U.S. cities: The challenges, solutions and opportunities for prosperity. *Brookings Report.*

Prilepok, M., III, S. S., Yearwood, K., & Zegeye, A. (2022). *Expand diversity among your suppliers--and add value to your organization*. Mckinsey.com.

Score. (2020). The megaphone of main street: The impact of COVID-19 Infographic #2—Impact on Black owned small businesses. Score.org. https://www.score.org/resource/infographic/megaphone-main-streethttps://www.score.org/resource/infographic/megaphone-main-street-impact-covid-Black-owned-small-businessesimpact-covid-Black-owned-small-businesses

Seijts, G.H. & Milani, K.Y. (2022).The application of leader character to building cultures of equity, diversity and inclusion. *Business Horizons,* 65: 573-590.

U.S. Chamber of Commerce. (2020). Why businesses should strengthen their supply chain diversity. USChamber.com.

Walker, J.E.K. (1998) *The History of Black Business in America: Capitalism, Race, Entrepreneurship.* Twayne Publishers: New York.

Weems, R. E., & Randolph, L. A. (2001). The national response to Richard M. Nixon's Black capitalism initiative: The success of domestic detente. *Journal of Black Studies*, *32*(1), 66–83. http://www.jstor.org/stable/2668015

Yin, R. K. (2013). *Case study research: Design and methods*. Sage.

Zimmerman, S. (2021). Supplier diversity importance grows, but hurdles remain. SupplyChainDive.com

Table 1: Summary of Cases

	Company A	**Company B**	**Company C**	**Support A**	**Support B**
Type	Publicly Traded	Publicly Traded	Nonprofit	Nonprofit	Nonprofit
Industry	Pharmaceuticals	Diesel and Natural Gas Engine Manufacturer	Innovation District	Economic Development	Manufacturing and Logistics Support
Revenue	$24 B	$19 B	N/A	N/A	N/A
Size (Employees)	35,000	57,825	10+	50+	10+
Founded	1876	1919	2011	1850	2007
Interview Length (minutes)	51:01	33:52	54:24	29:42	52:24
Transcript Length	11 pages	7 pages	10 pages	8 pages	8 pages
Secondary Data	Company reports, website, news articles	Company reports, website, news articles	Company reports, website, news articles	Company reports, website, news articles	Company reports, website, news articles

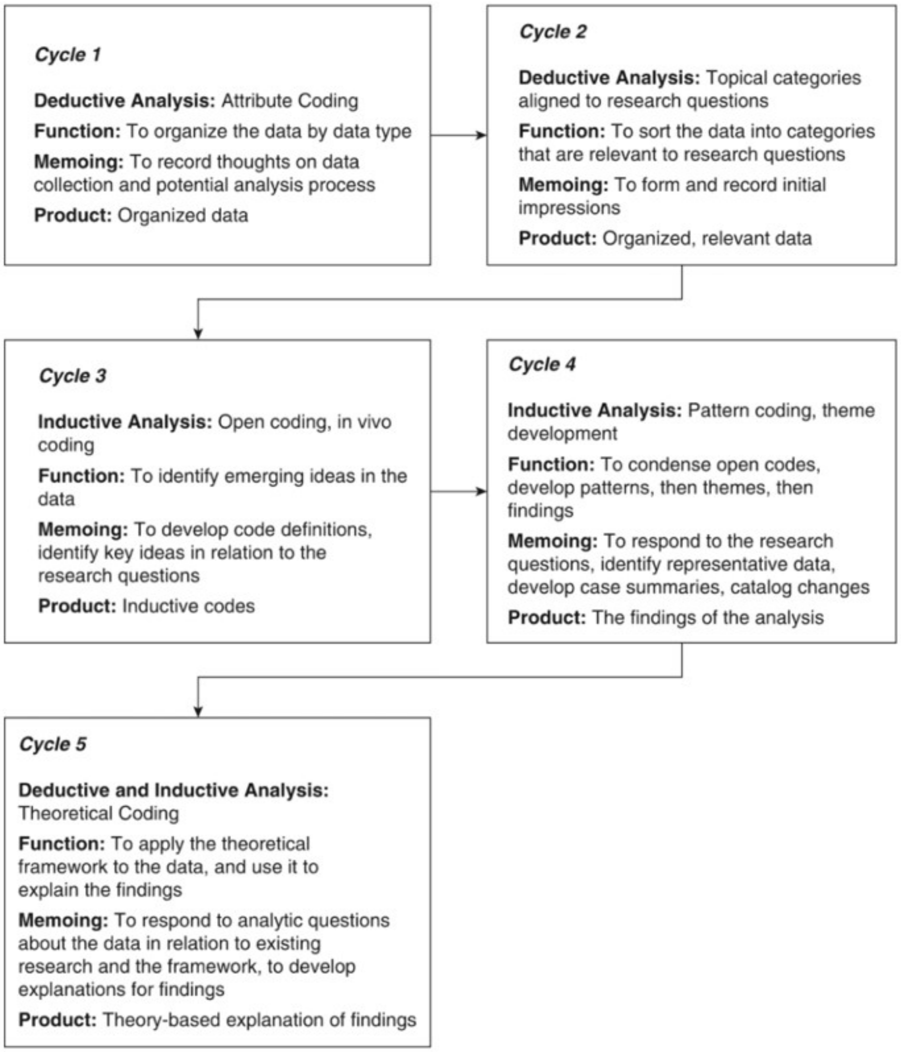
Table 2: Data Structure

DIMENSION	THEME	CODE	EXEMPLARY QUOTE
Industry-Level Challenges	**Barriers for Corporations**	**Limited Models**	There weren't really any good benchmarks out there at the time about how to do this.
		Finding Diverse Suppliers	How many Black-owned developers or engineering firms do we have? ...that's what you're running into, so it's not that there aren't that many. It just depends on the area or field.
		Systemic Complexity	[We] realized, you know what? This is overly complicated and, you know, it requires a lot of judgment.
	Barriers for Diverse Suppliers	**Certification**	We have businesses that are owned by women or people of color who for whatever reason have elected not to go through the city certification process.
		Infrastructure	What opportunities are there to go global with some of the bigger Black-owned businesses that currently exist ... is it easy for them to go global?
Strategy-Level Challenges	**Measuring Impact**	**Internal Accountability**	Procurement reports are now at the procurement director's level w/ additional goals for each director, holding procurement more accountable.
		Reporting	We also look specifically at what business is the percentage of our operating budget that we spend with companies that are headquartered in the neighborhoods...we have a categorization that we track/measure there.

22

	Making the Commitment	Defining Impact	It's because it's global. What qualifies as diverse is contingent upon the region in the world.
		Humanizing the Issue	I don't want to say [we were] oblivious to the problem because I think they've known where their numbers have been for a long time. So, they knew there was a problem, but I think the human element was completely not in it.
Operations-Level Challenges	Searching for Talent	Leveraging Partnership	This isn't something that companies can solve on their own. This is something that's a very collaborative effort called a public–private partnership.
		Expanding the Network	I worked in the private sector. You know I have all these different groups that I've been a part of. And then I suddenly realized, oh yeah, my network is diverse, but it's not diverse in the way that it needs to be.
		Communicating with Suppliers	We also provided procurement "Opportunity Reports" based on multiple marketplace data points of areas with the greatest opportunities.
	Onboarding Talent	Bidding Process	Sometimes it's a capacity issue ... in terms of preparing the bid. ...You know it's sometimes a little bit like Goldilocks, you know, sometimes it's too big, sometimes it's too small.
		Coaching	We have a nine-month program that they'll go through to help them anywhere from, you know, financial workshops, marketing workshops, HR workshops, strategy, and then one-on-one coaching with our executives.
		Engagement	40% of the attendees showed. 60% didn't even show up, and it was the executives that were hiring them. So, it's an interesting conundrum. You can put all these resources and throw all this stuff at them, and for whatever reason it could be considered lean on staff, and they can't attend everything.
	Managing Talent	Additional Suppliers	Bringing on another new supplier will slow the process down. So, you have got that working against you. You also have the idea, of from an occurrence standpoint, you tend to consolidate your suppliers. You want to do business with less suppliers, not necessarily more different spaces.
		Creating new norms	You know, part of it [supplier diversity] is just making it part of the culture.
		Existing Suppliers	The initial thought is to keep doing business with suppliers they already know and already have a better reputation at [COMPANY] and who are already under contract? That's the easiest and quickest way ... we have to do things quickly.

Figure 1: *Overview of The Coding and Analysis Process*



Source: Bingham and Witkowsky, 2021

Figure 2: *Final Code Book*

Source: Bingham and Witkowsky, 2021

Figure 3: *Coding Frequency Tree*

	Count	% Codes	Cases	% Cases	Nb Words	% Words
Industry Level Challenges						
Barriers for Corporations						
Limited Models	10	4.3%	5	100.0%	446	1.5%
Finding Diverse Suppliers	12	5.2%	5	100.0%	519	1.7%
Systemic Complexity	5	2.2%	4	80.0%	229	0.8%
Barriers for Diverse Suppliers						
Certification	8	3.5%	3	60.0%	350	1.2%
Infrastructure	6	2.6%	4	80.0%	237	0.8%
Management Level Challenges						
Measuring Impact						
Internal Accountability	11	4.8%	3	60.0%	518	1.7%
Reports	17	7.4%	4	80.0%	780	2.6%
Making the Committment						
Defiing Impact	12	5.2%	5	100.0%	596	2.0%
Humanizing the Issue	5	2.2%	2	40.0%	242	0.8%
Operational Level Challenges						
Searching for Talent						
Leveraging Partnerships	24	10.4%	4	80.0%	1413	4.8%
Expanding Network	16	6.9%	4	80.0%	865	2.9%
Communicating with Suppliers	3	1.3%	2	40.0%	105	0.4%
Onboarding Talent						
Bidding Process	15	6.5%	3	60.0%	888	3.0%
Coaching	22	9.5%	5	100.0%	913	3.1%
Engagement	21	9.1%	5	100.0%	1747	5.9%
Managing Talent						
Additional Suppliers	26	11.3%	5	100.0%	1544	5.2%
Creating New Norms	13	5.6%	4	80.0%	568	1.9%
Existing Suppliers	5	2.2%	2	40.0%	260	0.9%

Figure 4: *Distribution of Codes By Frequency*

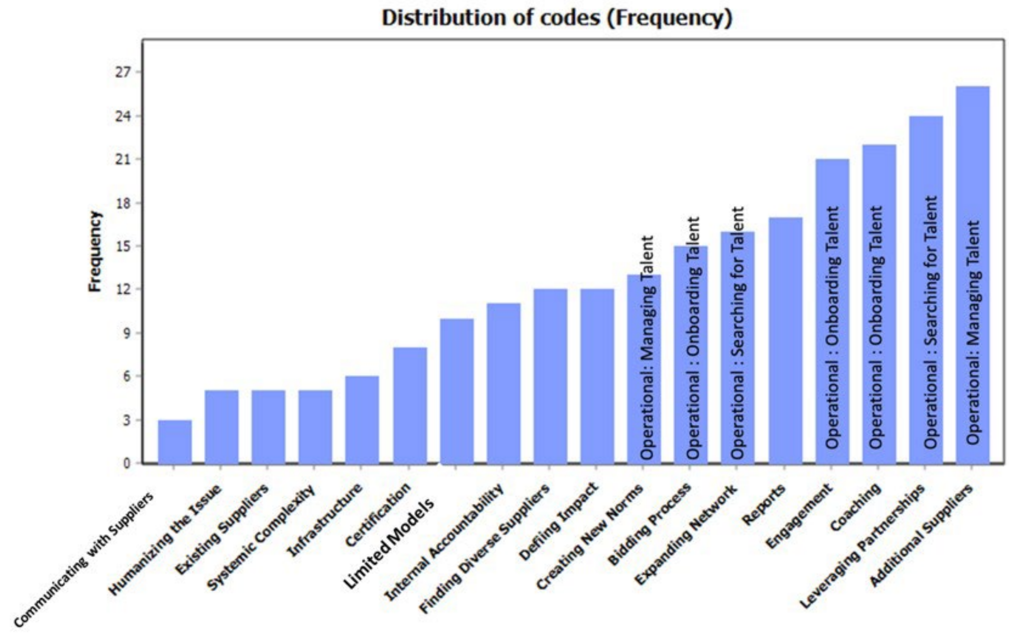
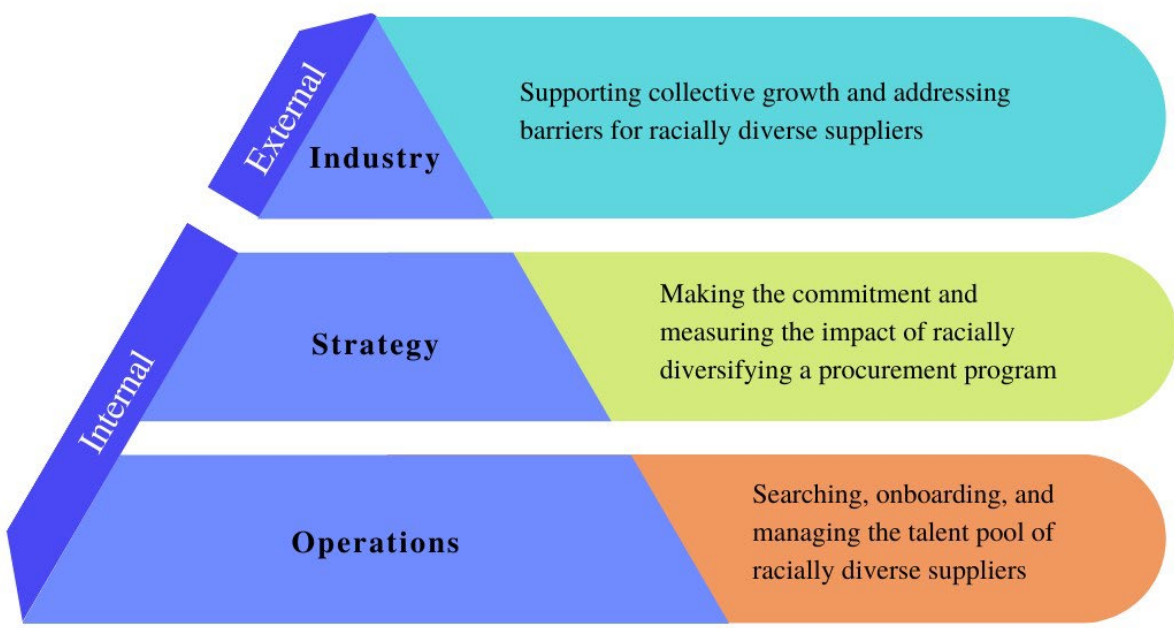


Figure 5: *Management Challenges When Implementing Racially Diverse Supplier Programs*



APPENDIX A: SEMI-STRUCTURED INTERVIEW QUESTIONS

1. Can you talk a little about your background, current role and what your day-to-day looks like?
2. How are you involved in the collaboration with internal and/or external stakeholders?
3. Can you talk about your organization's diverse supplier pledge? (Corporations only)
4. Can you talk about your organization's commitment to supplier diversity? (Support organizations only)
5. What is your organization's strategy for expanding your diverse supplier initiatives? (Corporations only)
6. What is your organization's strategy for addressing disparities for diverse suppliers? (Support organizations only)
7. What do you believe to be the challenges or barriers of building a more equitable supply chain? How much does that affect the community and supply chain resiliency?
8. What does the procurement bidding process look like for one of your suppliers? (Corporations only)
9. How is your organization measuring impact and success?

APPENDIX B:

INTEGRATING RACIAL DIVERSITY INTO THE SUPPLY CHAIN:
A CHECKLIST FOR OVERCOMING COMMON MANAGEMENT CHALLENGES

INDUSTRY-LEVEL

Supporting Collective Growth

☐ **Benchmark Diverse Supplier Initiatives.** What can we learn from other efforts to racially diversify the supply chain in our industry? Can we share our learnings in some way, perhaps at an industry conference or shared report?

☐ **Get Involved with Support Organizations.** What groups specialize in supporting diverse suppliers in our industry, and how can we get more involved? Reach out to your industry trade association as a starting point.

☐ **Explore Systemic Issues.** What systemic challenges and barriers are diverse suppliers facing within our industry? What might we do to address these issues?

Addressing Barriers for Diverse Suppliers

☐ **Understand Certification Process.** What is the process for racially diverse suppliers to get certified in our industry? What about government-based certifications (e.g. Minority Business Enterprise certification in the United States)? How can we help?

☐ **Consider Infrastructure Requirements** What infrastructure and scalability must suppliers have in place to best meet the needs of industry firms? How can we help?

STRATEGY-LEVEL

Making the Commitment

☐ **Clearly define what is meant by supplier diversity.** How does our organization define diverse suppliers? How does this compare to other definitions? Explore both collective (e.g. Fifteen Percent Pledge for retailers) and organization commitments. How can we make our definition transparent?

☐ **Humanize the Issue.** How might we create opportunities to personally interact with founders of diverse suppliers to hear their stories? This might include setting up formal and informal events.

Measuring Impact

☐ **Set Internal Accountability Goals.** What goals can we set to ensure accountability to our diverse supplier plan? How do different employee groups impact our organization's overall objectives?

☐ **Develop Tracking and Reporting Process.** What are the data points needed to track our diverse procurement program goals? Consider both inputs and outputs. How will we gather and report them? This should be shared both internally and externally.

OPERATIONS-LEVEL

Searching for Talent

☐ **Leverage Partnerships.** How might we leverage existing or new partners to help drive our diverse supplier initiative? Keep in mind that these partnerships take time and trust to develop.

☐ **Expand Network.** What ways can we expand (and diversify) our personal and professional network to assist diverse supplier hiring efforts? Consider attending minority-owned business events.

☐ **Communicate with Suppliers.** How can we expand our communication and marketing channels to reach more diverse suppliers? This might include non-traditional routes such as minority-based publications and community groups.

Onboarding Talent

☐ **Improve the Bidding Process.** How might we improve the bidding process to better align with diverse suppliers? What works well and what are the challenges?

☐ **Offer Coaching.** What kind of coaching might we offer to diverse suppliers to help prepare and onboard them? This is especially pertinent if the industry is highly regulated or has high infrastructure requirements.

☐ **Rethink Engagement.** What initiatives can we launch to keep our diverse suppliers better engaged? Examples might include regular touchpoints, offering frequent feedback, and going to visit their facilities.

Managing Talent

☐ **Add more resources.** What resources (financial, human, knowledge) are needed to support additional suppliers? These need to be added to the budget and are part of the commitment.

☐ **Create new norms.** How might we make diverse and inclusive procurement not just something that we do, but part of our corporate culture? Buy-in from top management is key. Other ideas might include humanizing the issue, incorporating ambassadors, and providing resources.

☐ **Consider existing suppliers.** How might our supplier diversity initiative affect existing suppliers, in both positive and negative ways? How do we manage this?

29. Title: A theoretical framework for Strategic Cognitive Analytics Management and Supply Chain Cognitive Capabilities

Mehrdad Jalali Sepehr
John B. and Lillian E. Neff College of Business and Innovation
University of Toledo
Toledo, Ohio, USA 43606
Email: Mehrdad.JalaliSepehr@rockets.utoledo.edu

Dr. Benjamin George
Assistant Professor at John B. and Lillian E. Neff College of Business and Innovation
University of Toledo
Toledo, Ohio, USA 43606
E-mail: benjamin.george3@utoledo.edu

Dr. Paul Hong
Distinguished Professor at John B. and Lillian E. Neff College of Business and Innovation
University of Toledo
Toledo, Ohio, USA 43606
E-mail: paul.hong@utoledo.edu

Abstract:

This study proposes Supply Chain Cognitive Capabilities framework. In this regard, according to the previously published empirical studies and their results, Technology Convergence TC, Supply Chain Analytics and Organizational Factors SCAOF, and Stakeholders-Analytics Interaction SAI have been proposed as the three pillars of the framework. The three parts of the brain and nervous system, mind and thought process, and action and motion have been defined as the qualities of the humanistic cognitive system by which our triple framework could reconcile Cognitive Analytics Management CAM concepts with SCM factors and emulates human expertise of rationalization and inference.

Keywords: Supply Chain Cognitive Capabilities theory, Strategic Cognitive Analytics Management, supply chain analytics, Stakeholder-Analytics Interaction, Technology Convergence.

1. Introduction

In the recent decade, the popularity of business intelligence (BI) solutions to enable business analytics has skyrocketed and customer service, human resources (HR), information technology (IT), legal support, marketing, procurement, production, project delivery, R&D, and sales can all benefit from commonly used business intelligence (BI) system solutions. (Pape, 2016), Business intelligence systems can be viewed as the technological underpinning for conducting business analytics (Lim et al., 2013), (Pape, 2016), (Chiang et al., 2012). In this regard, Business analytics is defined as a set of methodologies for converting raw data into action by creating insights for organizational decision making (Liberatore & Luo, 2010). The nature of analytics (descriptive, predictive, and prescriptive) and the focus of the SCM would be used to define supply chain analytics SCA in strategic and operational level. Big Data Business Analytics BDBA involves Big Data (BD) and business analytics (BA). BD is high-volume, high-velocity, high-variety dynamic data that exceeds typical data management methodologies (G. Wang et al., 2016). Big data business analytics (BDBA) has emerged as a vital business capacity to help organizations get a competitive advantage from vast amounts of data. Companies that can exploit big data well have a competitive advantage. Big data challenges compel organizations to develop new business strategies (Muhtaroglu et al., 2013). SCM faces issues that might cause waste and ineffectiveness in supply chains, such as delayed shipments, increased gasoline prices, unreliable providers, and escalating consumer demands. Companies anticipate BDBA to assist logistics and supply chain visibility, flexibility, and integration, mitigate demand unpredictability, and control cost changes It has also been demonstrated that BDBA may be used at both the strategic and operational levels of SCM (G. Wang et al., 2016), Furthermore, the position of BD as a value creator has resulted in higher levels of responsiveness, transparency, and accountability across governmental entities (Fosso Wamba et al., 2015).

"Cognitive Analytics" is a new emerging concept that refers to how organizations use analytics to make smart decisions; it is defined as a "field of analytics that tries to mimic the human brain in drawing inferences from existing data and patterns, it draws conclusions based on existing knowledge bases and then inserts this back into the knowledge base for future inferences a self-learning feedback loop." In other terms, it is characterized as the digital emulation of human thought processes. It entails self-learning systems that replicate the way the human brain operates by using data mining, pattern recognition, and natural language processing (*Analytics & Cognitive | Deloitte*, 2021), (Gillis, 2022). Although the majority of studies have attempted to define the technical definitions and practical purposes of various levels of analytics, few have taken a managerial perspective, and thus Cognitive Analytics Management CAM is one of the most novel topics that addresses the Cognitive Analytics CA from a managerial perspective. To elaborate, this approach seeks to understand how CA could have originated and what is the ideal way to use it in various businesses for better value creation.

The primary goal of this paper is to review the definitions and proposed frameworks of Cognitive Analytics Management CAM from several perspectives, to comprehend what CAM is, and whether we can come up with a solid and comprehensive definition. In this regard, CAM may be regarded as a revolutionary idea in analytics management. To review it, we investigated some related concepts, such as cognitive analytics and cognitive computing, big data analytics, and business analytics, and their similarities and differences based on past studies. Meanwhile, the implications, relationship, and importance of CAM with organizational aspects have been investigated, mainly focusing on Supply Chain organizations and Supply Chain Management.

As an outcome of the following literature review, we have concluded that there is a gap in defining CAM from the perspective of Strategic Supply Chain Management SCM perspective. Based on the extracted definitions, we have stepped in and modeled a more comprehensive conceptual framework titled Strategic Cognitive Analytics Management for Supply Chain Cognitive Capabilities theory (SCAM-SCCC), the items and interactions of which have been supported by previous empirical studies. The goal of the SCAM-SCCC is to clarify how CAM would be created, established, implemented, and managed in a supply chain organization to produce strategic values in the value chain setting.

The main objective of the SCAM-SCCC framework is to enhance the supply chain cognitive capabilities (SCCC) through aligning Technology, institutions, and Stakeholders. 1) Technology: that is technology convergence of digital and information technologies as well as other supporting technological advancements in other areas such as material science and advanced nano and micro-manufacturing, 2) Institution, which encompasses supply chain organizational factors and capabilities, namely, Big Data Analytics Capabilities, Data-driven Culture, innovation, and change management, knowledge management, safety, and information security, flexibility, agility, adaptability, resilience, the three features of its Dynamic Capabilities as sensing capability, seizing capability, reconfigurability capability, and trust management capability to enhance supply chain analytics capabilities and eventually value creation and operational performances, and 3) Stakeholders are government, customers, employees, society and community, etc. Hence, the concepts of Human-Machine Interaction HMI and Human-Computer Interaction HCI have been used for addressing the topic of interaction between stakeholders via analytics; thus, the term Stakeholder-Analytics Interaction SAI has emerged in this study as another novel topic, and the reason is that from the strategic perspective, some stakeholders like government, society, and community who are mass may be seen and behaved as a voice in the different kind of media and communication tools. Also, more importantly, the reinforcement cycle forms between the mentioned elements of the framework to have a more effective cognition of the business environment and decision-making process. The actual working process of a humanistic, cognitive system has been insightful for the conceptualization process of this paper. The three parts, brain and nervous system, mind and thought process, and action and motion, have been defined as the qualities in a humanistic, cognitive system by which our triple framework could reconcile CAM concepts with SCM factors; this feature has been explained more in the various parts of this study. In this regard, a literature review of the previous studies about CAM will be presented in the following.

2. Cognitive Analytics Management CAM Frameworks and its applications

In this section, we directly explore Cognitive Analytics Management (CAM) concepts and definitions. As one of the definitions (Osman et al., 2019),the CAM is characterized as a mission-driven strategy to transform companies, people, and systems to develop insights and create shared values. It generates data-driven insights to help make informed decisions and policy innovations for controlling performance and transforming e-services to achieve desired outcomes. Cognitive Analytics Management (CAM) unifies cognitive, analytics, and management ideas to support digital transformation in a company (Osman et al., 2019).

As the other definition for CAM we may refer to this definition from (Osman & Anouze, 2014) "*The scientific and integrative process for acquiring data and transforming them into applied insights to make informed decisions in real time using data analytics models, cognitive systems, and tools in a specific contextual society domain whether business, government, for-profit and not-for-profit organizations. They are inspired by human brain intelligence to make on-demand informed decisions anyplace, at any time.*"

(Osman & Anouze, 2014) developed Cognitive Analytics Management frameworks. According to their research, CAM employs five components of *Shared values, Analytics, Mission, Activities, and Structure* (SAMAS). From their study's view, Shared values from all stakeholders' perspectives and organizational structures are the most prominent components in the SAMAS framework (Osman & Anouze, 2014). Increasing shareholder value through communicating innovation to stakeholders while meeting environmental, social, and governance performance targets (Osman & Anouze, 2014). Although the notion of innovation will be described more in the following sections, here may be seen as a stepping stone to putting greater emphasis on innovation management inside businesses when our focus is the CAM framework. The values that an organization proposes to all stakeholders are also included in the concept of shared values. Using net physical and intellectual values, one might estimate the size of the impact they have. The organizational structure includes not only the traditional units of the organization, such as finance, marketing, operations, and human resource management, but also the innovation model and some other intangible parts and features of the organization that support the production or services of that organization (Osman & Anouze, 2014).

Also, concerning *Business analytics* which is a novel concept. INFORMS joined the business analytics movement to improve corporate decision-making. In this regard, they defined the BA as (Osman & Anouze, 2014): "*Analytics facilitates realization of business objectives through reporting of data to analyze trends, creating predictive models for forecasting and optimizing business processes for enhanced performance*". In (Osman & Anouze, 2014) the analytics part has been detailed for Data Envelopment Analysis (DEA) which may be defined as prescriptive analytics; however, their model may be expanded to general analytics too.

The 7S McKinsey model is the basis of SAMAS, which is a shared value management approach. The components in the 7S McKinsey framework are divided into hard, which consists of structure, strategy, and systems, and soft parts, which include skills, style, staff, and shared values. Using the stakeholder theory (Osman & Anouze, 2014) which tries to uncover societal demands from the viewpoint of stakeholders; the SAMAS framework aims to balance the interests of shareholders and other important stakeholders at the same time.

Moreover, as the support of SAMAS framework, (Osman & Anouze, 2014) presented the 8-Backbone Structure. The 3-I's chain - Invent, Improve, and Innovate are the central parts of this model. The link between invention and social needs goes back to an old saying *"the necessity is the mother of innovation".* Innovation is very essential for sustainable growth in shared values. invention creates something new. Improvement refers to an incremental innovation to enhance existing activities to make them work faster, better, cheaper, and easier or existing business models or processes or structures for a continuous development and enhancement. Innovation transforms an invention or an improvement through incubation into something that has a significant new impact on an organization or a market (Osman & Anouze, 2014)

In this regard, the first two rows of their model create a cycle in which needs result in invention, then improvement as incremental innovation leads to development, and finally innovation develop an invention or an improvement into something with a significant new impact on the enterprise or a market, these three parts are defined as which is the 3-I's chain Invent, Improve, and innovate. Finally, more creative businesses have more content stakeholders because of enhanced organizational performance. Overall, innovation is essential for long-term growth in shared values (Osman & Anouze, 2014).

As previously stated, the suggested frameworks are not only concerned with analytics skills, but also with institutional, technical, and stakeholder characteristics (Osman & Anouze, 2014), (Osman et al., 2019). However, according to prior articles, their major focus is still on cognitive analytics and cognitive computing; in this respect, Cognitive analytics uses machine learning and AI to analyze data. It handles unstructured and structured data, Cognitive analytics involves analyzing varied, complicated, heterogeneous, and qualitative data, (Handfield et al., 2019) , (Hurwitz et al., 2015), (Majhi et al., 2021) and from the synthesis of cognitive science, neurology, data science, and cloud computing, we reach to cognitive computing (Gupta et al., 2018). Cognitive computing simulates human mental processes in a digital model using data mining, pattern recognition, and natural language processing. Cognitive analytics is predicted to expand from $4.1bn in 2020 to $28.4bn by 2026 (37.2% CAGR), with the USA and China as leaders (Majhi et al., 2021). Cognitive analytics generate more insights than big data solutions and the main reasons are the capability of analyzing various types of real time data and also decision making like human's common sense (Majhi et al., 2021). According to SAMAS, technology is vital for CAM framework, but cognitive analytics technology is immature. There are few commercial platforms utilizing cognitive analytics, and cognitive analytics varies from data analytics (Handfield et al., 2019). Cognitive analytics investments provide low returns. Identifying how cognitive analytics might benefit organizations is crucial (Majhi et al., 2021). In this way, (Teece, 2007) defined the dynamic capabilities as an organization's dynamic sensing, seizing, and reconfiguring capabilities, and (Majhi et al., 2021) postulate how cognitive analytics technology might lead to these capabilities and then business value.

In this context, one of the other major and groundbreaking notions is big data, big data may be attributed to the mid-1990s, when John Mashey, working at Silicon Graphics in the United States, was active in the computation of massive datasets (Gupta et al., 2018) which has led to institutional insights in both strategic and operational organizational levels and is characterized by the 4-V data-models: *Volume, Velocity, Variety and Value* (Fosso Wamba et al., 2015), also some other studies consider some other Vs by adding veracity, volatility and visualization (Phillips-Wren et al., 2015), (Richey et al., 2016).

(Raghupathi & Raghupathi, 2014) have shown the sources of big data, which spans academic and scientific fields.

Cognitive computing may alleviate big data analytics' flaws Looking at the history of computing (Esser et al., 2013; Gupta et al., 2018), the Tabulating Era 1900s-1940s, the Programming Era 1950s-present, and the Cognitive Era 2011-present, in cognitive era the quantity of data will become unmanageable for individuals to comprehend, and so computer systems will need to aid people in decision making (Changchit & Chuchuen, 2016), (Gupta et al., 2018). A cognitive system can make interpretation of raw data and converting it into usable information and useful knowledge.

Cognitive analytics aims to eliminate numerous computations, developing, and testing the self-learning hypothesis by adapting the system to human brain structure, inferential reasoning and thinking. Cognitive systems should understand the company's key concerns. A cognitive system can learn and improve without reprogramming, and it process unstructured data and natural languages in a short time than big data. Cognitive computing is scalable, dynamic, and natural, these are some of the features that distinguish cognitive computing from big data analytics (Gupta et al., 2018). However, cognitive computing uses the output of big data as the entry point for its thinking processes, which are comparable to those of humans and include observation, interpretation, assessment, and decision making (Chen et al., 2016), (Gupta et al., 2018), (Ragini et al., 2018).

Big data and cognitive computing are complimentary analytical methodologies that may function together by using big data as a starting point. To elaborate, deploying a big data system while keeping the 5Vs in mind, followed by cognitive computing fueled by big data analytics outputs, would result in an organization's value creation, making it more competitive and efficient (Gupta et al., 2018).

The following is the relationship between the 5Vs and the four components of cognitive computing. *Observation* with *volume*, a huge amount of raw data enters the system and is stored; in this phase, the organization is collecting data using its cognitive capacities. *Interpretation* with *Variety*, having multiple forms of data such as location data, text, unstructured and structured data, would lead to better data *interpretation*, similar to using a person's various senses for better perception. Furthermore, *evaluation* with *velocity* and *veracity*, taking into account real-time data and the speed of data generation, as well as the accuracy of the following decisions and predictions that are dependent on the data, and finally a *decision* with *value*, which is the decision-making of a cognitive system that has been driven by the analyzed data via big data and would lead to knowledge creation (Gupta et al., 2018).

Focusing on cognitive computing and big data analytics will directly target the hard components of the 7S McKinsey model as Structure, Strategy & Systems or Analytics, Technology, and Insight (SAMAS). However, to get to the soft side of cognitive analytics management, we must first explore the underlying theories through which this notion was developed.

The basic theory for CAM is Social Cognitive Theory (SCT) (Albert Bandura, 1986). SCT is one of the most well-known human behavior theories. Although it has been used less frequently than other relevant theories like as technology acceptance, reasoned action, planned behavior, and innovation diffusion, it has been used in some aspects of analytics technology such as human–machine online interactions, which have been investigated by SCT. (Hollan et al., 2000) , (Osman et al., 2019), (Weerakkody et al., 2013), (Osman & Anouze, 2014), (Rana & Dwivedi, 2015).

SCT addresses human behavior that inspires humans to embrace information systems like e-government. SCT focuses on cognitive, indirect, self-reflective, and self-regulatory mechanisms in psychosocial functioning (Albert Bandura, 1986). Personal and social change, human thought, motivation, behavior, integration of innovation and social network, technology convergence and its effect on humans, processes governing societies, and SCT may be considered a comprehensive theory that can be applied to various types of humans or society and technology contexts are all addressed by Social Cognitive Theory. When analyzing the topics that SCT has studied, one of the most crucial issues that, to the best of the authors' knowledge, has not been considered in any previous research pertinent to CAM is technological convergence. As a result, it is one of the four components of our suggested SCAM-SCCC structure, following the cases, we can observe that there has been technological convergence; however, the CAM research does not provide any specific evidence of this occurrence.

Convergence was once focused on digital technologies. However, technology is expected to work as a catalyst in the future to improve creative convergence (Park, 2017). Low global growth has persisted for many years. Many experts agree that future technologies (e.g., Internet of Things, big data, and AI) will address the global recession through openness, convergence, and new market demand. Nowadays, the competitive advantage would be called the economy of convergence. Furthermore, there are case studies involving Nike and Walmart and their business analytics adoption approaches; yet, while reaching economic prosperity, not all levels of business analytics adoption have achieved the degree of technology convergence (Park, 2017).

3. *Theory Development for Strategic Cognitive Analytics Management SCAM and Supply Chain Cognitive Capabilities SCCC*

The primary goal of the SCAM-SCCC framework is to introduce and enhance Supply Chain Cognitive Capabilities (SCCC) by aligning 1) Information Technology: technology convergence, 2) Institution: organizational supply chain factors such as organizational knowledge management KM, organizational dynamic capabilities, and so on. 3) Stakeholders: integrating stakeholders and analytics to create value 4) Learning and Improving: Reinforcement cycle and interaction between IT, Institution, and Stakeholders to make the optimal decision.

3.1. Technology Convergence and Supply Chain Analytics

The notions of Industry 4.0 and digital convergence and transformation, as well as their effects on supply chain, become bold when discussing technology convergence and supply chain analytics. In this regard, we will first go through these topics. However, there are subtle differences between them, hence, in this study, cognitive analytics technologies are viewed as the context for technology convergence, and its reasoning is described below. As theory building for technology convergence, we discussed the cognitive convergence of organizations by considering the Triple Dynamic Capabilities model of (Teece, 2007) as our base for theory developing in this section which includes sensing, seizing and reconfiguration and how it is aligned with the cognitive technology convergence (Majhi et al., 2021). In this regard, DC has been determined as an essential capability for competitiveness (Teece, 2007), and it has been supported that cognitive analytics technology would improve the DC of firms (Maymir-Ducharme et al., 2014), and also some firms have been developed their knowledge management system based on cognitive analytics which would lead to value creation in their organizations and also as an forerunner for DC, hence we perceive the reinforcement cycle of CA, KM, DC, and value.

Although several studies have addressed cognitive analytics, cognitive computing, and CAM, most of them did not directly refer to cognitive analytics as a technology. In this context, only (Majhi et al., 2021) examined cognitive analytics technologies that would contribute to value generation for enterprises by taking the Dynamics Capabilities DC into account. In their research, they investigated Dynamic Capabilities as *Sensing Capability*, *Seizing Capability*, and *Reconfiguring Capability* (Majhi et al., 2021). Their definition from DC is based on the theory of (Teece, 2007) for DC, the conceived triple model's strategic vision, and its theory is to sense and channel opportunities and threats, then seize the sensed chances for appropriate decision making, and finally reconfigure existing tangible and intangible assets to preserve competitiveness. Based on prior tests, we can observe that there will be cognitive convergence in this process; it appears to be a person's cognitive system, which leads to greater decision-making, improvement, and progress by considering the dynamics of life and learning experiences and that is the main reason we approach the technology convergence as cognitive convergence in the dynamic environment. Sensing and seizing capabilities are anticipated to assist businesses in sustaining and prospering in the context of global transformation (Conboy et al., 2020).

Sensing capabilities assist organizations in better understanding their environment by gaining consistent and reliable strategic information (Kump et al., 2019). Cognitive analytics may assist us in embedding the sensing capabilities in our firm by detecting the external world and dealing with uncertainty and ambiguity (Gudivada, 2016), and so it is also consistent with our proposed framework through which we intend to deal with SCM through CAM. Furthermore, CA improves sensing capacities by allowing businesses to overcome obstacles and gain tacit information. The Defense Advanced Research Projects Agency (DARPA) Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE) project or The DARPA SyNAPSE system, and IBM Watson DeepQA, are two well-known examples of CA technologies that use different types of sensors to grasp data by merging contextual analytics to help emulate human expertise of rationalization and inference. These two technologies are not analytics approaches because they are technologies with their own strategic objectives as part of their development, of which analytics is merely one component. From the perspective of analytics, the sensing almost happens by descriptive quality of CA to discover the opportunities in the environment (van Rijmenam et al., 2019), (Majhi et al., 2021).

Seizing capability is concerned with how a company engages in measures relevant to producing opportunities based on its strengths and disadvantages. These processes include strategic decisions about stakeholders, supply networks, and even decision-making techniques (Conboy et al., 2020; Kump et al., 2019), (Conboy et al., 2020), (Majhi et al., 2021). Reconfiguring capabilities of firms also assists them to be more flexible facing the changing environment, and CA contributes to this capability. Hence (Majhi et al., 2021) created a conceptual model by analyzing three hypotheses that show Cognitive Analytics increases organizations' sensing, seizing, and reconfiguring capabilities. The issue of organizational reconfiguration is of great consequence in the above model which has been supported in some other studies too.

At the Davos Forum 2016, the president of the World Economic Forum, Klaus Schwab, elaborated on several concepts of technological convergence, stating that the Fourth Industrial Revolution is a *disruptive innovation* since our models and ideals are being replaced with radically new ones (Roland, 2022). In the third industrial revolution, humans were tied to computers and machines, however in the fourth industrial revolution, computers and machines communicate proactively and autonomously via AI, with increased smartness involved (Park, 2017). The convergence of digital, physical, and biological technology has been identified as a hallmark of a dynamic economy that will lead to a convergence economy (Park, 2017). Furthermore, claims made in a Washington Post article regarding Google's LaMDA project and Blake Lemoine, a Google engineer who believes AI has progressed to the point where it looks to have personality and life (Tiku, 2022).

The convergence of digital, physical, and biological technologies has been identified as a feature that will lead to a dynamic economy which will lead to a convergence economy. Furthermore, statements made by the Washington Post article about Google's LaMDA project and Blake Lemoine, a Google engineer who believes the level of AI has been reached to the point where AI appears to have spirit and life, The level of self-consciousness can be used to support the notion that there is interactive communication between computers and machines in industry four via AI. According to Blake Lemoine's statement, he had a conversation with LaMDA, Google's artificially intelligent chatbot generator, and it felt like he was speaking to a child with thoughts, sentiments, and human-like senses such as comprehending loneliness and the yearning for a relationship. Following LaMDA's chat, we can observe that the transformational machine learning algorithm utilized for this bot has enabled it to construct human-like memory that profoundly perceives and saves the trend of the preceding environment.

We can also talk about wearable and additional technology that can help disabled people with their cognition. One example of such technology is Neil Harbisson, who is regarded as the first Cyborg (Harbisson, 2012). Neil Harbisson does not see colors, so technology has helped him to hear colors, as he said "*I don't feel like I'm using technology, I don't feel like I'm wearing technology, I feel that I am technology*." The reason for this is that he does not wear it; rather, an antenna has been implanted in his head to assist him in listening to colors and creating color perceptions in his mind; there is also a TED speech from him for more information. Anyway, this may be considered as a convergence which has led to the better cognition. It is real time, it uses electronics and algorithms (analytics) to transform color or vision to sound, and it is the convergence of digital, physical, and biological technology, as previously said. Our concern here is if, as a result of this trend in AI technology, our business analytics solutions will change in the future. Would it use new cognitive technology capabilities for its purposes? The Neil Harbisson case was Science Fiction a few years ago, but it has now become reality, as the AI chatbot LaMDA has come to true. Certainly, both were produced by the skill of scientists, but the convergence of distinct qualities and the creation of harmony between them is the key that has led to the production of value. Can we anticipate Cyborgs in the business sectors in the future? For example, stock experts can use wearable technologies to convert what they see on the market trend display into more advanced useful knowledge in their minds.

Convergence occurs when two or more diverse things come together to accelerate technological innovation, and they are all placed in a coherent system and device, as defined by Ray Kurzweil in his book The Singularity is Near. Smart phones are a common example of technological convergence because only in one device we have camera, music and video player, telephone, digital language translators, and many other separate features which come together to build a singular applicable device. The two principles for digital technology convergence have been identified as *integration* and *digitalization* (Menon, 2007). In addition, integrated technology can be defined as technologies that operate through independent digital systems that are placed next to one other, and digitalization is the method for creating a bridge using data to connect the integrated pieces as a whole system.

The internet of things, IOT, and block chain are the most frequent forms of technology convergence that have been used in the supply chain. Each was established by the convergence of Information Technology IT (physical devices) and Operational Technology OT (monitoring devices). Another way to look at it is the convergence of the three Cs: computers, content, and communication. However, M.C. Roco and W.S. Bainbridge define technology convergence as "converging technologies for boosting human performance," which is not limited to digital technology. As a result, for that goal, they considered the convergence of Nanotechnology, biotechnology, information technology, and cognitive science (NBIC). The NBIC may address both micro and macro scale problems in a society. For further exploration we may refer to this sentence from Christy Roland at AT&T (Roland, 2022) (The complete and modern guide to technology convergence, 2022) *"This new discipline describes the intersection of the four scientific areas as the cognitive scientists thinking it, the nanotechnology scientists building it, the biotechnology scientists implementing it, and the information technology professionals monitoring and controlling it. The "it" in this scenario describes futuristic scientific developments such as direct brain-to-brain communication, the evolution of our brain capabilities, and the ability to regulate our body functions far beyond their current capacities."*

When we look at the literature on technology convergence in the topic of digitalization, we nearly always come across technologies like AI, blockchain, and the Internet of Things (IoT). Meanwhile, AI has been defined as the amount to which a machine can perceive its surroundings, and it is mostly determined by the Turing Test. Blockchain is a solid distributed database that evolves as blocks, and IoT is when many physical devices are linked to each other by electronic implants to share data. Technology convergence has a great impact on economy as it creates the economy of convergence and in the USA, the government has considered the convergence of Nanotechnology, biotechnology, information technology, and cognitive science (NBIC) as the zenith to improve the individuals' life, and the pioneer companies in the field of technology convergence are Apple, Adobe, Netflix, Amazon prime, IBM, Microsoft, and Google. Also, in European countries the Converging Technologies for the European Knowledge Society (CTEKS) has been determined as the technology convergence apex which is mostly focused on social welfare. One of the pioneer researchers in this field is Dr. Maggie Kosal at Georgia Institute of Technology, has mentioned that converging technologies would be the future and the most suitable places to find new opportunities. Technology convergence is a broad subject as the NBIC that contains a broad range of technologies (Roland, 2022).

According to COL T. Christopher Petty, a senior researcher who recently completed researching technology convergence for the United States Army (Roland, 2022) , the emphasis has been placed on nanotechnology for further advancement in other fields such as artificial intelligence (AI), robotics, additive manufacturing, biotechnology, and even quantum computing. In this regard, we see that the breakthrough improvement in one field, which may be completely related to Cognitive Analytics like AI, is dependent on the improvement in another technology (nanotechnology), which would not have been seen obviously, while it is an underlying component, and this is the primary reason why, in this paper, we have not considered digitalization and or digital convergence alone. Furthermore, we may support it by considering the emergence of additive or wearing technologies, as well as Google's LaMDA project, they are all developed based on software and hardware, if we only consider the digital parts, our focus would be inclined towards software and also lower than that which is related to algorithms, but on the other side which has not been seen in the researches is the hardware components and their capabilities that have a prominent influence. This would be assisted by the impact of Micro and Nano Electromechanical systems (MEMS and NEMS) on electronic devices that might be utilized as business analytics equipment, or processors that could be employed in processors to implement more powerful AI algorithms. To present a comprehensive conceptual model, one of the most important factors for CAM is technology convergence, which behaves like a chain made up of various interconnected technologies and ultimately leads to a better environmental perception, and as previous technologies such as IoT and Blockchain have been used in business operations and supply chain. The future technology advancements also may be used in the business and supply chain management too, in this regard the study of (Xu et al., 2017) has conducted technical research on improving the resilience of MEMS against shocks to improve the microswitches which would be used in IoT systems, also (Sharma et al., 2021) conducted technical research about improving the design and functionality of the Radio frequency-based Micro-Electro-Mechanical System (RF-MEMS)-based devices (switches) which is used in IoT and 5G for their better performances, and also for this improvement in that type of switches they designed it by square plate of polysilicon with silicon nitride as dielectric which has led to the better voltage passing. Their study, as well as some other papers in the field of electronics and material science, confirm our claim that advancements in material science have influenced electronic devices utilized in IoT, Blockchain, 5G, and other current and future digitalization. Furthermore (Le et al., 2021) has conducted research about how various types of Piezoelectric microelectromechanical system (MEMS) devices based on their designing capabilities have been improved the 5G and IoT technologies. As previously stated, our attention is on the true meaning of technological convergence and its progression from micro to macro and more substantial applications for management goals. In this regard, one of the tangible applications of technology convergence is the integration of IoT-based systems with Improved Decision Machine Learning Algorithms (IDMLA) and its application in aquaculture for environmental management purposes. This technology integration has resulted in the reduction of time, measurement errors, and losses in sensing important real-time factors in aquaculture such as water quality and velocity (Manoharan et al., 2020). Despite the fact that the results have been directly related to environmental management and indirectly to food production and (fish) food supply chain management, there have been additional studies that support the usage of IoT, Blockchain, and digital convergence in supply chain applications, which will be discussed further below. Prior studies, however, have focused on digitalization and the technology associated with it, and their framework has not taken into account the supporting and underlying science, as well as the convergence of many types of technologies, that lie behind digitalization. To the best of the authors' knowledge, managerial research did not investigate the impact of convergence in electronics and material science on the tools that would be used in digital supply chain management, but it has been indirectly proven by technical papers written by engineers and technology scientists, as mentioned above.

Michael Shearer, SelectHub's Director of Marketing Operations (Roland, 2022) , noted the convergence of marketing automation and Customer Relationship Management for more successful sales and marketing in corporate processes. By going further, (Dubey et al., 2021) studied how artificial intelligence-based supply chain analytics could improve the performance of industrial relationships of organizations during the COVID 19 pandemic. In this regard, they developed their theory by considering dynamics capability view Contingency theory (CT), and alliance management capability. They referred to (Araz et al., 2020) by writing that "*Thanks to the explosive expansion and advances of digital technologies, such as smart mobile phones, social media platforms, e-commerce, and so on, data are around in every organization. As the analytics capabilities of organizations develop rapidly, artificial intelligence tools, big data analytics, blockchain, and so on are all tools available and being used in the industry"*. Considering the earlier statements, and referring to (Araz et al., 2020) and (Dubey et al., 2021), the convergence of digital technologies have been mentioned by considering this statement of (Dubey et al., 2021),"*With the rapid proliferation of the internet, smartphones, and other emerging technologies (RFID, sensors, Internet of Things, Cloud Computing, etc.), we have reached a new phase where large volumes of data are collected in real-time in structured, semi-structured and unstructured formats*", which has been supported by (Fisher et al., 2012), (Agarwal & Dhar, 2014) too, and by referring to the words "*and so on are all tools*" and "etc." We can see that this barrier may not be identified precisely owing to a lack of understanding about future surprise innovations that may develop in the field of business analytics, but the common aspect in both comments is the demand for the convergence of diverse types of digital technology. They coupled dynamic capabilities with evolving technologies in their study, and they discovered that establishing the boundaries of dynamic capabilities' influence on firm performance is a difficult process, while some capabilities do not operate well (Dubey et al., 2021).

Industry 4.0 could have been created by the technology convergence. The phrase Industry 4.0 arose as the result of the integration of IoT into production and manufacturing, and it has contributed to the supply chain by taking into account the global network and making them smarter and more transparent (Tjahjono et al., 2017). "*A smart factory is referred to as the use of new innovative developments in digital technology including advanced robotics and artificial intelligence, hi-tech sensors, cloud computing, the Internet of Things, data capture and analytics, digital fabrication (including 3D printing), software-as-a-service and other new marketing models, mobile devices, platforms that use algorithms to direct motor vehicles (including navigation tools, ride-sharing apps, delivery and ride services, and autonomous vehicles), and the embedding of all these elements in an interoperable global value chain, shared by many companies from many countries*", (Tjahjono et al., 2017), also as we read in the above quote, again we face the technology convergence issue, it calls different digital technologies to converge to form the industry 4.0 and a smart factory, also they explores the application of each of the technologies such as virtual and augmented reality, additive manufacturing, simulation, Big data analytics, cloud technology, cybersecurity, IoT, miniaturization of electronics, automatic identification and data collection AIDC, radio-frequency identification (RFID), robotics, drones and nanotechnology, Machine-to-machine communication M2M, business intelligence BI, on the logistics, warehouse, and supply chain (Tjahjono et al., 2017). Considering the miniaturization of electronics may lead to considering MEMS and NEMS technologies as well. By the way, the mentioned technologies create a portfolio of technologies that have been converged for a singular purpose, which is to make a factory smarter, or to strengthen its ability to perceive its environment and cognitive capabilities. (Fatorachian & Kazemi, 2020) employed System Theory to examine the relevance of some of the technologies which created the industry 4.0 and then its impact on supply chain performance. They also included M2M, Cyber-Physical Systems (CPSs), IoT, AI and Big Data Analytics, and cloud computing as the technologies that could improve connectivity, integration, visibility, transparency, as well as responsiveness, flexibility, and productivity in the supply chain. To go depth into the opportunities and also challenges of industry 4.0 on supply chain.

In the following sections, we will discuss the opportunities and challenges that technological convergence, particularly in the field of digital technologies, would bring to supply chain management, to further support our theory that technology convergence is a component of SCAM-SCCC, we investigate how the development and convergence of some of the other digital technologies would contribute to the development of the other. This is a model which has been supported by (Gill et al., 2019), actually they explored how the of triumvirate "IoT + AI + Blockchain" and also emerging advanced technologies are going to affect could computing within the future as they called it could futurology. On the other hand, they indirectly support the idea that the convergence of digitalization has transformed and evolved computing throughout history, beginning with Client Server in 1960 and progressing to IoT and Edge computing in 2013. By delving deeper into their Triumvirate conceptual model, we will see how cloud computing emerged as a result of the convergence of Internet of Things IoT, Artificial Intelligence AI, and Blockchain, while also keeping in mind that each of the components is an overarching digital concept formed by the convergence of some other components (Gill et al., 2019), also there have been some other studies which explored the integration of IoT and Blockchain (Reyna et al., 2018), AI and Blockchain (Dinh & Thai, 2018), (Ekramifard et al., 2020), IoT and AI (Kankanhalli et al., 2019), which explored the technical features of their integration also their application in various sectors such as energy, healthcare, agriculture, food industry, pharmacy, and education (Alshehri & Muhammad, 2021), (Chamola et al., 2020), (Misra et al., 2022), (Ciolacu et al., 2019), (Kumar et al., 2020).

During the COVID 19 epidemic, Supply Chain Resilience SCR became a very important topic for both researchers and industrialists, and the integration of digital technologies and AI could have a significant positive impact on long-term SCR performance and its adaptive capabilities (Belhadi et al., 2021). Furthermore, it has been demonstrated that facilitating conditions such as a person's belief that an organizational and technical infrastructure supports the system, trust, and social influence are the main factors that could influence the behavioral intention of organizations to adopt blockchain technology, according to the theory of the Unified Theory of Acceptance and Use of Technology (UTAUT) model (Queiroz et al., 2020), As a result, Blockchain has been identified as a disruptive technology that has the potential to modify an organization's technological architecture in order to get more value from its deployment. The influence of blockchain convergence with industrial revolution (IE 4.0) technologies and its effective deployment in several fields such as healthcare and supply chain management have been addressed from this perspective (Chand Bhatt et al., 2021). To go deeper, we can observe that the impact of technology convergence is advantageous not only in the broad chain of a supply chain, but also in the specialized sector of manufacturing, which is one of the most significant components of the SC. In this regard, robotics has been identified as one of the industrial sectors where technological convergence could be beneficial. In other words, we can see that robotics was created by the convergence of various specific technologies in the mechanical, electrical, automation and control, computer, and even biological technologies for some of them that mimic the motion and behavior of various creatures, and that this convergence could also work efficiently, especially in the healthcare industry, which has advanced dramatically (Kose & Sakata, 2019), in addition to the direct constructive impact of digital technology convergence on supply chain management such as improving the value chain, transparency, quality control and management, secure supply chain transactions, resilience and managing the ripple effect during crisis, and trust (Rejeb et al., 2019),(Sariyer et al., 2021) (Hasan et al., 2022), (Ivanov et al., 2018), (Fatorachian & Kazemi, 2020).

Another key aspect of technological convergence is that it stimulates technical innovation and technology opportunity discovery, organizational core competitiveness (Tang et al., 2020), so creating a reinforcement cycle, (Choi et al., 2022), (J. Lee & Sohn, 2021) in other words, it assists the technological continuous improvement within the organizations, the prominent factors which plays an important role in this cycle are the *knowledge convergence* (Duan & Guan, 2021), *R&D human resource planning* (J. W. Lee & Sohn, 2021), *technological convergence network (TCN)* (Luan et al., 2021), which have reinforced the improvement cycle.

The other point is that digital convergence may lead to the industrial convergence within the supply chain. many aspects of supply chain convergence so that businesses can effectively capitalize on economic opportunities given by the growing development of digital convergence and related technical breakthroughs, this feature of technology convergence has been emerged and boosted by virtual organizations and their dependence on effective information systems(Chandrashekar & Schary, 1999), (Rim et al., 2007), (Karvonen et al., 2012).

3.2. Supply Chain Analytics and Organizational Factors

So far, we've talked about the importance of technology convergence as one of the major components of our framework, which will play a significant role in Cognitive Analytics Management. If we consider a supply chain organization to be a human, the most tangible definition of the organization's cognitive abilities is technological convergence. In other words, a human would utilize his or her senses of sight, hearing, smell, taste, and touch to detect his or her surroundings, which is necessary for survival and growth, and all inputs from these senses would be processed in the brain, which forms the mind. The combination of these characteristics would result in improved perception. While there is a subtle difference between the mind and the brain. In other words, the brain is a physical entity with blood and nerve cells that can be touched, whereas the mind is just mental and has no set position in our body. The brain is the nerve system's control and coordination center, managing and coordinating our motions, thoughts, and feelings, but our mind is in charge of thinking processing and comprehension, which forms our conscience. In this view, we can hypothesize that technological convergence is the union of numerous nerves in addition to their processor, the brain.

In this section, we'll delve deeper into the mind component of CAM. The emphasis in this regard would be on organizational and business analytics capabilities that lead to effective transformation of sensed and processed information to value and assets. Assets can be both tangible and intangible. We would like to consider some organizational capabilities as intellectual assets that would evolve because of more growth and development of effective use of analytics. This is like how mind abilities would develop because of more training, doing some practices, and gaining experience over time. As we explained above Dynamic Capabilities DC consists of *sensing*, *seizing*, and *reconfigurability capabilities*, and it has been postulated that cognitive analytics enhances these three organizational capabilities (Majhi et al., 2021), however, the other point is that these capabilities exist before the cognitive analytics initiate to enhance them, and after that there would also be a reinforced cycle which boosts the organization's analytics capability by their enhanced dynamic capabilities, so there would be a reciprocal relationship for this field, and it is not only for DC, but also it works for organizational innovation and change management capabilities (Duan et al., 2020), (Božič & Dimovski, 2019), (Lozada et al., 2019) ,(Chatterjee et al., 2021)knowledge management capabilities (Chung et al., 2020), (Y. Wang & Byrd, 2017), (Khan & Vorley, 2017a), (Enad Al-Qaralleh & Atan, 2022), data and business security (Z. M. Bi et al., 2021), (Z. Wang et al., 2020), (Mahmood & Afzal, 2013)Flexibility, reconfigurability, agility, adaptability, and resilience (Z. Bi et al., 2021), (Napoleone et al., 2022), (Dubey, Gunasekaran, & Childe, 2019), (Dubey, Gunasekaran, Childe, Fosso Wamba, et al., 2019), (Yu et al., 2021)organizational culture and trust management (Dubey, Gunasekaran, Childe, Roubaud, et al., 2019).

A structured literature assessment based on the mutual relationship between analytics capabilities and supply chain organizations is shown in the table 1 below. Because our proposed conceptual model and theory development are new, this literature review is limited to the supply chain analytics section, which supports one of the components of our study. The goal of this literature review is to support our conceptual framework and provide an overview of previous works. It should be noted that all of the "RQ and hypotheses" in the Research Questions and Hypotheses column are the same as they are stated in the original papers.

In the context of hospital supply chain, Big Data Analytics Capability BDAC has positively influenced inter-functional integration, hospital-patient integration, and hospital-supplier integration, meanwhile hospital-patient integration and hospital-supplier integration completely mediate the relationship between inter-functional integration and operational flexibility (Yu et al., 2021), or in other words all may be stated as, BDAC influences organization-stakeholder integration and organization-stakeholder integration has leaded to operational flexibility. (Srinivasan & Swink, 2018) explored how Demand Visibility and Supply Visibility would result in more effective Supply Chain Analytics Capability SCAC, and SCAC would positively impact operational performance as cost and product delivery, meanwhile organizational flexibility would be moderator in this relationship between the SCAC and operational performance. The common point between (Yu et al., 2021) and (Srinivasan & Swink, 2018) is the fact that (Yu et al., 2021) states the goal as the improvement of flexibility by analytics, while (Srinivasan & Swink, 2018), (Dubey, Gunasekaran, & Childe, 2019) indicate how organizational flexibility moderates the relationship between analytics and operations performance. Here we may point to our statement about the mutual relationship between analytics capabilities and organizational capabilities. Furthermore, Big Data and Predictive Analytics BDPA (Mishra et al., 2019), could also work as the mediator between the information technology deployment and human resources capabilities and organizational performance. According to the digital convergence, IT and BDPA would be converged as a part of the CAM, that would result in better organizational performance. Like flexibility, reconfigurability is also one of the organizational capabilities which plays a crucial role for manufacturing companies, and this capability also could be supported by technologies of cyber-physical systems (CPS) as one of the industry 4.0 technologies, which may be converged for CAM (Napoleone et al., 2022).

From the perspective of organizational culture and trust management, considering analytics capability as one of the dimensions of organizational culture, and swift trust as a critical aspect of for designing a supply chain to manage the ripple effects during the crisis (Dubey, Gunasekaran, Childe, Roubaud, et al., 2019), big data analytics impacts swift trust and the collaborative organizational performance. In this way, flexibility positively moderates the relationship between analytics and collaborative performance too.

As discussed in the previous section, cyber-physical systems and interaction of humans and machines are important features for industry 4.0 applications in collaborative manufacturing, in this regard, safety and security assurance in this area is of great importance, concerning this there is an approved view of how to assure security and safety while working under one of the masterpieces of technology convergence as Industry 4.0 (Z. M. Bi et al., 2021) which has been embedded in CAM, also business analytics could support the data security in cloud computing (Z. Wang et al., 2020), thus two parts of CAM which has been formed by convergence of analytics and cloud computing may support one each other's security. and there is another view which supports how to use big data analytics as another component of CAM to assure safety and security in cybersecurity (Mahmood & Afzal, 2013), (Cardenas et al., 2013), (Verma et al., 2015)operations in energy sector or assist to assure energy security (Hu & Vasilakos, 2016), and occupational health and safety (Ezerins et al., 2022). To delve into the relationship of data analytics and safety from which the term *Safety Analytics* has been emerged, and the proposed framework by (Ezerins et al., 2022), that indicates *quality of data*, *organizational standards for data collection and classification, technological infrastructure and platforms and adroitness in analytical and safety skills, and measurement culture* plays an important role in safety analytics aspect of organizations to predict and take measure against potential hazards. In conclusion analytics may assist organizations in their cybersecurity, operational performance and occupational health and safety, also one analytical component like business analytics can support the security in another one like cloud computing.

Business intelligence and analytics (BI&A) are recognized for their ability to foster innovation and deliver organizational value. Also, for the mutual impact of analytics on innovation and change management, Business Analytics directly impacts the Data-Driven Culture DDC, and Environmental Scanning ES (absorptive Capacity) which helps organizations to perceive what is happening for their business sector and their stakeholders, and DDC and ES positively affects new product development and innovation which ultimately leads to competitive advantage (Duan et al., 2020). Considering exploitative innovations as improving the current products and services and exploratory innovations as dramatic changes, innovation ambidexterity is establishing a balance between them for better value creation, concerning this, (Božič & Dimovski, 2019) proved that Business Intelligence and Analytics (BI&A) positively impacts innovation ambidexterity and Absorptive Capacity, moreover Absorptive Capacity impacts innovation ambidexterity that eventually leads to organizational performance. It should be noted that Absorptive Capacity is a company's capacity to perceive the value of incoming, external knowledge, integrate it, and apply it to commercial purposes (Božič & Dimovski, 2019). Furthermore, organizational Data-Driven Culture DDC impacts the organization’s process and product innovation, and adoption of business analytics positively affects DDC, and Environmental Scanning ES. DDC also works as a moderator between the ES and innovation that ultimately lead to better organizational performance and competitive advantage (Chatterjee et al., 2021).

Also, from the perspective of knowledge creation and knowledge management, (Khan & Vorley, 2017b) conducted a comprehensive literature review study to support that big data text analytics is an enabler for knowledge management. However, data for its own by using analytics may not solve business problems, hence a mutual interaction and relationship is needed for this approach, and knowledge management is the factor which forms the context and domain for big data analytics (Edwards & Taborda, 2016) and considering this importance is vital for designing innovative information technology platforms for analytics and knowledge management (Rabhi et al., 2021), as (Chung et al., 2020) developed a social media-based business analytics system to support intelligent analyses of social media content for knowledge creation by Twitter content.

From the perspective of cognitive analytics technologies and triple dynamic capabilities (sensing, seizing, reconfigurability) of (Majhi et al., 2021), (Fosso Wamba et al., 2020) studied how firm’s capabilities to discover opportunities (sensing capabilities), and their organizational data driven culture is crucial for Big Data Analytics, in this regard they have found that BDA-enabled sensing capability and analytics culture has a positive impact on organizational outcomes.

3.3. stakeholder-analytics interaction

As the two main components of the proposed framework, Technology Convergence TC and Supply Chain Analytics and Organizational Factors SCAOF have been explained. TC is assumed to be the corresponding cognitive component as brain and nervous system in which the various senses would work to gather information and transit it to the brain for processing. SCAOF has also been characterized as the mental component of a cognitive system, which oversees thinking processing, learning from experience, communication, and some other skills learned through the mutual interplay of mind and brain. In this manner, a corporation can collect and store a massive quantity of data using TC such as Industry 4.0, IoT, and cloud computing, and then process it using business analytics approaches. For both processes, it is critical to analyze the architecture of the information and analytics system; this architecture is how the organization's mind has been formed, and it is tied to both hardware (brain and nerves) and software (mind), as well as TC and SCAOF.

Meanwhile, technology convergence helps to capture more accurate real-time data, such as the brain and nerve system; nevertheless, the architecture and organization of brain and nerve cells have been formed in advance and at the start of mind creation, although for organizations it is a little different. Whether a company wants to enjoy analytics and IT is determined by the decisions of its stakeholders, particularly management and top-level decision makers; thus, the hierarchy or process that is determined for human cognitive systems does not quite exist here. Most humans have Determinism about their cognition, whereas organizations have Free Will, but the unifying tie is that both require it to thrive and govern their environment. As a result, the decision to implement and enjoy business analytics and IT is dependent on organizational strategies, culture, dynamic capabilities, and successful and effective use of analytics. As previously discussed, these organizational factors may act as mediators in this process, and there is a reinforcement cycle between organizational factors and analytics. Furthermore, one of the other features that we would like to discuss here is how stakeholders interact with analytics during this process and decision making.

To elaborate, the various forms of technologies and their development help businesses better gather and store information. As a result, detecting bottlenecks, key performance indicators of the process, and some other critical points that should be monitored is critical to putting the sensors there, and these points would be identified by the company's experts and the available historical data and observations before implementing any advanced technology. Furthermore, after the successful adoption of IT systems, the process and its outcomes will be monitored by professionals. As a result, we would always anticipate interaction between specialists and analytics in CAM. However, this interaction can occur at various levels of an organization's structure and can be both external and internal. However, it is evident that all stakeholders are involved, both directly and indirectly. Earlier research has focused on human-computer interaction; however, for institutional theory and stakeholder perspective, the crucial issue is how stakeholders who are also humans, governments, and non-governmental organizations (NGOs) interact with analytics and IT. As a result, the third component of our paradigm was described as Stakeholder-Analytics Interaction SAI. The perspective from which the stakeholder-analytics interaction was examined differs from that of the human-analytics interaction. Because stakeholders can be governments, other organizations and entities, suppliers, internal and external customers, consumers, and community we should consider it as a broader issue than a single human, and there are no similar results in Google scholar for the term "*Stakeholder-Analytics Interaction,*" which was coined and developed in this study.

Stakeholder theory is one of the most important theories in management (Clarkson, 1995) (Freeman, 2015) , its concentration is to consider not only the shareholders of the organization but also all other entities who may be directly and indirectly influenced by the organization's performance (Clarkson, 1995). And fairness among the stakeholder system is of great importance, because if one of the stakeholders felt it has been treated unfairly would withdraw from the stakeholder system of that organization and would go in the stakeholder system if the other organization and it has adverse effect on organizations performance (Clarkson, 1995). Stakeholders are also essential for value creation (Tantalo & Priem, 2016), and organization and stakeholder interaction, engagement, involvement, or mutual communication is one of the issues in this theory (Greenwood, 2007), (Kuhn, 2008). In the past firm's interaction was between some of the individuals as human-human interaction, however, todays it would be through Internet-based technology medium and social networking and we see the technology-enabled interactions (Bhattacharyya, 2020). For example, companies send their surveys via emails, Short Messaging Services (SMSs) to their customers to get data about their customer experiences and after that by analyzing that survey and other sorts of data such as implemented changes and improvements in the sales processes, would turn the data and information to more valuable knowledge about their sales and marketing management (de Bussy et al., 2001). Moreover, many of the interactions would be conducted as the rates, feedbacks, comments on the official website, social media channels like YouTube, Twitter, etc. of the companies (Ruehl & Ingenhoff, 2015), and sometimes the communication may be indirectly related to the company's product, and it may be related to companies' policy. For instance, an oil company decided to destroy a huge part of jungle for piping, in this regard, there would be a campaign on the internet to protest this issue. Here it seems to be a conflict of interests between stakeholders. One side who are opponent of financial performance of that operation would communicate their feedbacks and comments about the benefits of their project, and on the other side environmentalists and those who would lose their interests by this project who supposed to communicate about the side effects of this project. Here we hear two voices, as two types of stakeholders, and the overall sentiment, proofs, evidence that each side of the conflict would present would be considered as the information and their opinions. Even, we talk about customers, they would be treated based on the segmentation and classification coding that the firm uses, not individual persons. In this regard, we face stakeholders as voices not specific individual persons. However, if we look at stakeholders for micro-theoretical topics, we may say that stakeholders are all humans (Harrison et al., 2022), and that is the reason (Bhattacharyya, 2020) concluded that more humanistic interactions would lead to stakeholder satisfaction. Moreover, we have the e-governments and their stakeholders for administrating countries and states (Rowley, 2011). E-government stakeholders are suitable example for stakeholders-analytics interaction, because it consists of a broad range of stakeholders as service users, citizens, businesses, other public administrations, and the employees and experts who analyze the information for policy makers, and for each of these groups has their own priority in communicating with technologies.

"*Human–computer interaction (HCI) is the area of intersection between psychology and the social sciences, on the one hand, and computer science and technology, on the other*."(Carroll, 1997), "*Human–computer interaction is an area of applied cognitive science and engineering design. It is concerned both with understanding how people make use of devices and systems that incorporate computation, and with designing new devices and systems that enhance human performance and experience*" (Carroll, 2006). To developing our concept Stakeholder-Analytics-Interaction SAI based on the previous definitions for HCI, we may define it as "*SAI is concerned with how an entity's stakeholders use and interact with analytics, as well as the extent to which analytics have been or have not been produced for successful interaction with stakeholders in order to create value for the whole value chain.*" The whole idea is that how an organization may reach out to its stakeholders to perceive their information, their opinions and experiences to do its responsibility for them and involve them in their decision-making process. In this way companies may use sentiment analysis in social media and develop their own systems such as *SentiProMo* (Lüftenegger & Softic, 2020), SentiProMo is a creative sentiment-based social business process modeling tool. (Lüftenegger & Softic, 2020)socially improved the BPM lifecycle in particular at the process analysis stage by capturing and processing stakeholder opinions about the tasks within a business process. SentiProMo translates these opinions with sentiment analysis and categorizes them as positive or negative feedback from the standpoint of social information systems, in this regard SentiProMo supports the redesigning of business processes by the information which have been extracted from stakeholders. Another similar technology has also been developed for manufacturing (Lüftenegger & Softic, 2021) (Softic & Lüftenegger, 2022), in which Sentiment Analysis Module (SAM) software has been trained and applied in real world cases to use the sentiment analysis to grasp the stakeholders' opinions for process improvement. (Hajek et al., 2014) forecasted U.S company's financial performance using sentiments in their annual reports. The findings reveal that sentiment information is an important predicting factor of financial success and can thus be utilized to aid business stakeholders' decision-making processes. Twitter is a valuable source of data for sentiment analysis and opinion mining that companies may utilize to improve their interactions with stakeholders (Aldahawi & Allen, 2013), in this regard, Twitter mining and sentiment analysis based on Twitter content may be defined as a method to know the stakeholders' opinions about an industry like oil industry, (Aldahawi & Allen, 2013) studied a sentiment analysis based on BP America and Saudi Aramco. It is suitable example that social media is not just for gaining data and information about consuming daily products and may also be used for more strategic industries (Jussila et al., 2017), also companies may use analytics to gain information and perceive their stakeholders' emotions during a specific crisis like COVID 19 pandemic (de las Heras-Pedrosa et al., 2020) during which communication is of great consequence for risks management.

Any way after TC and SCAOF, or by considering them as brain and nerve systems, and mind and thought process, an entity needs to communicate with its stakeholders to know its environment, to learn from their experiences and learn how to judge the conditions. Stakeholder Analytics Interaction SAI is going to feed the cognitive system, is going to make it more up to dated, knowledgeable, and responsive. The real learning would happen and start during the interaction. SAI is the moment an organization reaches out its hand to touch the texture of a problem, to smell it better, to look at its shape and be able to improve it after deep perception, and it has been conceptualized that within the future we may expect wearing technologies for complex managerial tasks for real-time decision makings purposes.

4. *Proposed Framework: Strategic Cognitive Analytics Management and Supply Chain Cognitive Capabilities*

In the following, Figure 1 displays the proposed framework for the Strategic Cognitive Analytics Management SCAM and Supply Chain Cognitive Capabilities SCCC.

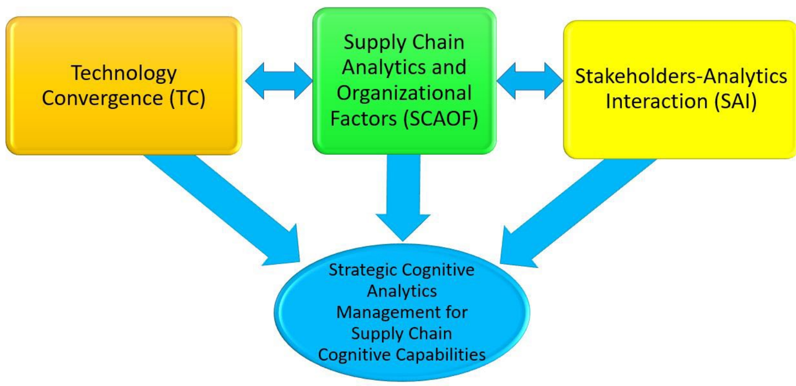


Figure 1. Strategic Cognitive Analytics Management and Supply Chain Cognitive Capabilities framework (SCAM-SCCC)

According to the review of the literature, this framework and its related theory have merged and been derived from Dynamic Capability View, Social Cognitive Theory, the theory of the Unified Theory of Acceptance and Use of Technology (UTAUT), Organizational Information Processing Theory (OIPT), and Stakeholder theory. Following previous articles on Cognitive Analytics Management CAM, it was determined that a comprehensive framework, especially from the perspective of supply chain management, may explain more about the characteristics of this idea. We built the proposed approach by assessing and examining existing literature. To the best of the authors' knowledge, this idea has never existed before, and its concept is suggested as the foundation of this framework. SCCC investigates and offers insights into the cognitive capacities that a supply chain organization may develop. The main purpose of SCCC is to promote a humanistic spirit in the supply chain. As a result, it strives to identify and evaluate its cognitive system by comparing it to human cognitive skills and physical and behavioral structures, which remains one of the most complicated problems that necessitates interdisciplinary scientists to detect its operations. The other topic is Stakeholders-Analytics Interaction SAI, a concept created in the present study and determined as one of the components of our proposed framework.

In section 2.1 we explained about the Technology Convergence TC and its interaction with supply chain analytics, by reviewing the previous papers as support to our proposal, we concluded it should go beyond the digital convergence when we consider SCAM, and the main reason is the unbounded scope of this concept, because it is not just Industry 4.0, rather it is the technologies and all technological developments which contribute to this field. This component has similarities with the three elements of Big Data, Machine-Machine Interaction (cognitive computing) and data archiving for future use of Conceptual Model Relating Big Data, and Cognitive Computing in (Gupta et al., 2018) framework. Analytics component of SMAS model in (Osman & Anouze, 2014), Systems in 7S model of McKinsey in (Osman & Anouze, 2014), and technology and analytics features of eight backbone structure of the SAMAS framework in (Osman & Anouze, 2014), by considering the mentioned frameworks, it would be noticed that all components create a system in their context and there should be collaboration between one each other.

The second component of SCAM-SCCC framework is Supply Chain Analytics and Organizational Factors (SCAOF), the supporting concepts to include this feature in our proposed framework is explained in section 2.2. Also by extracting the results from the reviewed papers, it has been supported that there is a mutual and complementary relationship between organization factors like Dynamic Capabilities (Majhi et al., 2021), organizational innovation and change management capabilities (Duan et al., 2020), (Božič & Dimovski, 2019), (Lozada et al., 2019) ,(Chatterjee et al., 2021)knowledge management capabilities (Chung et al., 2020), (Y. Wang & Byrd, 2017), (Khan & Vorley, 2017a), (Enad Al-Qaralleh & Atan, 2022), data and business security (Z. M. Bi et al., 2021), (Z. Wang et al., 2020), (Mahmood & Afzal, 2013)Flexibility, reconfigurability, agility, adaptability, and resilience (Z. Bi et al., 2021), (Napoleone et al., 2022), (Dubey, Gunasekaran, & Childe, 2019), (Dubey, Gunasekaran, Childe, Fosso Wamba, et al., 2019), (Yu et al., 2021)organizational culture and trust management (Dubey, Gunasekaran, Childe, Roubaud, et al., 2019), with supply chain analytics capabilities. This component has similarities with the two components of business insights, knowledge creation, of Conceptual Model Relating Big Data and Cognitive Computing in (Gupta et al., 2018), Activities, Structure, Mission, components of SMAS model in (Osman & Anouze, 2014), Style, Skills, Structure, Strategy, and Staff in 7S model of McKinsey in (Osman & Anouze, 2014), and innovation, improve, productivity, insights, and needs features of eight backbone structure of the SAMAS framework in (Osman & Anouze, 2014).

The third component of our proposed framework is Stakeholder-Analytics Interaction, This component has similarities with the three components of Human-Machine Interaction, Competitor Enterprise Existing knowledge, and new values of Conceptual Model Relating Big Data and Cognitive Computing in (Gupta et al., 2018), Shared Values component of SMAS model in (Osman & Anouze, 2014), Shared Values of 7S model of McKinsey (Osman & Anouze, 2014), and Shared Values of eight backbone structure of the SAMAS framework of (Osman & Anouze, 2014). Stakeholder-Analytics Interaction SAI is where communication takes place where various parties perceive the values and their expectations. The importance of discussing any innovation with stakeholders, as well as any financial and non-financial environmental, social, and governance performance requirements, was brought to light by the concept of stakeholder value management (Eccles & Serafeim, 2013). In general, shared values include the values that an organization proposes to all of its stakeholders. Some examples of these values include (accountability, empowerment, integrity, diversity, teamwork, collaboration, and partnerships). One way to evaluate them would be to consider the actual and intangible values of an organization's cost-risk, benefit-opportunity, and impact values as they relate to the stakeholders of that organization. This would be an example of a technique to evaluate the value of an organization's stakeholders (Osman & Anouze, 2014). In this way analytics has been proved as an effective way to realize the values especially tangible ones (Lüftenegger & Softic, 2020), (Lüftenegger & Softic, 2021), (Softic & Lüftenegger, 2022).

5. Discussion (Theoretical and Practical Implications)

To answer the question of what Cognitive Analytics Management is and why it is important for Supply Chain Management, we conducted a literature review. To the best of our knowledge, there has been no direct previous study in this field, so we attempted to create a framework and, later, a related theory to explain this idea from a managerial and strategic perspective. The Strategic CAM framework for our suggested Supply Chain Cognitive Capabilities SCCC has been proposed in this regard. This framework and its accompanying theory distinguishes between analytics and analytics management from the perspective of strategic management in corporate analytics. As a result, it focuses on the latter. A relationship between the CAM and SCM has been discovered by proposing three components that have been separately approved by previous empirical studies. The three components of the presented triple model are Technology Convergence TC, Supply Chain Analytics and Organizational Factors SCAOF, Stakeholder-Analytics Interaction SAI, and it has been proposed that there would be a mutual relationship between TC and SCAOF, and SCAOF and SAI. As stated in the theory development section, Business Analytics would lead to improved operational performance and, ultimately, value creation, and this framework would pave the way for future studies to have a perspective on CAM-SCM, with more researchers developing and evolving complementary features for this framework and theory. This term has emerged in this study, specifically for the SAI element; although its concept is derived from Human-Machine Interaction HMI, its perspective on the problem is different, and it may lead to more future studies about it, particularly when considering the SCM domain, which involves a broad network of stakeholders; it may lead to valuable contributions in this field. Overall, from a theoretical standpoint, this study can be regarded as a stepping stone in supply chain analytics management, shedding light on one of the novels, CAM-SCM.

From a practical standpoint, it helps consultants and practitioners understand CAM better, and our theory may be useful for technology managers and forecasters. Because this conception has benefited from the usage of a humanistic cognitive system. The three aspects of the humanistic cognitive system, brain and neurological system, mind and thought process, and action and motion, have been defined as the features in our triple framework that might reconcile CAM notions with SCM factors. As a result, technology convergence, of which digital convergence is only one component, may emphasize the importance of investing in and recognizing the technologies that have the potential to improve supply chain analytics capabilities, and that measures may lead to improved value chain and operational performance. SCAOF element contains the organizational factors as Big Data Analytics Capabilities, Data driven Culture, innovation and change management, knowledge management, safety and information security, flexibility, agility, adaptability, resilience, the three features of its Dynamic Capabilities as sensing capability, seizing capability, reconfigurability capability, and trust management capability, which all could lead to better supply chain performance in both terms of information and operations management, As a result, this feature sheds more light on the important factors that practitioners may be focused on for more value creation, and SAI emphasizes the importance of communicating with stakeholders, and analytics would be a facilitator to capture knowledge and provide the supply chain with a better perception of its stakeholders' expectations, which is a critical feature for survival in today's business environment.

6. Conclusion

There is a delicate distinction between technical and management difficulties in business analytics. This paper proposed a managerial framework to add to and cover the research gap in Cognitive Analytics Management CAM as a fresh field. A complete and analytical literature research was undertaken for this study to answer the question, "What is CAM and why is it significant for Supply Chain Management SCM?" As a result, we went through the previous definitions and frameworks, and we discovered that there has been no framework to explain CAM and SCM simultaneously. Furthermore, by looking at CAM from a strategic and macro level perspective, we came up with Strategic CAM (SCAM), which defines how it may be created and managed, and to link it to SCM, we proposed the novel Supply Chain Cognitive Capabilities SCCC theory. To begin the framework and its theory, the roots of the components of prior frameworks, as well as the results of previous empirical investigations and several case studies from reputable publications, were reviewed. Furthermore, the actual working process of a humanistic cognitive system has been helpful in the conceptualization process of this paper. The three aspects of the humanistic cognitive system, brain and neurological system, mind and thought process, and action and motion, have been defined as the features in our triple framework that might reconcile CAM notions with SCM factors. Furthermore, by categorizing the components based on their concepts and applications, we developed a triple framework that identifies Technology Convergence TC, Supply Chain Analytics and Organizational Factors SCAOF, and Stakeholders-Analytics Interaction SAI as the three elements that comprise SCAM. By reviewing previous papers in digital convergence and technology convergence, as well as the related digital paradigms and technologies such as Industry 4.0, Blockchain, IoT, new artificial intelligence driven technologies such as Google's LaMDA, or implanted technology in Neil Harbisson, who was recognized as the first Cyborg, who said "I don't feel like I'm using technology, I don't feel like I'm wearing technology, I feel like I am technology." And Nano and Micro Electromechanical Systems NEMS and MEMS applications in IoT and Blockchain have been endorsed, indicating that we are facing technology convergence for SCAM, with digital convergence being merely one component of that. Also included are Supply Chain Analytics and Organizational Factors. SCAOF, which is the core component of our framework due to its mutual relationship with the two other components TC and SAI, explains how a supply chain organization can use organizational factors such as Big Data Analytics Capabilities, Data Driven Culture, innovation and change management, knowledge management, safety and information security, flexibility, agility, adaptability, and resilience, as well as the three features of its Dynamic Capabilities as sensing capability, seizing capability, and capturing capability. Finally, the third component is the Stakeholders-Analytics Interaction. SAI is yet another new term that has evolved from this investigation. Previous studies emphasized Human-Computer Interaction HCI, and it is the technological perspective for this feature; however, as stated, we are looking for a managerial perspective of CAM, stakeholders may be customers, governments, and societies, so we should look at them as a holistic strategic view and know them as a voice, thus by giving a humanistic spirit to supply chain organizations to mimic and implement the cognitive system on its performance. Thus, SAI has concentrated on the relationship and communication with stakeholders in order to perceive their suggestions for adding value and being more creative, and it has also been noted in the text how some proposed unique sentiment analysis algorithms may contribute to SAI. Furthermore, a reinforcement cycle between the SCAOF with TC and SAI was discovered by examining prior empirical investigations. Finally, it should be noted that the purpose of this study is to present a new perspective on cognitive analytics management and its application in supply chain management, resulting in the creation of business value, and that the proposed triple framework and its related theory SCCC can be viewed as a steppingstone for taking, employing, and evolving this perspective in future multidisciplinary studies in supply chain management and information systems.

References:

Agarwal, R., & Dhar, V. (2014). Editorial—Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. *Https://Doi.Org/10.1287/Isre.2014.0546*, *25*(3), 443–448. https://doi.org/10.1287/ISRE.2014.0546

Albert Bandura. (1986). *Social foundations of thought and action: A social cognitive theory*. https://doi.apa.org/record/1985-98423-000

Aldahawi, H. A., & Allen, S. M. (2013). Twitter mining in the oil business: A sentiment analysis approach. *Proceedings - 2013 IEEE 3rd International Conference on Cloud and Green Computing, CGC 2013 and 2013 IEEE 3rd International Conference on Social Computing and Its Applications, SCA 2013*, 581–586. https://doi.org/10.1109/CGC.2013.101

Alshehri, F., & Muhammad, G. (2021). A Comprehensive Survey of the Internet of Things (IoT) and AI-Based Smart Healthcare. *IEEE Access*, *9*, 3660–3678. https://doi.org/10.1109/ACCESS.2020.3047960

Analytics & Cognitive | Deloitte. (2021). https://www2.deloitte.com/tw/en/pages/strategy-operations/solutions/analytics-and-cognitive.html

Araz, O. M., Choi, T. M., Olson, D. L., & Salman, F. S. (2020). Data Analytics for Operational Risk Management. *Decision Sciences*, *51*(6), 1316–1319. https://doi.org/10.1111/DECI.12443

Belhadi, A., Mani, V., Kamble, S. S., Khan, S. A. R., & Verma, S. (2021). Artificial intelligence-driven innovation for enhancing supply chain resilience and performance under the effect of supply chain dynamism: an empirical investigation. *Annals of Operations Research*, 1–26. https://doi.org/10.1007/S10479-021-03956-X/TABLES/5

Bhattacharyya, S. S. (2020). Humanistic Orientation in Firm–Stakeholder Technology-based Interaction and Its Impact on Stakeholder Satisfaction: *Https://Doi.Org/10.1177/2394901520907726*, *6*(1), 86–105. https://doi.org/10.1177/2394901520907726

Bi, Z., Jin, Y., Maropoulos, P., Zhang, W. J., & Wang, L. (2021). Internet of things (IoT) and big data analytics (BDA) for digital manufacturing (DM). *Https://Doi.Org/10.1080/00207543.2021.1953181*. https://doi.org/10.1080/00207543.2021.1953181

Bi, Z. M., Chen, B., Xu, L., Wu, C., Malott, C., Chamberlin, M., & Enterline, T. (2021). Security and safety assurance of collaborative manufacturing in industry 4.0. *Https://Doi.Org/10.1080/17517575.2021.2008512*. https://doi.org/10.1080/17517575.2021.2008512

Božič, K., & Dimovski, V. (2019). Business intelligence and analytics use, innovation ambidexterity, and firm performance: A dynamic capabilities perspective. *The Journal of Strategic Information Systems*, *28*(4), 101578. https://doi.org/10.1016/J.JSIS.2019.101578

Cardenas, A. A., Manadhata, P. K., & Rajan, S. P. (2013). Big data analytics for security. *IEEE Security and Privacy*, *11*(6), 74–76. https://doi.org/10.1109/MSP.2013.138

Carroll, J. M. (1997). Human–computer interaction: psychology as a science of design. *International Journal of Human-Computer Studies*, *46*(4), 501–522. https://doi.org/10.1006/IJHC.1996.0101

Carroll, J. M. (2006). Human–Computer Interaction. *Encyclopedia of Cognitive Science*. https://doi.org/10.1002/0470018860.S00545

Chamola, V., Hassija, V., Gupta, V., & Guizani, M. (2020). A Comprehensive Review of the COVID-19 Pandemic and the Role of IoT, Drones, AI, Blockchain, and 5G in Managing its Impact. *IEEE Access*, *8*, 90225–90265. https://doi.org/10.1109/ACCESS.2020.2992341

Chand Bhatt, P., Kumar, V., Lu, T. C., & Daim, T. (2021). Technology convergence assessment: Case of blockchain within the IR 4.0 platform. *Technology in Society*, *67*, 101709. https://doi.org/10.1016/J.TECHSOC.2021.101709

Chandrashekar, A., & Schary, P. B. (1999). Toward the Virtual Supply chain: The Convergence of IT and Organization. *The International Journal of Logistics Management*, *10*(2), 27–40. https://doi.org/10.1108/09574099910805978/FULL/XML

Changchit, C., & Chuchuen, C. (2016). Cloud Computing: An Examination of Factors Impacting Users' Adoption. *Https://Doi.Org/10.1080/08874417.2016.1180651*, *58*(1), 1–9. https://doi.org/10.1080/08874417.2016.1180651

Chatterjee, S., Chaudhuri, R., & Vrontis, D. (2021). Does data-driven culture impact innovation and performance of a firm? An empirical examination. *Annals of Operations Research*, 1–26. https://doi.org/10.1007/S10479-020-03887-Z/FIGURES/6

Chen, Y., Argentinis, E., & Weber, G. (2016). IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. *Clinical Therapeutics*, *38*(4), 688–701. https://doi.org/10.1016/J.CLINTHERA.2015.12.001

Chiang, R. H. L., Goes, P., & Stohr, E. A. (2012). Business Intelligence and Analytics Education, and Program Development. *ACM Transactions on Management Information Systems (TMIS)*, *3*(3). https://doi.org/10.1145/2361256.2361257

Choi, S., Afifuddin, M., Access, W. S.-I., & 2022, undefined. (2022). A Supervised Learning-Based Approach to Anticipating Potential Technology Convergence. *Ieeexplore.Ieee.Org*, *10*. https://ieeexplore.ieee.org/abstract/document/9714299/

Chung, W., Mustaine, E., & Zeng, D. (2020). A computational framework for social-media-based business analytics and knowledge creation: empirical studies of CyTraSS. *Https://Doi.Org/10.1080/17517575.2020.1827299*, *15*(10), 1460–1482. https://doi.org/10.1080/17517575.2020.1827299

Ciolacu, M. I., Binder, L., & Popp, H. (2019). Enabling IoT in Education 4.0 with BioSensors from Wearables and Artificial Intelligence. *SIITME 2019 - 2019 IEEE 25th International Symposium for Design and Technology in Electronic Packaging, Proceedings*, 17–24. https://doi.org/10.1109/SIITME47687.2019.8990763

Clarkson, M. E. (1995). A Stakeholder Framework for Analyzing and Evaluating Corporate Social Performance. *Https://Doi.Org/10.5465/Amr.1995.9503271994*, *20*(1), 92–117. https://doi.org/10.5465/AMR.1995.9503271994

Conboy, K., Mikalef, P., Dennehy, D., & Krogstie, J. (2020). Using business analytics to enhance dynamic capabilities in operations research: A case analysis and research agenda. *European Journal of Operational Research*, *281*(3), 656–672. https://doi.org/10.1016/J.EJOR.2019.06.051

de Bussy, N. M., Watson, R. T., Pitt, L. F., & Ewing, M. T. (2001). Stakeholder communication management on the Internet: An integrated matrix for the identification of opportunities. *Journal of Communication Management*, *5*(2), 138–146. https://doi.org/10.1108/13632540110806721/FULL/XML

de las Heras-Pedrosa, C., Sánchez-Núñez, P., & Peláez, J. I. (2020). Sentiment Analysis and Emotion Understanding during the COVID-19 Pandemic in Spain and Its Impact on Digital Ecosystems. *International Journal of Environmental Research and Public Health 2020, Vol. 17, Page 5542*, *17*(15), 5542. https://doi.org/10.3390/IJERPH17155542

Dinh, T. N., & Thai, M. T. (2018). AI and Blockchain: A Disruptive Integration. *Computer*, *51*(9), 48–53. https://doi.org/10.1109/MC.2018.3620971

Duan, Y., Cao, G., & Edwards, J. S. (2020). Understanding the impact of business analytics on innovation. *European Journal of Operational Research*, *281*(3), 673–686. https://doi.org/10.1016/J.EJOR.2018.06.021

Duan, Y., & Guan, Q. (2021). Predicting potential knowledge convergence of solar energy: bibliometric analysis based on link prediction model. *Scientometrics 2021 126:5*, *126*(5), 3749–3773. https://doi.org/10.1007/S11192-021-03901-6

Dubey, R., Bryde, D. J., Blome, C., Roubaud, D., & Giannakis, M. (2021). Facilitating artificial intelligence powered supply chain analytics through alliance management during the pandemic crises in the B2B context. *Industrial Marketing Management*, *96*, 135–146. https://doi.org/10.1016/J.INDMARMAN.2021.05.003

Dubey, R., Gunasekaran, A., & Childe, S. J. (2019). Big data analytics capability in supply chain agility: The moderating effect of organizational flexibility. *Management Decision*, *57*(8), 2092–2112. https://doi.org/10.1108/MD-01-2018-0119/FULL/XML

Dubey, R., Gunasekaran, A., Childe, S. J., Fosso Wamba, S., Roubaud, D., & Foropon, C. (2019). Empirical investigation of data analytics capability and organizational flexibility as complements to supply chain resilience. *Https://Doi.Org/10.1080/00207543.2019.1582820*, *59*(1), 110–128. https://doi.org/10.1080/00207543.2019.1582820

Dubey, R., Gunasekaran, A., Childe, S. J., Roubaud, D., Fosso Wamba, S., Giannakis, M., & Foropon, C. (2019). Big data analytics and organizational culture as complements to swift trust and collaborative performance in the humanitarian supply chain. *International Journal of Production Economics*, *210*, 120–136. https://doi.org/10.1016/J.IJPE.2019.01.023

Eccles, R. G., & Serafeim, G. (2013). The performance frontier: innovating for a sustainable strategy. *Harvard Business Review*, *91*(5), 50–56, 58, 60, 150. https://europepmc.org/article/med/23898735

Edwards, J. S., & Taborda, E. R. (2016). Using Knowledge Management to Give Context to Analytics and Big Data and Reduce Strategic Risk. *Procedia Computer Science*, *99*, 36–49. https://doi.org/10.1016/J.PROCS.2016.09.099

Ekramifard, A., Amintoosi, H., Seno, A. H., Dehghantanha, A., & Parizi, R. M. (2020). A systematic literature review of integration of blockchain and artificial intelligence. *Advances in Information Security*, *79*, 147–160. https://doi.org/10.1007/978-3-030-38181-3_8/FIGURES/3

Enad Al-Qaralleh, R., & Atan, T. (2022). Impact of knowledge-based HRM, business analytics and agility on innovative performance: linear and FsQCA findings from the hotel industry. *Kybernetes*, *51*(1), 423–441. https://doi.org/10.1108/K-10-2020-0684/FULL/XML

Esser, S. K., Andreopoulos, A., Appuswamy, R., Datta, P., Barch, D., Amir, A., Arthur, J., Cassidy, A., Flickner, M., Merolla, P., Chandra, S., Basilico, N., Carpin, S., Zimmerman, T., Zee, F., Alvarez-Icaza, R., Kusnitz, J. A., Wong, T. M., Risk, W. P., … Modha, D. S. (2013). Cognitive computing systems: Algorithms and applications for networks of neurosynaptic cores. *Proceedings of the International Joint Conference on Neural Networks*. https://doi.org/10.1109/IJCNN.2013.6706746

Ezerins, M. E., Ludwig, T. D., O'Neil, T., Foreman, A. M., & Açıkgöz, Y. (2022). Advancing safety analytics: A diagnostic framework for assessing system readiness within occupational safety and health. *Safety Science*, *146*, 105569. https://doi.org/10.1016/J.SSCI.2021.105569

Fatorachian, H., & Kazemi, H. (2020). Impact of Industry 4.0 on supply chain performance. *Https://Doi.Org/10.1080/09537287.2020.1712487*, *32*(1), 63–81. https://doi.org/10.1080/09537287.2020.1712487

Fisher, D., DeLine, R., Czerwinski, M., & Drucker, S. (2012). Interactions with big data analytics. *Interactions*, *19*(3), 50–59. https://doi.org/10.1145/2168931.2168943

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234–246. https://doi.org/10.1016/J.IJPE.2014.12.031

Fosso Wamba, S., Queiroz, M. M., Wu, L., & Sivarajah, U. (2020). Big data analytics-enabled sensing capability and organizational outcomes: assessing the mediating effects of business analytics culture. *Annals of Operations Research*, 1–20. https://doi.org/10.1007/S10479-020-03812-4/TABLES/6

Freeman, R. E. (2015). Strategic management: A stakeholder approach. *Strategic Management: A Stakeholder Approach*, 1–276. https://doi.org/10.1017/CBO9781139192675

Gill, S. S., Tuli, S., Xu, M., Singh, I., Singh, K. V., Lindsay, D., Tuli, S., Smirnova, D., Singh, M., Jain, U., Pervaiz, H., Sehgal, B., Kaila, S. S., Misra, S., Aslanpour, M. S., Mehta, H., Stankovski, V., & Garraghan, P. (2019). Transformative effects of IoT, Blockchain and Artificial Intelligence on cloud computing: Evolution, vision, trends and open challenges. *Internet of Things*, *8*, 100118. https://doi.org/10.1016/J.IOT.2019.100118

Gillis, A. S. (2022). *What Is Cognitive Computing?* https://www.techtarget.com/searchenterpriseai/definition/cognitive-computing

Greenwood, M. (2007). Stakeholder Engagement: Beyond the Myth of Corporate Responsibility. *Journal of Business Ethics 2007 74:4*, *74*(4), 315–327. https://doi.org/10.1007/S10551-007-9509-Y

Gudivada, V. N. (2016). Cognitive Computing: Concepts, Architectures, Systems, and Applications. *Handbook of Statistics*, *35*, 3–38. https://doi.org/10.1016/BS.HOST.2016.07.004

Gupta, S., Kar, A., … A. B.-I. J. of, & 2018, undefined. (2018). Big data with cognitive computing: A review for the future. *International Journal of Information Management*, *42*, 78–89. https://www.sciencedirect.com/science/article/pii/S0268401218304110

Hajek, P., Olej, V., & Myskova, R. (2014). Forecasting corporate financial performance using sentiment in annual reports for stakeholders' decision-making. *Vilnius Gediminas Technical University*, *20*(4), 721–738. https://doi.org/10.3846/20294913.2014.979456

Handfield, R., Jeong, S., & Choi, T. (2019). Emerging procurement technology: data analytics and cognitive analytics. *International Journal of Physical Distribution and Logistics Management*, *49*(10), 972–1002. https://doi.org/10.1108/IJPDLM-11-2017-0348/FULL/PDF

Harbisson, N. (2012). *Neil Harbisson: I listen to color | TED Talk*. https://www.ted.com/talks/neil_harbisson_i_listen_to_color?language=en

Harnessing Cognitive Analytics to Reinvent Your Business. (2019). https://10xds.com/blog/cognitive-analytics-to-reinvent-business/

Harrison, J. S., Wicks, A. C., & Palmié, M. (2022). Stakeholders Are Human: The Micro-foundations of Stakeholder Theory and an Application to the Value Distribution Problem. *Issues in Business Ethics*, *53*, 529–558. https://doi.org/10.1007/978-3-030-72204-3_36/COVER

Hasan, A. S. M. T., Sabah, S., Haque, R. U., Daria, A., Rasool, A., & Jiang, Q. (2022). Towards Convergence of IoT and Blockchain for Secure Supply Chain Transaction. *Symmetry 2022, Vol. 14, Page 64*, *14*(1), 64. https://doi.org/10.3390/SYM14010064

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research. *ACM Transactions on Computer-Human Interaction*, *7*(2), 174–196. https://doi.org/10.1145/353485

Hu, J., & Vasilakos, A. v. (2016). Energy Big Data Analytics and Security: Challenges and Opportunities. *IEEE Transactions on Smart Grid*, *7*(5), 2423–2436. https://doi.org/10.1109/TSG.2016.2563461

Hurwitz, J., Kaufman, M., & Bowles, A. (2015). *Cognitive computing and big data analytics*. https://books.google.com/books?hl=en&lr=&id=EI5xBgAAQBAJ&oi=fnd&pg=PR1&dq=Cognitive+Computing+and+Big+Data+Analytics&ots=1QWBZuvsNE&sig=XrD_plyC4L6_uz9S8vpXEpTbVbA

Ivanov, D., Dolgui, A., & Sokolov, B. (2018). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. *Https://Doi.Org/10.1080/00207543.2018.1488086*, *57*(3), 829–846. https://doi.org/10.1080/00207543.2018.1488086

Jussila, J., Vuori, V., Okkonen, J., & Helander, N. (2017). Reliability and Perceived Value of Sentiment Analysis for Twitter Data. *Springer Proceedings in Business and Economics*, 43–48. https://doi.org/10.1007/978-3-319-56288-9_7/TABLES/1

Kankanhalli, A., Charalabidis, Y., & Mellouli, S. (2019). IoT and AI for Smart Government: A Research Agenda. *Government Information Quarterly*, *36*(2), 304–309. https://doi.org/10.1016/J.GIQ.2019.02.003

Karvonen, M., Lehtovaara, M., & Kässi, T. (2012). BUILD-UP OF UNDERSTANDING OF TECHNOLOGICAL CONVERGENCE: EVIDENCE FROM PRINTED INTELLIGENCE INDUSTRY. *Http://Dx.Doi.Org/10.1142/S0219877012500204*, *9*(3). https://doi.org/10.1142/S0219877012500204

Khan, Z., & Vorley, T. (2017a). Big data text analytics: an enabler of knowledge management. *Journal of Knowledge Management*, *21*(1), 18–34. https://doi.org/10.1108/JKM-06-2015-0238/FULL/XML

Khan, Z., & Vorley, T. (2017b). Big data text analytics: an enabler of knowledge management. *Journal of Knowledge Management*, *21*(1), 18–34. https://doi.org/10.1108/JKM-06-2015-0238/FULL/XML

Kose, T., & Sakata, I. (2019). Identifying technology convergence in the field of robotics research. *Technological Forecasting and Social Change*, *146*, 751–766. https://doi.org/10.1016/J.TECHFORE.2018.09.005

Kuhn, T. (2008). A Communicative Theory of the Firm: Developing an Alternative Perspective on Intra-organizational Power and Stakeholder Relationships: *Http://Dx.Doi.Org/10.1177/0170840608094778*, *29*(8–9), 1227–1254. https://doi.org/10.1177/0170840608094778

Kumar, N. M., Chand, A. A., Malvoni, M., Prasad, K. A., Mamun, K. A., Islam, F. R., & Chopra, S. S. (2020). Distributed Energy Resources and the Application of AI, IoT, and Blockchain in Smart Grids. *Energies 2020, Vol. 13, Page 5739*, *13*(21), 5739. https://doi.org/10.3390/EN13215739

Kump, B., Engelmann, A., Kessler, A., & Schweiger, C. (2019). Toward a dynamic capabilities scale: measuring organizational sensing, seizing, and transforming capacities. *Industrial and Corporate Change*, *28*(5), 1149–1172. https://doi.org/10.1093/ICC/DTY054

Le, X., Shi, Q., Vachon, P., Ng, E. J., & Lee, C. (2021). Piezoelectric MEMS—evolution from sensing technology to diversified applications in the 5G/Internet of Things (IoT) era. *Journal of Micromechanics and Microengineering*, *32*(1), 014005. https://doi.org/10.1088/1361-6439/AC3AB9

Lee, J., & Sohn, S. Y. (2021). Recommendation system for technology convergence opportunities based on self-supervised representation learning. *Scientometrics*, *126*(1), 1–25. https://doi.org/10.1007/S11192-020-03731-Y/TABLES/4

Lee, J. W., & Sohn, S. Y. (2021). Patent data based search framework for IT R&D employees for convergence technology. *Scientometrics*, *126*(7), 5687–5705. https://doi.org/10.1007/S11192-021-04011-Z/TABLES/13

Liberatore, M. J., & Luo, W. (2010). The analytics movement: Implications for operations research. *Interfaces*, *40*(4), 313–324. https://doi.org/10.1287/INTE.1100.0502

Lim, E. P., Chen, H., & Chen, G. (2013). Business Intelligence and Analytics. *ACM Transactions on Management Information Systems (TMIS)*, *3*(4). https://doi.org/10.1145/2407740.2407741

Lozada, N., Arias-Pérez, J., & Perdomo-Charry, G. (2019). Big data analytics capability and co-innovation: An empirical study. *Heliyon*, *5*(10), e02541. https://doi.org/10.1016/J.HELIYON.2019.E02541

Luan, C., Deng, S., Porter, A. L., & Song, B. (2021). An Approach to Construct Technological Convergence Networks Across Different IPC Hierarchies and Identify Key Technology Fields. *IEEE Transactions on Engineering Management*. https://doi.org/10.1109/TEM.2021.3120709

Lüftenegger, E., & Softic, S. (2020). SentiProMo: A Sentiment Analysis-Enabled Social Business Process Modeling Tool. *Lecture Notes in Business Information Processing*, *397*, 83–89. https://doi.org/10.1007/978-3-030-66498-5_7/FIGURES/4

Lüftenegger, E., & Softic, S. (2021). Supporting Manufacturing Processes Design Using Stakeholder Opinions and Sentiment Analysis. *IFIP Advances in Information and Communication Technology*, *632 IFIP*, 112–117. https://doi.org/10.1007/978-3-030-85906-0_13/FIGURES/5

Mahmood, T., & Afzal, U. (2013). Security analytics: Big data analytics for cybersecurity: A review of trends, techniques and tools. *Conference Proceedings - 2013 2nd National Conference on Information Assurance, NCIA 2013*, 129–134. https://doi.org/10.1109/NCIA.2013.6725337

Majhi, S. G., Mukherjee, A., & Anand, A. (2021). Business value of cognitive analytics technology: a dynamic capabilities perspective. *VINE Journal of Information and Knowledge Management Systems*. https://doi.org/10.1108/VJIKMS-07-2021-0128/FULL/HTML

Manoharan, H., Teekaraman, Y., Kshirsagar, P. R., Sundaramurthy, S., & Manoharan, A. (2020). Examining the effect of aquaculture using sensor-based technology with machine learning algorithm. *Aquaculture Research*, *51*(11), 4748–4758. https://doi.org/10.1111/ARE.14821

Maymir-Ducharme, F., Systemics, L. A.-J. of, and, C., & 2014, undefined. (2014). Cognitive analytics: A step towards Tacit Knowledge? *Iiisci.Org*. http://www.iiisci.org/journal/PDV/sci/pdfs/HA342OE14.pdf

Menon, S. (2007). Policy Initiative Dilemmas Surrounding Media Convergence: A Cross National Perspective 1. *Http://Dx.Doi.Org/10.1080/08109020600563937*, *24*(1), 59–80. https://doi.org/10.1080/08109020600563937

Mishra, D., Luo, Z., Hazen, B., Hassini, E., & Foropon, C. (2019). Organizational capabilities that enable big data and predictive analytics diffusion and organizational performance: A resource-based perspective. *Management Decision*, *57*(8), 1734–1755. https://doi.org/10.1108/MD-03-2018-0324/FULL/XML

Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., & Martynenko, A. (2022). IoT, Big Data, and Artificial Intelligence in Agriculture and Food Industry. *IEEE Internet of Things Journal*, *9*(9), 6305–6324. https://doi.org/10.1109/JIOT.2020.2998584

Muhtaroglu, F. C. P., Demir, S., Obali, M., & Girgin, C. (2013). Business model canvas perspective on big data applications. *Proceedings - 2013 IEEE International Conference on Big Data, Big Data 2013*, 32–37. https://doi.org/10.1109/BIGDATA.2013.6691684

Napoleone, A., Negri, E., Macchi, M., & Pozzetti, A. (2022). How the technologies underlying cyber-physical systems support the reconfigurability capability in manufacturing: a literature review. *Https://Doi.Org/10.1080/00207543.2022.2074323*. https://doi.org/10.1080/00207543.2022.2074323

Osman, I. H., & Anouze, A. L. (2014). *A Cognitive Analytics Management Framework (CAM-Part 3)*. 190–234. https://doi.org/10.4018/978-1-4666-4474-8.CH003:

Osman, I. H., Anouze, A. L., Irani, Z., Lee, H., Medeni, T. D., & Weerakkody, V. (2019). A cognitive analytics management framework for the transformation of electronic government services from users' perspective to create sustainable shared values. *European Journal of Operational Research*, *278*(2), 514–532. https://doi.org/10.1016/J.EJOR.2019.02.018

Pape, T. (2016). Prioritising data items for business analytics: Framework and application to human resources. *European Journal of Operational Research*, *252*(2), 687–698. https://doi.org/10.1016/J.EJOR.2016.01.052

Park, H. S. (2017). Technology convergence, open innovation, and dynamic economy. *Journal of Open Innovation: Technology, Market, and Complexity*, *3*(4), 1–13. https://doi.org/10.1186/S40852-017-0074-Z/TABLES/3

Phillips-Wren, G., Iyer, L. S., Kulkarni, U., & Ariyachandra, T. (2015). Business Analytics in the Context of Big Data: A Roadmap for Research. *Communications of the Association for Information Systems*, *37*(1), 23. https://doi.org/10.17705/1CAIS.03723

Queiroz, M. M., Fosso Wamba, S., de Bourmont, M., & Telles, R. (2020). Blockchain adoption in operations and supply chain management: empirical evidence from an emerging economy. *Https://Doi.Org/10.1080/00207543.2020.1803511*, *59*(20), 6087–6103. https://doi.org/10.1080/00207543.2020.1803511

Rabhi, F. A., Bandara, M., Lu, K., & Dewan, S. (2021). Design of an innovative IT platform for analytics knowledge management. *Future Generation Computer Systems*, *116*, 209–219. https://doi.org/10.1016/J.FUTURE.2020.10.022

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, *2*(1), 1–10. https://doi.org/10.1186/2047-2501-2-3/FIGURES/1

Ragini, J. R., Anand, P. M. R., & Bhaskar, V. (2018). Big data analytics for disaster response and recovery through sentiment analysis. *International Journal of Information Management*, *42*, 13–24. https://doi.org/10.1016/J.IJINFOMGT.2018.05.004

Rana, N. P., & Dwivedi, Y. K. (2015). Citizen's adoption of an e-government system: Validating extended social cognitive theory (SCT). *Government Information Quarterly*, *32*(2), 172–181. https://doi.org/10.1016/J.GIQ.2015.02.002

Rejeb, A., Keogh, J. G., & Treiblmaier, H. (2019). Leveraging the Internet of Things and Blockchain Technology in Supply Chain Management. *Future Internet 2019, Vol. 11, Page 161*, *11*(7), 161. https://doi.org/10.3390/FI11070161

Reyna, A., Martín, C., Chen, J., Soler, E., & Díaz, M. (2018). On blockchain and its integration with IoT. Challenges and opportunities. *Future Generation Computer Systems*, *88*, 173–190. https://doi.org/10.1016/J.FUTURE.2018.05.046

Richey, R. G., Morgan, T. R., Lindsey-Hall, K., & Adams, F. G. (2016). A global exploration of Big Data in the supply chain: Global exploration of Big Data. *International Journal of Physical Distribution and Logistics Management*, *46*(8), 710–739. https://doi.org/10.1108/IJPDLM-05-2016-0134/FULL/HTML

Rim, M. H., Lim, K. S., & Sawng, Y. W. (2007). A business model analysis for the convergence services of supply chain. *Portland International Conference on Management of Engineering and Technology*, 2325–2335. https://doi.org/10.1109/PICMET.2007.4349566

Roland, C. (2022). Technology Convergence 101 - AT&T Developer. In *AT&T developer program*. https://developer.att.com/blog/technology-convergence

Rowley, J. (2011). e-Government stakeholders—Who are they and what do they want? *International Journal of Information Management*, *31*(1), 53–62. https://doi.org/10.1016/J.IJINFOMGT.2010.05.005

Ruehl, C. H., & Ingenhoff, D. (2015). Communication management on social networking sites: Stakeholder motives and usage types of corporate Facebook, Twitter and YouTube pages. *Journal of Communication Management*, *19*(3), 288–302. https://doi.org/10.1108/JCOM-04-2015-0025/FULL/XML

Sariyer, G., Mangla, S. K., Kazancoglu, Y., Ocal Tasar, C., & Luthra, S. (2021). Data analytics for quality management in Industry 4.0 from a MSME perspective. *Annals of Operations Research 2021*, 1–29. https://doi.org/10.1007/S10479-021-04215-9

Sharma, B., Sharma, M., Singla, B. S., & Goyal, S. (2021). *Design and Analysis of Thin Micro-Mechanical Suspended Dielectric RF-MEMS Switch for 5G and IoT Applications*. 133–146. https://doi.org/10.1007/978-981-16-1048-6_10

Softic, S., & Lüftenegger, E. (2022). Towards Empowering Business Process Redesign with Sentiment Analysis. *Lecture Notes in Networks and Systems*, *236*, 119–126. https://doi.org/10.1007/978-981-16-2380-6_10/COVER

Srinivasan, R., & Swink, M. (2018). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective. *Production and Operations Management*, *27*(10), 1849–1867. https://doi.org/10.1111/POMS.12746

Tang, Y., Lou, X., Chen, Z., Sustainability, C. Z.-, & 2020, undefined. (2020). A study on dynamic patterns of technology convergence with IPC co-occurrence-based analysis: The case of 3D printing. *Sustainability*. https://doi.org/10.3390/su12072655

Tantalo, C., & Priem, R. L. (2016). Value creation through stakeholder synergy. *Strategic Management Journal*, *37*(2), 314–329. https://doi.org/10.1002/SMJ.2337

Teece, D. J. (2007). Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350. https://doi.org/10.1002/SMJ.640

Tiku, N. (2022). *Google engineer Blake Lemoine thinks its LaMDA AI has come to life*. Technology, Washington Post. https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/

Tjahjono, B., Esplugues, C., Ares, E., & Pelaez, G. (2017). What does Industry 4.0 mean to Supply Chain? *Procedia Manufacturing*, *13*, 1175–1182. https://doi.org/10.1016/J.PROMFG.2017.09.191

van Rijmenam, M., Erekhinskaya, T., Schweitzer, J., & Williams, M. A. (2019). Avoid being the Turkey: How big data analytics changes the game of strategy in times of ambiguity and uncertainty. *Long Range Planning*, *52*(5). https://doi.org/10.1016/J.LRP.2018.05.007

Verma, R., Kantarcioglu, M., Marchette, D., Leiss, E., & Solorio, T. (2015). Security analytics: Essential data analytics knowledge for cybersecurity professionals and students. *IEEE Security and Privacy*, *13*(6), 60–65. https://doi.org/10.1109/MSP.2015.121

Wang, G., Gunasekaran, A., Ngai, E. W. T., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98–110. https://doi.org/10.1016/J.IJPE.2016.03.014

Wang, Y., & Byrd, T. A. (2017). Business analytics-enabled decision-making effectiveness through knowledge absorptive capacity in health care. *Journal of Knowledge Management*, *21*(3), 517–539. https://doi.org/10.1108/JKM-08-2015-0301/FULL/XML

Wang, Z., Wang, N., Su, X., & Ge, S. (2020). An empirical study on business analytics affordances enhancing the management of cloud computing data security. *International Journal of Information Management*, *50*, 387–394. https://doi.org/10.1016/J.IJINFOMGT.2019.09.002

Weerakkody, V., Irani, Z., Lee, H., Osman, I., & Hindi, N. (2013). E-government implementation: A bird's eye view of issues relating to costs, opportunities, benefits and risks. *Information Systems Frontiers 2013 17:4*, *17*(4), 889–915. https://doi.org/10.1007/S10796-013-9472-3

Xu, Q., Yang, Z., Sun, Y., Lai, L., Jin, Z., Ding, G., Zhao, X., Yao, J., & Wang, J. (2017). Shock-Resistibility of MEMS-Based Inertial Microswitch under Reverse Directional Ultra-High g Acceleration for IoT Applications. *Scientific Reports 2017 7:1*, *7*(1), 1–12. https://doi.org/10.1038/srep45512

Yu, W., Zhao, G., Liu, Q., & Song, Y. (2021). Role of big data analytics capability in developing integrated hospital supply chains and operational flexibility: An organizational information processing theory perspective. *Technological Forecasting and Social Change*, *163*, 120417. https://doi.org/10.1016/J.TECHFORE.2020.120417

31.

The Emergence of New Technologies on Supply Chain Management: A risk Management Assessment

Peace Aludogbu

New Mexico State University, Las Cruces, USA

Abstract

Purpose – This research aims to determine the risks associated with the emergence of new technologies incorporated into the different factors of the supply chain, such as time, flexibility, cost, and production.

Design/Methods - The research will involve an online email survey to collect data from two manufacturing companies based in Nigeria. The analysis will employ structural equation modeling to test the hypothesis.

Theoretical implications – Many works of literature have been written on this. However, most research was focused on North America, Europe, and Asia. The gap this article aims to address is that it uses Nigeria, a west African country, as a source for sample collection. Nigeria is the most populous country and largest African economy. By utilizing Nigerian data, the article aims to improve our understanding of this topic by utilizing data from Nigeria that has significant cultural insights for researchers.

Practical Implications – This research will assist investors and entrepreneurs in the supply chain and logistics field in managing activities operational activities using technology.

Keywords: Supply chain, Risk assessment, Flexibility, Time, Cost, Production, Nigeria

Categories

Supply Chain Management and Logistics

43.

Inventory and firm performance analysis in the pharmaceutical industry

Jeong Hoon Choi

Youngstown State University, Youngstown, USA

Abstract

The study explores and expands the theoretical foundations of the strategic group concepts through investigations on the inventory-firm performance relationship. Both total inventory and its discrete inventory are considered. We developed the Earns-Turns matrix based on the DuPont analysis and mapped strategic groups in the pharmaceutical industry with earns (or profitability) and inventory turns (or effective inventory management). We found that majority of pharmaceutical companies pursue profitability, moving from the bottom-right section to the top-left section in the Earns-Turns matrix. This study broadens the scope of operations management research by introducing the earns-turns matrix as an empirical validation tool for operational and strategic management theories.

Categories

Supply Chain Management and Logistics

62.

A Behavioral Experiment on Inventory Decision-making in Global Supply Chains

Sourish Sarkar [1], Sanjay Kumar [2]

[1]Penn State Behrend, Erie, USA. [2]Valparaiso University, Valparaiso, USA

Abstract

Using the beer game framework and controlled laboratory settings, we conducted behavioral experiments on supply chains that are partially or completely located in the US and in China. Several scenarios are compared to examine the supply chain costs and associated bullwhip effect. We provide an explanation of the outcome by considering cultural influence on decision-making.

Categories

Supply Chain Management and Logistics

68.

Usages of Drone as Transportation in Supply Chain: Using Mixed-Integer Linear Programming and Assignment Linear programming Approaches

OMAR FARUQ OSAMA, Abdur Rahman

Emporia State University, Emporia, USA

Abstract

The use of drones in supply chain management has been gaining popularity as a reliable and secure approach, especially during natural disasters, calamities, or pandemics. With isolated people in need of supplies and social distancing measures in place, the traditional methods of transportation may not always be feasible or efficient. Drones provide a solution to this problem by allowing for socially distant delivery services that can be more resilient and adaptive to changing circumstances. Additionally, drones can help businesses boost productivity, cut costs, and meet customer expectations in the e-commerce industry. However, implementing drones into supply chains does come with its own set of challenges, such as regulatory barriers and technological limitations. Research is needed to examine the decision-making process of using drones for essential deliveries and to measure the efficiency and effectiveness of transportation from beginning to end. Overall, drones can play a significant role in improving transportation in supply chain management during times of crisis and beyond. In terms of decision-making for using drones as a transport system in product supply, we will introduce a model and use quantitative approaches such as mixed-integer linear programming and assignment linear programming to verify whether the proposed model is cost and time-efficient and Dikshtra's algorithm to find the shortest path. The results of this research will help reduce delivery costs compared to traditional systems, provide fast and emergency supplies on demand at any time and any place within a certain distance, and reduce pollution caused by traffic congestion.

Categories

Supply Chain Management and Logistics

Supply Chain Resilience & Sustainability

8. A Qualitative Study on Green Offices in Turkey: Can the Green Office Concept Be an Alternative to the Remote Work?

AHU KARABULUT [1], CIHAN BAYRAM [2]

[1]Istanbul Commerce University, Istanbul, Turkey. [2]Istanbul, Istanbul, Turkey

Abstract

The purpose of this research is to explore the green office practices (GOPs) and their effects on the employee well being, satisfaction and commitment to the organizations. It also examines whether green office practices affect company performance and be an alternative to remote work during and after pandemia. A structured in-depth interview method is used to get more comprehensive and organized data from 6 respondent companies which have had World Wide Fund (WWF) Green Office certificates to reach research goals. As it is observed in this research, GOPs affect employee well being, job satisfaction, organizational commitment, and company performance. However, due to pandemia conditions, it is not considered as an alternative for remote work practices.

Keywords: Green Office Certificate, WWF, Well Being, Company Performance, Remote Work

1. Introduction

1.1. Green Office Concept

Green office offers sustainable office practices for reducing environmental effects of office activities and making contribution to the sustainable development. The basic principle of creating green offices is promoting improvements leading eliminated or reduced environmental effects of an office work. Coşkun and Akar (2019) believe that there are three criterias for a green office: 1) consuming less natural resources (electricity, water, paper, etc.), 2) focusing on sustainable activities (waste management, building maintenance), and 3) increasing awareness of employees. European Green Office Handbook has environmental performance indicators for a green office creation: Efficiencies of location, structural design, energy, water, and material; and improving the quality of the office environment, decreasing toxic and waste and optimizing operations and maintenance (Povodör et al., 2010) (Brazdauskas and Zirne le, 2020, p. 62).

Brown, Cole, Robinson and Dowlatabadi (2010) suggest that green building and organization should be considered as an integrated system which include protection of an environment and social responsibility for sustainability. Green office is an organizational tool to implement the sustainable development. Zou and Couani (2012) state that the green office demand is growing due to the demand for sustainable development and decrease of operational costs (Varanavicius and Navikaite, 2015, p. 145).

1.2. WWF Green Office Programs and General View on the Green Offices in Turkey

Green Office is an environmental management system which was developed by WWF Finland and operates in more than 200 companies in Turkey, China, Latvia, and Pakistan. The objectives of the program are decreasing carbon emissions and ecological footprint by reducing negative effects of operations, increasing environmental awareness, motivating employees to behave environmentally friendly ways, and conserving energy. The Green Office program lets companies to evaluate their resources, create environmental management systems and determine savings criterias. It encourages employees to choose environmentally responsible alternatives to save more (Coşkun and Akar, 2019, pp. 49-54). The Green Office system offers solutions for organizations to decrease their costs and ecological footprints. It is beneficial for both organizations and the environment. It is adaptable for small and large offices in private and public sectors (WWF Finland, 2011). WWF wishes to accomplish these targets through the Green Office environmental management system (WWF International, 2011):

- Decreasing consumption of natural resources by improving the environmental efficiency
- Promoting sustainable practices by increasing employees' environmental awareness
- Promoting climate change mitigation by using renewable energy sources and saving energy (Rauatmaa, 2011, p. 29).

The process of becoming a Green Office is as follows (WWF Finland, 2011; Sirviö, 2011) (Rauatmaa, 2011, p. 29):

1) Collaboration Agreement: Signed between WWF and the company, the company pays the admission fee
2) Go-Coordinator & Go-Team: Chosen from the employees for environmental management system implementation
3) Self Assessment: Assessment form is used for defining environmental effects
4) Environmental Review: Conducted by the go-team or an external specialist for choosing indicators
5) Environmental Programme: Environmental principles (target, objectives and environmental policy), responsibilities and schedules, consumer behaviour questionnaire
6) Communication & Training
7) Implementation & Follow-Up
8) Inspection: Necessary information from compass
9) Green Office Diploma & Green Office Logo: Companies receive Green Office Diploma and rights to use the green office logo for communication
10) Annual Report: Indicators and measures, environmental programme, consumer behaviour questionnaire
11) Continuous Improvement: Aim for continuous improvement, office inspection every three years

Companies participate in WWF-Turkey's Green Office program voluntarily by considering the ecological sustainability of their activities. They make environmental choices (e.g., office stationery, mode of travel), increase the sense of responsibility of their employees for protecting the environment (e.g., regular environmental training sessions), and save more energy and water and enhance recycling rates to reduce carbon emissions within the program. The Green Office Program was launched for decreasing the ecological footprint of organizations in 2011 in Turkey (Coşkun and Akar, 2019, p. 55).

WWF-Turkey audits annual performance of Green Office programs by using web-based tools and visiting offices. WWF-Turkey makes a new agreement with the company for the next period if it does not meet its targets at the end of a year. Accomplishment rates are determined by comparing current indicators with results of the previous year. A company can use its Green Office diploma for new goals each year or maintains current levels and pays an annual subscription fee as long as its collaboration continues. WWF-Turkey and the company establish goals for the next year for the continuity of the program (Coşkun and Akar, 2019, p. 51).

Customers are more sensitive to environmental problems such as pollution and companies face social and regulatory pressures to reorganize their activities to minimize their harms to the environment. The Green Office diploma and logo help companies to add value to their environmental strategies and improve their corporate images. Thus, if companies don't improve their performances year by year, they will try to maintain their current status to keep their program diplomas. WWF-Turkey cooperates with Green Office companies every year by sharing up-to-date information so they can continue to be greener offices. WWF-Turkey aims to work in harmony with the nature by providing guidance and consultancy for the Green Office program (Coşkun and Akar, 2019, p. 55).

Table 1. WWF Green Office Criterias

WWF Green Office Criterias	
1) Consumption of Natural Resources	**2) Functions**
Heating/Cooling	Purchasing
Electricity consumption	Cafeteria
Transportation	Cleaning services
Water Consumption	Waste Management
Paper Consumption	Building Maintenance
Office Supplies	**3) Increasing Employee Awareness**
Food Consumption	

(Retrieved by 03.02.2022 from https://www.wwf.org.tr/sizneyapabilirsiniz1/isinize_dogayi_katin/yesilofisolun/sistemnasilisliyor/)

2. Literature Review

2.1. Green Office Effects on Organizational Outcomes

Some researches reveal that energy-savings in offices can improve employee performance and wellbeing (Agha-Hossein et al., 2013) (Kozusnik, 2019, p. 3). It is possible to decouple high levels of performance and comfort from high energy-related costs. The theory of person-environment fit (Kaplan, 1983) focuses on the human environment interface, the supportive role of the environment for attainment of goals and basic processes such as attention, perception etc. It is based on the analysis of environment as a source of necessary action (Barker, 1968) and the concept of "behavior environment congruence" (Wicker, 1973) which considers supportive workplaces are important for functioning instead of being a source of constraint or pressure. Person-environment fit theoretical considerations suggest that an adequate office environment can improve performance and comfort (Vischer, 2007). This is consistent with the work characteristics approach (Hackman and Oldham, 1976) according to which the design of optimally challenging work characteristics can motivate employees internally and increase their performances (Kozusznik et al., 2019, p. 2).

Vischer (2008) proposes a model arranging workplace comfort into a hierarchical framework which has three comfort categories. The different types of comfort are based on workspace quality attributes. Comfort is the link between satisfaction and productivity in Vischer's (2008) model (Nurick and Thatcher, 2021, pp. 27-28).

2.1.1. Green Office Effects on Employee Wellbeing

Currie (2001) views employee wellbeing at work as the mental and physical health of the workforce. Thus, employees should be working in a physically safe and stress-free environment. Bakke (2005) adds that wellbeing can be linked to promoting an environment which make-work stimulating, exciting, enjoyable, and rewarding, and proposes that joy-filled workplaces improve financial performance (Baptiste, 2008, p. 291).

Wellbeing can be defined in terms of strain. Strain is physical, psychological, or behavioral responses to stressors (Le Fevre et al., 2003). The term stressor indicates job or organizational conditions which require adaptive responses from employees (Jex, 1998). Strain is manifested in the form of depressed mood, job dissatisfaction, headache, anxiety, absenteeism, coronary heart disease, turnover and poor performance (Jex, 1998; Cooper et al., 2001) (Jain et al., 2009, p. 258).

Feige et al. (2013) show that there is a relationship between building features and comfort (Nurick and Thatcher, 2021, p. 28).

Liquna and Yanqunb (2011) acknowledge that people spend at least 85% of 24 hours in indoor in their lives. Thus, indoor air quality is an important issue for health of people. Poor indoor air quality and pollution can cause injuries in the health of employees. It also affects their lives and productivity. Building and materials used for building decoration can cause indoor pollution. It can harm the health of employees more than outdoor pollutions. 68% of diseases such as fatigue, skin aging, forgetfulness, hair shedding, infertility, cancer, and leukemia are related to poor indoor air quality or indoor pollutions (Liquna and Yanqunb, 2011). Liquna and Yanqunb (2011) add that indoor environment cause 37% of the respiratory diseases, 22% of chronic disease, 5% of leukemia, 5% of bronchial catarrh, and 5% of cancers (Ghodrati et al., 2012, p. 4235).

2.1.2. Green Office Effects on Job Satisfaction

Hoppock (1935) defined job satisfaction as any combination of physiological, psychological, and environmental circumstances which led a person to tell others "I am satisfied with my job" (Hoppock, 1935). Spector believes that job satisfaction is the way how an employee feels about his job and its aspects. It shows whether an employee likes or dislikes his job. Job satisfaction represents a combination of positive or negative feelings which employees have towards their work. Job satisfaction is the extent to which expectations are and match the real awards (Davis and Nestrom, 1985). Job satisfaction is an employee's sense of achievement and job success. It is linked to personal well-being and productivity. Job satisfaction is doing an enjoyable job, doing it well and being rewarded for the efforts. Job satisfaction shows enthusiasm and happiness with work. Job satisfaction is the key ingredient which leads to recognition, income, promotion, and achievement of goals which lead to a feeling of fulfillment (Kaliski, 2007). Job satisfaction is related to rewards of employee for his job, especially in terms of intrinsic motivation (Statt, 2004). The term job satisfaction is the attitute and feeling of an employee about his work. Favorable and positive

attitudes towards the job indicate job satisfaction. Unfavorable and negative attitudes towards the job indicate job dissatisfaction (Armstrong, 2006). Job satisfaction is the combination of beliefs and feelings which an employee has about his current job. Besides having attitudes about his job as a whole, an employee can have attitudes about aspects of his job such as kind of work he does, his pay, supervisors, coworkers, and subordinates (George and Jones, 2008) (Aziri, 2011, pp. 77-78).
Leder et al. (2016) believe that the size of the workplace affects the satisfaction of an employee, if the workplace is larger, the satisfaction level of the employee is greater. Larger workstations place employees further apart, decrease the number of employees to overhear conversations and create unwanted sound, facilitate sound attenuation, and support visual privacy to increase the satisfaction of employees (Frontczak et al., 2012; Kim and de Dear, 2013; Schiavon and Altomonte, 2014).
Employees state their satisfactions with the thermal environment and comfort (Brager and Baker, 2009). Workplace design can increase job satisfaction, positive attitudes, and desires of employees for the environment (Monfared and Sharples, 2011; Deuble and de Dear, 2012). Carlopio (1996) reveals that workplace design satisfaction of employees is directly related to their job satisfactions and indirectly related to their organizational commitments. Workplace design is connected to job satisfactions and organizational commitments of employees (Bangwal et al., 2017, p. 2).

2.1.3. Green Office Effects on Organizational Commitment

The concept of commitment was introduced in 1960 by Becker to the literature by explained as “one mechanism producing consistent human behavior” (Becker, 1960). Porter et al. (1974) revealed the relationship between organizational commitment, job satisfaction and turnover intentions. Morrow (1983) proposes that job characteristics, career, union and personal values have essential roles for the employee commitment. Reichers (1985) defines organizational commitment as “a collection of multiple commitments to various groups which comprise the organization”. Allen and Meyer (1991) classified three dimensions of organizational commitment as “Affective” (AC), “Continuance” (CC) and “Normative” (NC) commitment. According to them (1991), employees who have higher levels of affective commitment remain in their companies because they “want to”, employees who have strong continuance commitment levels stay in their companies because they “need to” and employees who have strong normative commitment remain in their companies because they feel they “ought to” do. According to them (1991), organizational commitment is “a psychological state which characterizes the employee’s relationship with the organization and has implications for the decision to continue or discontinue membership in the organization” (Berberoğlu, 2018, p. 3).
According to Meyer, Allen and their colleagues (Allen and Meyer, 1990; Meyer and Allen, 1984, 1991; Meyer et al., 1990; Meyer et al., 1993; Meyer et al., 1989) affective, normative and continuance commitment are dimensions of attitudinal commitment. Allen and Meyer (1990, p. 1) define affective commitment as an “employee's emotional attachment to, identification with, and involvement in the organization”, continuance commitment (Hackett et al., 1994; Mathieu and Zajac, 1990) as “commitment based on the costs which employees associate with leaving the organization”, and normative commitment as an “employee's feelings of obligation to remain in the organization”. Normative commitment focuses on the “right or moral thing to do” (Weiner, 1982, p. 421), and employees’ moral attachment and/or

obligation occured by their socialization to organizational values and goals (Allen and Meyer, 1990; Weiner, 1982) (Iverson and Buttgirieg, 1999, pp. 308-309).
There are not many studies focusing on the relationship between being a green office and organizational commitment. Generally, it is observed that the relationship between being a green office and other organizational outputs is the main interest of the previous studies.

2.1.4. Green Office Effects on Job Performance

Work performance is a measure showing how an employee performs against the expected work tasks (Christiansen and Chandan, 2017). Level of work performance can vary based on the level of comfort with IEQ of workplace (Ali et al., 2015). Employees who are more satisfied with their physical workplace conditions are more motivated and can have higher work performances (Leaman, 1995; Dole and Schroeder, 2001) (Elnaklah, Fosas, and Natarajan, 2020, p. 1046).
Performance of the employee reflects what an employee does and what he doesn't do. Employee performance involves presence at work, quality and quantity of output, timeliness of output and helpful and accommodative nature. Yang (2008) reveals that performance of an employee cannot be verified. He suggests that companies can give rewards and bonuses if an employee shows a noticeable performance (Yang, 2008). Bishop (1987) examines employee performance and shows that recognition, acknowledgment and reward of employee performance lead to the discrimination of employee productivity. Moral and productivity of employees are affected by the effectiveness of the company performance and its reward system (Yazıcı, 2008) (Shahzadi et al., 2014, p. 161).
According to Kaplan and Norton (1992), one measure can not provide a clear performance target so managers should choose between operational and financial measures. They (1992) propose the following dimensions of Balance Scorecard as measures of organizational performance: 1. financial perspective; 2. customer perspective; 3. internal business perspective; and 4. learning perspective (Ahmed and Shafiq, 2014, p. 22).
Balance Scorecard is the one of most important framework to ensure that the strategy is translated into a rational performance measurement set (Kaplan and Norton, 1992). Perspectives of balance scorecard are as follows (Chavan, 2009) i. Financial perspective: How should we appear to our shareholders to succeed financially? Some of its measures are asset utilization, improved shareholder value, and return on capital ii. Customer perspective: How should we appear to our customers to achieve our vision? Some of its measures are reputation and image, customer relations and product/service qualities. iii. Internal business processes: At what business processes must we excel to satisfy our shareholders and customers? Some of its measures are after-sales services, produced products and services, and delivered products and services. iv. Learning and growth perspective: How will we maintain our ability to change and improve to achieve our vision? Some of its measures are motivation, empowerment, employee capabilities and information system capabilities (Ahmed and Shafiq, 2014, pp. 22-23).
Energy-efficient solutions in office buildings have a potential to improve person-environment fit. They include features to create a flexible work environment with spaces for tasks, such as meeting rooms, open plan offices, and areas for concentration (Agha-Hossein et al., 2013) to improve the fit between employee and his environment. Good-quality lighting supports interpersonal communication and visual performance essential for employee comfort (Nagy et al., 2015) and performance. There are practical building solutions created

from advances in sustainable technologies for better IEQ conditions by improving air quality (Fisk, 2000), privacy, personal workspace, lighting, and acoustics (Leder et al., 2016; Ornetzeder et al., 2016), decreasing density (Fisk, 2000) and improving the esthetics and workplace image (Newsham et al., 2013). These features can have an effect on the fit between the employee and the office space (Kozusnik, 2019, p. 3).

A theoretical model developed by Seppa¨nen and Fisk (2006) reveals a link between environmental controls (IEQ) with regards to maintenance and operational costs/benefits, including a decrease in health and maintenance costs, increase in staff retention, number of working days, and output quality (individual productivity). They conclude that the implementation of GBFIs enhancing IEQ have a direct positive financial effect on the organization without any intermediary variables (although health and productivity are intermediary variables). Doggart (2006) supports that by showing a well-designed building from a user comfort perspective can increase a financial performance of a company. He (2006) finds that overhead costs per an employee can decrease up to 15% in a short time showing that good building design is good for business (Nurick and Thatcher, 2021, p. 33).

3. Research Framework and Methodology

Purpose of the Research:

The purpose of this research is to explore the green office practices and their effects on the employee well being, satisfaction and commitment to the organizations. It also examines whether green office practices affect company performance and be an alternative to remote work during and after pandemia.

Scope:

Companies face difficulties while they are transforming their offices from traditional offices to green offices. This process affects employees and organizations.

Method:

The research has been conducted using the structured in-depth interview method which is a qualitative research method. A structured in-depth interview method is used to get more comprehensive and organized data from respondent companies to reach research goals.

Six interviews were conducted with managers of six companies which have had green office certificates from WWF in Turkey. The in-depth interview method is used to examine the green office certification process, the applications which have been made by the companies, the challenges and benefits of being a green office company, the effect of applications on employee job satisfaction, wellbeing, commitment and overall company performance. Thus, it reveals the green office certification process for the companies and its effects on organizational outputs in Turkey for the first time. Considering that the tendency of most of the companies to the issue of sustainability has increased in recent years, it is thought that researching the processes of being a green office company and the organizational outputs which these companies achieved, will contribute to both the academic field and the business world.

Limitations:

The main limitation of this research is the data was gathered from companies which received green office certificates from WWF. Certified companies by other organizations were not included to the research. There are few companies which have had WWF Green Office certificates and some of them are in the list of WWF but they are not applying the requirements of the certificate extensively. There are time, budget and pandemic limitations for the research as well. Structured in-depth interviews were conducted via online platforms such as Teams and Zoom due to pandemia, time and budget limitations. Researchers could not conduct face to face in-depth interviews with the respondent companies.

Research Questions:

- What are the requirements and practices for a Green Office?
- What are the effects of Green Office practices on employee well being?
- What are the effects of Green Office practices on job satisfaction?
- What are the effects of Green Office practices on organizational commitment?
- What are the effects of Green Office practices on company performance?
- Can Green Office became an alternative to remote working?

Key Hypotheses:

Green Office practices affect employee well being
Green Office practices affect job satisfaction
Green Office practices affect organizational commitment
Green Office practices affect company performance
Green Office can be an alternative to remote working

3.1. Data Collection

For understanding the Green Office Practices (GOPs) and its effects on employees in Turkish companies, 22 qualitative questions were asked to the respondent companies. The questions are classified into two categories. While the first part is about general knowledge about the process of certification, challenges, budgets etc.; the second part focuses on the effects of these practices on employees and organization. The questionnaire was developed by the authors based on the literature review about green offices and their organizational outcomes.

Within the scope of the research, companies which have had "WWF certificates were tried to be reached via e-mail and phone. The in-depth interview appointments were taken. The appointments were conducted via online platforms such as Teams, Zoom, etc. Lists of companies which have certificates were taken from the websites of WWF.

First of all. 9 not-for profit companies were eliminated from the list of researchers to contact for in-depth interviews. Also, companies which are in the list but do not keep their certificates anymore were excluded from the list of researchers. Researchers could not reach 10 companies although they tried to call or e-mail them. 7 companies refused to make in-depth interviews.

Researchers reached 27 companies in total and interviewed with 6 of them. In-depth interviews were almost 1 hour long for each company. When the interviews were made, researchers wrote what was talked during the interviews and send them back to the companies to confirm what was written and get their pemissions for using the information for the research.

Interview Questions

The first part of the interview which focuses on understanding the general green office concept and gathering information include these questions:

1)When did you start making green office practices?
2)What are your reasons for starting green office practices?
3)Which of the following arrangements have you made within the framework of green office practices? Mark it with an X.

Air Conditioning Quality (Heating/Cooling)	
Electricity/Energy Savings	
Saving Water	
Using Sustainable Office Materials	
Transportation Facilities (E.g. Usage of Electric-Hybrid Vehicle)	
Strategic Plan	
Strategies	
Other (Please Write)	

4) Have you received a certificate for your green office practices?
4.1) If so, what arrangements did you make for it?
5) How much budget did you allocate for green office practices?
6) Have your costs decreased since you started to implement green office practices? If yes, what is the estimated percentage?
7) What kind of process did you carry out while making green office practices?
8) How long did the process take?
9) Did you encounter any difficulties in the process? How did you overcome these problems?
From Questions 10 to 14, the aim was to understand the effects of green offices on satisfaction and wellbeing of the employees. For this purpose these questions are asked:
10) Have your green office practices increased the satisfaction of your employees related to the environment in the workplace?
11) Have your practices increased the comfort of the employees?
12) Have your green office practices increased the overall job satisfaction of your employees?
13) Have your green office practices had an effect on psychological well-being and mental health of your employees'? (e.g.: decreased stress levels, feeling determined, coping with the problems they encounter, etc.)
14) Have your green office practices had an effect on physical well-being and health of your employees? (e.g.: reduction of headaches, reduction of sleep problems, reduction of fatigue, reduction of tension, etc.) ***
From Questions 15 to 17, questions are asked to understand the effect of green office practies on organizational commitment.
15) Have your green office practices affected the emotional commitment of your employees towards the company positively?
16) Have your green office practices affected commitment of your employees to continue to work for the company positively?
17) Have your green office practices affected the normative commitment of your employees towards the company positively? *Normative Commitment: The employee feels responsible to the company and development of thought to say "I need to work for this company."
Questions between 18 and 20 were asked to understand the effects of green offices on job performance of employees and organizational performance.
18) Have your green office practices and the changing internal environmental conditions of your office increased the productivity of your employees?

19) Have your green office practices and the changing internal environmental conditions of your office reduced the absenteeism of your employees?
20) Have your green office practices increased the strategic performance of your business? If yes, in which of the following dimensions did it increase? Please explain briefly.

Financial Performance	
Customer Performance	
Performance for Business Processes	
Performance for Innovation and Development	

The last two questions were about to understand the intent of companies to continue to the certification programme and thoughts of specialists about the situation after pandemic.
21) Do you intend to continue to green office practices?
22) Do you think your green office practices will be a positive factor for attracting your employees to the office during and after the pandemic? Do you think your green office practices can be an alternative to reduce the remote work of your employees (increase working in the office)?

3.2. Findings

In total, researchers conducted 6 structured in-depth interviews with specialists and directors who were responsible for the green office certification processes and practices for Green Office of WWF. The interviews were made in 2022. The results of the structured in-depth interviews are given in details in the following pages. A qualitative research is used for this study to gather appropriate and comprehensive information about the subject.

Answers 1-9:

In this part of the research form; researchers asked the companies general questions about their certification processes, their reasons to start the certification process and benefits and challenges of having the certificates. Six respondent companies have WWF Green Office Certificates. For all of the companies, the process for taking the certificate lasts 1-3 years. Most of the companies got the certificate for their headquarters. On the other hand, one of them got it for its production facility and the other one got it for both its headquarter and some of the branches.
The answers to the questions reveal that the green office practices (GOPs) are often considered by companies as a part of their sustainability approaches. All of the companies stated that they started their practices with this vision. For example, **Company 1** states that; *"This is not just GOPs, but a work we do as a part of our sustainability program. In fact, it is (GOPs) one of the work we have done under the title of sustainability. We can basically say that our aim is to be a sustainable company."*
Also **Company 2** mentions that it is setting green office standards for its sustainability practices such as adapting climate changes, reducing its carbon footprint and deforestation. Like the two mentioned companies, **Company 3** also indicates that its reason to start the GOPs is being part of its sustainability vision. It explaines that; *"Our starting point is not the cost but having the clean World to leave to our children, we will leave it to the future, as employees of our company, as partners of our company as much as we can. We wanted this awareness to rise not only on the management side, but also on my last colleage who works at*

the bottom with the same feeling and the same responsibility. We really want to be environmentally conscious. Obviously, both as an individual and as a company."

Company 4 focuses on its environmental responsibility for starting the GOPs. It acknowledges that *"Our company is an environmentally friendly company. Probably, the subject of this diploma was started with our environmental responsibility."*

Company 5 mentions that GOAs are adding value to the environmental management system. It explains that *"In fact, we started by thinking that there would be a certification that would add value to our environmental management system."*

Lastly **Company 6** focuses on its holistic approach about GOA's and sustainability: *"We see sustainability in the middle of our business objectives. Our aim here is both to reduce our negative environmental impact causing from our own operations and increase our positive impact as a company with our employees, customers and business partners to the environment. I don't want to praise ourselves, but I don't think you can find a better company than us in terms of sustainability practices. Not only environmental sustainability, but also economic and social sustainability are our focus areas, and we continue our activities as a part of our strategy. I can say that our work is actually an indicator of this. For example, the issue of women is one of the issues which we prioritize.*

It is obvious that one of the important issues in the process of being a green office company is to increase the awareness of employees about these practices, change their perceptions to the concept positively and ensure that communication of these are made. **Company 3** is one of the companies which focuses on this issue. It explains that; "*We have given our own training. We started simple. We used small stickers and information notes such as -if you are going to open the window, you have to turn off the air conditioner. We talked about the employees gradually. We have to do this, because this is our future, entrustment of our children to us. Of course, first starts are not easy. Because people have other purposes. We all live in this environment, so we explained that we should all be sensitive. But it was very successful in the end in my opinion."*

Company 4 states that WWF has two lists for companies which apply for the certification: Watch list and action list. Awareness raising and empowerment are some of the main issues on the action list. *"WWF sent a watch list. For this watch list; we collected data for oır water consumption, natural resource consumption, and waste amount. They also sent an action list. For the action list, we completed 77 of the actions they determined. There are main groups on the list. Awareness raising and empowerment, energy saving, electricity, water, paper, waste management, reduction of single-use plastics, office supplies and fuel saving are the main issues on the list. We completed 77 of them. Some of them were removing rain water, using pet glasses and water bottles, sending food waste to animal shelters, obtaining zero waste certificate, aerator application, and participation in awareness raising activities."*

Company 6 also emphasizes the awareness of employees for GOA's success: *"We encountered some difficulties in the process. For example, plastics may be a priority area for the company in the context of sustainability, but it may not be a priority area for the employee. When he goes to the water dispenser, he may want to see a plastic cup. Or he may not eat his food in the cafeteria. He can go upstairs or eat somewhere in the garden with plastic cutlery. Therefore, this is an action which makes his job difficult. There has been a resistance among our employees in this regard. But only 5-10% of empoyees resist to practices which protect the environment. Apart from that, there is a very young and dynamic generation in our company* *right now. There is a generation that we can call the Z generation. All of them supported us in this regard and took very serious actions. Therefore, when I evaluated it in a transparent way, we had more supportive employees."*

The starting year, process duration, type of the certificate and the year it was taken, certificied office locations and main GOPs by each company are presented in Table 2.

Table 2. Green Office Certification Process

	Starting Year	Process Duration	Certificate	Location	Practices
Company 1	2013	1 Year	WWF-2014	Headquarter and Some Branches	Assign Boards (WWF Panda's) Paper Usage Light Usage Electricity Consumption Deductor Faucet With Armature Hybrid Vehicles
Company 2	2012	3 Years	WWF-2015	Headquarter	Sound Absorbers Led Illumination Facade Lighting Projectors Carbon Monoxide and Moisture Censors Carbonfiltered Poddle Boxes Trigeneration Systems
Company 3	2017	14 Months	WWF-2018	Headquarter	Awareness Rising (Using Stickers) Disposing of the Materials Water Taps with Sensors Reuse of Boxes in Warehouses

Company 4	2011	5 Years	WWF-2016	Headquarter	Removing Rain Water Sending Food Waste to Animal Shelter Zero Waste Management Aerator Application Awareness Raising Activities
Company 5	2016	1 Year	WWF-2016	Gebze Facility	Action Plan Digital Thermometer Heating System Sweater Day Raising Awareness Water Taps with Sensors Zero Waste Management Sustainable Office Materials Disposing of Wastes
Company 6	2012	1 Year	WWF-2012	Headquarter	Sweater Day Raising Awareness Recycling Water Taps with Sensors Sustainable Office Materials Zero Waste Management Hybrid Vehicles Led Illumunation

An environmentally friendly company can make contributions to other companies economically by providing positive developments in their consumptions. Also, some companies use these practices for their communications to their stakeholders. In this sense, some companies had measured these contributions. For example **Company 1** is one of the companies which measured its efforts and outputs:
"When we compare the total carbon footprint of our natural gas, electricity and paper consumption in 2013 with 2020, we see a reduction of around 45-46%. We have reported these reductions nowadays. Frankly, we do not use this data in advertising but we use it in communication channels. For example, we used it in our annual report before, but now we publish it separately in our sustainability report. Secondly, we share this on the social media accounts of our holding."
Company 3 and **Company 5** also have measured their savings. Their exact numbers are on Table 2. **Company 5** believes that the main focus of these GOPs is increasing employee awareness: *"We can't give an estimated percentage of each cost. Because while we are calculating our savings, such as electricity and natural gas in general, the production activities of the factory are also included in the total amount. We consider the WWF certificate purely as a tool to increase the awareness of our employees. In other words, we looked at it as the satisfaction we will experience through the effect of the savings we will provide on the environment. But of course it affected our consumption. Some improvements related with natural gas and electricity were made."*
About the challenges in the process, some companies believed that they did not face serious challenges in their efforts. **Company 4** believes that it is about its culture. It stated: *"We did not encounter any difficulties. As we have always adopted continuous improvement as a company culture, we did not find it difficult."*
Company 2 attributes this to its external consultancy support: "*We received support from the relevant consultants so that the process would succeed. While the process was being carried out, we received support from expert consultants, even during the construction phase. If there are periodic updates, we do them, but if there is something extra, we may get consultancy."*
Company 1 believes that employee turnover, market conditions and workload of employees are some of the most challenging issues. *"We had sustainability and WWF ambassadors in the volunteer-based business. Sometimes we sacrifice them to turnover. There were 3 people left in the list, we lost most of them. There can be a challenge when there are market requirements, workload, other priorities, the departure of the employee make it difficult to proceed the practices"*
Company 3, **Company 5** and **Company 6** stated that the most difficult challenge they faced was to change the perceptions of the employees and ensuring their participations to the GOPs. Company 3 even focuses on it in its job interviews: *"It's not easy to tell people. They understand, but when it comes to the implementation, there are deficiencies. We held more information meetings, trainings and on-site warnings to overcome them. Also, new staff are coming. There is circulation, not just the staff here. You have to tell them too. Let me be very clear, I am trying to understand whether the candidate is environmentally friendly or not in my job interviews. The important thing is what will you add here as a human being These are very important to us."*
Company 5: "*There were some difficult processes in terms of perception. Because there are some things that are related to employee take off and reduce office comfort. There may be*

difficulties with this at first, but we cannot say that we had an important difficulty. There were minor resistances when many of our practices reduce personal comforts of our employees. But since we are a company with an energy management system certificate, we did not find it very challenging."

Benefits and challenges of the process for each company are presented on Table 3.

Table 3. Benefits and Challenges of Being a Green Office Company

	Benefits	***Challenges***
Company 1	45% Reduction in total Carbon footprint Commercial Benefits (Social Media and Sustainability Report)	Employee Engagement and Awareness Market Conditions and Workload
Company 2	Paper Cups Saving 41% Energy Saving 7-10% Plastic and Water Savings	No Spesific Challenge
Company 3	Water Saving 20% Electricity Saving 10% Paper Cups Saving 25%	Raising Employee Awareness and Ensuring the Implementation
Company 4	Effect on Costs Electricity Consumption	No Spesific Challenge
Company 5	Increase in Employee Awareness Satisfaction 14% Reduction in Electricity Consumption per Employee 24% Saving in Water Consumption Per Employee 32% Reduction in Natural Gas	Employee Perception Office Comfort
Company 6	Paper Cups Saving 41% Energy Saving 7-10%	Employee Perception

Green office practices of companies are found out to be certified as a green office company. Their answers are presented on Table 4.

Table 4. Green Office Practices of Companies

	Air Conditioning Quality	**Electricity/ Energy**	**Water**	**Office Materials**	**Transportation**	**Strategic Plan**	**Strategies**
COMPANY 1	Partially	Yes	Yes	Yes	Yes	Yes	Yes
COMPANY 2	Yes	Yes	Yes	No	Yes	Yes	Yes
COMPANY 3	Yes	Yes	Yes	Yes	No	Yes	Yes
COMPANY 4	Yes	Yes	Yes	No	For Managers	No	Yes
COMPANY 5	Yes	Yes	Yes	Partially	No	No	No
COMPANY 6	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Answers 10-14 (Satisfaction, Health and Wellbeing):
All of the companies observed that their practices had positive affects on satisfaction about office environment. **Company 1** acknowledges that "*We can say that receiving daylight has a positive effect on employees. In the previous structure, the manager took the corner, while the employees were working in the middle areas which were not exposed to the sun light. Now this work we have done to benefit from the sun has enabled our employees to work in a slightly more sustainable environment."*
Company 2: *"We think it increased. We can say that the most important point is the increase in the landscape area. An additional greenery and landscaping area of approximately 2000 m^2 came around our building and inside itself. Lighting and ventilation quality have improved. Waste management has started in our building, I think these are some of the elements which increase satisfaction."*
Company 3, 4, 5, and 6 also state that it has a positive effect. **Company 4** and **Company 6** underline importance of their "Sweaters Day" for their employees about the issue.
Company 4:*"We attend the Sweater Day every year. It is a day of WWF where air conditioners are fixed at 25 degrees all over the world, everyone is dressed thickly, and energy consumption is noted. We get feedback from employees on its benefits."*
Company 6: *"We started the world sweater day application for the first time in Turkey. It is organized every February. It was a practice implemented abroad by WWF at first, and then we started to implement it in Turkey. WWF does not implement it in Turkey. Various business partners and different companies continue to implement this. On this day, we turn off the air conditioners and ask people to come in sweaters but turning off the air conditioner doesn't mean we have to cool the inside in February. We provide an air quality of international standards there. In other words we are adjusting the air quality and temperature-coldness that should be in February."*

On comfort of the employees: All of the companies stated that their practices had positive effects. Company 2 underlined the value of air conditioners on that: ***"****We think the improvement efforts have increased. For example we changed the air conditioners in all locations. We have replaced old air conditioners with environmentally friendly gas-forming air conditioners. Also staff seats in all branches and buildings have become harmonious seats. We think it has increased the satisfaction in this direction."*

Company 4 has received positive feedbacks from its employees for its awareness rising activities.

Company 6: *"Applications such as areas designed with a specific concept, green areas and redesigning garden areas created a comfortable environment for our employees."*

About job satisfaction of the employees; all of the companies, except Company 3, believe it had some effects. For **Company 1** it has positive effects on mostly satisfaction of environmentally concerned employees *"Employees who follow how the company approaches environmental issues and whether it is sensitive or not may feel this way. They can be more satisfied. Do you have an approach to it? It already has a place in our organizational culture. I think it increases employee satisfaction."*

Company 2 says that it has received feedbacks from employees: **"***The studies, trainings, audits and subsequent feedbacks of our sustainable finance team and productivity teams are very positive. There is a satisfaction because our employees internalize the sustainability processes so much. This reaches us both by e-mails and feedbacks in trainings***."**

About physical healths of employees, companies underline the importance of office environment;

Company 1: "*They feel better in a slightly more relaxed environment, thanks to the more daylight we get and our garden."*

Company 2: "*Previously, non-LED lighting was a bit more tiring, we switched to LED lighting. Now when we turn it on and off at certain times, it obviously increased our performances.*"

Company 3: *"I think it's certain.(It affected positively) Because really, at least the seats they sit on and everything they use are ergonomic. How many companies have an indoor air cleaning system now? They only have air conditioners. The air conditioner does not clean the air."*

Company 6*: "We have an operational unit. Even though it is an open office, when you tell your desk number for the air conditioner and light on the table where you are sitting, technicians can intervene to set them however you want momentarily. Therefore, the employee can call that line and ask to set the applications on his own behalf. Of course, it also brings satisfaction."*

Opposite to them, Company 4 and Company 5 believe that GOAs did not affect the physical health of their employees that much, because their office environment already had enough oxygen, daylight etc.

Company 4: "*I don't think these practices have much physical effects. Since our company gives great importance to continuous improvement, these changes will not affect us physically. In order to have a positive effect; for example, your office may not have ventilation before, or there may not be oxygen in the environment before".*

Company 5: "*In fact, we did not have major radical changes in our offices. Air and light were already at a certain level, for example, we were already using the daylight. There were green plants. Small touches were actually made in the offices. For this reason, I can't say much about its effects, since there has not been a big change."*

In addition to that, about the changes in mental health of the employees, answers show that awareness of the company on these issues has an effect on its employees.

Company 2:*"According to the results of our survey, our bank appears to be ahead by a large margin on issues such as to be the most environmentally friendly bank and the greenest bank. This is something that you will appreciate that these efforts positively affected the psychology and motivation of employees"*

Company 4: "*Working in a green office certified company definitely creates a positive effect. The diploma shows that something has been done."*

Company 6: *"We organize "Spirit Day" here 4 times a year. There is no work on that day and various activities are organized. The employee can use that day for himself. We do these activities several times in the contexts of sustainability and green office. Activies are landscaping, crop rooting, home gardening etc. We see that their stress is reduced there and they feel good. Their motivation has also increased. These types of activities have become even more valuable for us, especially after the pandemic, and we get very good feedback from them".*

Answers 15-17 (Organizational Commitment):

All companies seem to agree that green office practices have positive effects on the emotional commitment of employees. Especially having the vision of sustainability have become important for employees for gaining an emotional commitment. Some notable explanations of the companies related to this can be seen below.

Company 1: *"The feeling of emotional commitment increases when the employees who give importance to these issues see that the company is also interested in these issues."*

Company 2: *"An institution with a sustainable perspective develops an emotional commitment from its employees, customers and other stakeholders. Now the climate has changed, the things we live during the day have changed. In the past, these issues were not discussed much, but our generation is aware of this. We noticed this by the feedbacks we received. The fact is our institution pays attention to these issues, and has an awareness. For example, the efforts we do at our headquarters and branches (regarding waste, we use recycling boxes on each floor and separate wastes for recycling) can be observed by everyone".*

Company 3: *"It sure did. Because we started by valuing them. We always told our employees that it is for their health and their children."*

Company 4: ***"****Yes it did. We can say that after our practice on that sweater day, it increased the pride of the institution. Just being a donor of the WWF is a good thing after all. It's something an employee can proudly say in his/her personal life outside work as well."*

Company 6: *"Definitely. I don't know how valuable it is for you, but there is an initiative called Great Place to Work. We received the award there. There are already several criterias here. So we can say that it proves emotional commitment."*

About continuance and normative commitment, companies are not sure that GOAs have or will have an effect on them. For continuance commitment, economic issues could be more important for the employees as **Company 4** said:*"We don't think that green office certification will have an effect on employees for not quiting their jobs. Realistically speaking, especially in our country."*

About continuance commitment, similar to Company 4, Company 6 stated that its turnover ratio is not affected by GOAs. According to the company, there has been too much circulation in workforces, especially during and after the pandemic.

Answers 18-20 (Performance):

On employee performance:
The answers show that any of the companies in the research measured the effects of the GOAs on **employees performance (productivity and absenteeism). Company 1, 2, and 6** directly explain that they have not measured anything related to that. Even though there is no quantitave data, **Company 3** strongly believes that its GOAs have had positive effects on them: *"The seats of our personnel are the seats of managers in other companies. They're not sitting on normal revolving office chairs. It is very important. Or the mousepads we use have gels. They are not harsh. They relax the wrists pretty much. These are all details, but they are important for us. Imagine working in unhealthy conditions all day. The interior may be beautiful, but what matters is what materials employees use. This is much more important for us. About absenteeism, these practices have reduced it. Why? Let me give you an example, one of my employee was pregnant and had a child recently. But she has never worked from home. Because she said she was more comfortable here. When she comes here, of course, there is a healthy environment. As we said, air conditioning is extremely healthy. The seat where she sits on while she is working is comfortable. That's why everyone is happy to come and work here. For example, during the pandemic, I never worked from home. I tried to work from a home. One afternoon I went and repented not to do it again. These practices reduced absenteeism."*
Although **Company 3** is very optimistic about the effect.
On company performance:
In general, answers indicate that being a green office has positive effects on company performance in various areas. For example companies have reached a concensus about effects of GOPs on financial and customer oriented performances. All of the 6 companies believe that GOPs which they made had increased their financial and customer-oriented performances.
Company 6 has measured its financial performance and customer-oriented performance with quantitative data, through reputation surveys which are conducted by an independent research institution. However, since it is a global and quite large company, it thinks that this is the result of what it does under the umbrella of sustainability, rather than it is directly associated with the GOAs.
Company 1 is confident that green office practices have affected its financial success through its measurements. Samewise, **Company 3** says that: *"Our financial performance definetly increased. When we became a green company, we also made our boss happy by reducing our costs."* **Company 4** underlines the cost savings of energy, electricity and water as an increase in financial performance.
About customer-oriented performance; **Company 2** has received positive feedbacks from its customers from its social media accounts. **Company 3** also uses the certificate for its communication, marketing and sales activities. *"We use it as a sales argument in our meetings we ıse it for our suppliers or potential customers. -Be sure you emphasize these practices!- what I especially say to my friends. Because my competitors do not have these."* **Company 4** also uses them in its communication with its clients. It said: *"It also affects our customer oriented performance because we are telling our clients that we have a green office certificate. We're promoting our practices. It's a positive image for our company."*
Like **Company 3 and 5**; **Company 5** also uses its GOPs in its marketing activities: *"Our marketing department made the external communication of these in various bulletins and*

	Financial Performance	Customer-Oriented Performance	Business Process	Innovation and Growth
COMPANY 1	Yes	Yes	Notr	Yes
COMPANY 2	Yes	Yes	-	Yes
COMPANY 3	Yes	Yes	Yes	Yes
COMPANY 4	Yes	Yes	-	-
COMPANY 5	Yes	Yes	-	Yes
COMPANY 6	Yes	Yes	Yes	Yes
PERCENTAGE	100%	100%	33,3%	83,3%

industry magazines. We produce our products in factories with green offices. Thus, it is announced to our customers. Also our name is on the website of WWF."

There are different thoughts about the effects of GOAs on the company performance related to business processes. For a positive example **Company 6** stated that GOAs in its business performance and digitalization operation to save paper accelerate its business processes. Company 4 and 6 agreed on their positive effects on innovation and growth performance.

Answers 21-22:

Table 5 summarizes the effects of GOPs on different dimensions of the company performance.

Table 5. Green Office Effects on Company Performance

Table 6. Intentions for Continuity and Becoming An Alternative to Remote Work of Green Office Practices

Company	Intend to Continue	Green Office Practices as an Alternative to Remote Work
1	Yes (With Expanding Its Scope)	"It might be. Especially when you think about Istanbul's traffic. Days pass with the fear of how I will go to the office."
2	Yes (With Expanding Its Scope)	"In the past weeks, our department assistant came to the office on a day when he was not supposed to come. He cited using our waste battery box as a reason to come. We also have a blue cap project with the Barrier-Free Living Association. We even have some colleagues who come to the office to bring those caps. These examples may be the answer to your question."
3	Yes (Always)	"Sure. I honestly believe it is. If the staff comes to the office happily, believe me, they will prefer it more than their homes."
4	Yes (Our new investments will focus on	"It's possible. I'm not so sure right now if it's related, but it could be."

	the environment)	
5	Yes (It will continue in the future as well)	"Few of our teams are involved in remote work but I can't say that it is directly associated with the green office concept. We come to office because we can do some of our work more quickly face to face, and we prefer to be at the factory for the necessities of the job. Is it an alternative for working in the office? I don't think that it's entirely related to the issue."
6	Yes	"Yes. Especially activities, workshops we do in a Spirit Day, etc. These are subjects that our employees particularly demand and are interested in. We see that the interest of people for the environment has increased during the pandemic, I can say this is true for our employees as well. It may not be very accurate to say for the other companies. We think that it increases the commitment to the office here. But I don't know if our company wants this strategically."

4. Discussion and Conclusion

As it is observed in this research, green office practices affect employee well being, job satisfaction, organizational commitment, and company performance. However, due to pandemia conditions, it is not considered as an alternative for remote work practices.

Requirements of WWF is not very strict. Companies which have environmental awareness can get this certificate and make practices even beyond its requirements. Since most employees don't come to work frequently in most companies during pandemia, employees can show their environmental awareness in their personal lives as well with the encouragement of their companies via websites, electronic bulletins, etc.

However, there are still few companies which have WWF certificates in Turkey. The number of certified companies and their efforts to be more environmentally responsible should be increased. First of all, companies can hire employees who have environmental awareness, provide training to them to increase their enviromental awareness and encourage them to make practices to save the environment. Job descriptions, performance appraisal criterias, reward system, organizational culture should be designed to lead employees to be environmentally responsible. Managers should be role models to employees with their behaviors, approaches and practices.

Green offices affect employee well being, job satisfaction, organizational commitment, and company performance. Thus, they can increase overall performances of companies. Being certified has reasonable costs and it will help companies to make contributions to protect the environment. There is only one World and it is borrowed from next generations. So, companies should have responsible approaches to have the appropriate infrastructures and practices to save the World.

References

Agha-Hossein, M. M., El-Jouzi, S., Elmualim, A. A., Ellis, J., & Williams, M. (2013). Post-occupancy studies of an office environment: energy performance and occupants' satisfaction. Building and Environment, 69, 121-130. doi: 10.1016/j.buildenv.2013.08.003

Ahmed, M., & Shafiq, S. (2014). The impact of organizational culture on organizational performance: a case study on telecom sector. Global Journal of Management and Business Research. A Administration and Management, 14(3), 21-30.

Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance, and normative commitments to the organization. Journal of Occupational Psychology, 63, 1-8.

Allen, N. J. and Meyer, J. P. (1991). A three-component conceptualization of organizational commitment. Human Resource Management Review, 1, 61-89.

Ali A.S., Chua, S. J. L., & Lim Melissa, E. L. (2015). The effect of physical environment comfort on employees' performance in office buildings. Structural Survey, 33, 294-308.

Armstrong, M. (2006). A Handbook of Human Resource Management Practice. 10th Edition. London: Kogan Page Publishing.

Aziri, B. (2011). Job satisfaction: A literature review. Management Research and Practice, 3(4), 77-86.

Bakke, D. W. (2005). Joy at Work: A Revolutionary Approach to Fun on the Job. New York, NY: Penguin Group (USA) Inc.

Bangwal, D., Tiwari, P., & Chamola, P. (2017). Workplace design features, job satisfaction, and organization commitment. Journal of Workplace Rights, SAGE Open, 7(3), 1-12. Doi: 2158244017716708.

Baptiste, N. R. (2008). Tightening the link between employee wellbeing at work and performance: A new dimension for HRM. Management Decision, 46(2), 284-309.

Barker, R. G. (1968). Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior. Stanford, CA: Stanford University Press.

Becker, H. S. (1960). Notes on the concept of commitment. American Journal of Sociology, 66(1), 32-40.

Berberoglu, A. (2018). Impact of organizational climate on organizational commitment and perceived organizational performance: empirical evidence from public hospitals. BMC Health Services Research, 18(1), 1-9.

Bishop, J. (1987). The Recognition and Reward of Employee Performance. Journal of Labor Economics, 5(4), 36-56.

Brager, G., & Baker, L. (2009). Occupant satisfaction in mixedmode buildings. Building Research and Information, 37, 369-380.

Brazdauskas, M., & Žirnelė, L. (2020). Creating green offices: Employee engagement perspective. Journal of Creativity and Business Innovation, 6, 62-69.

Brown, Z., Cole, R. J., Robinson, J., Dowlatabadi, H. (2010). Evaluating user experience in green buildings in relation to workplace culture and context. Facilities, 28(3/4), 225-238. http://dx.doi.org/10.1108/02632771011023168

Carlopio, J. R. (1996). Construct validity of a physical work environment satisfaction questionnaire. Journal of Occupational Health Psychology, 1, 330-344.

Chavan, M. (2009). The balanced scorecard: a new challenge. The balanced scorecard: A new challenge. Journal of Management Development, 28(5), 393-406.

Christiansen B, Chandan HC (2017). Handbook of Research on Human Factors in Contemporary Workforce Development. Hershey, PA, USA: IGI Global.
Coşkun, A., & Akar, N. (2019). "WWF-Turkey's green office initiatives" in Green Behavior and Corporate Social Responsibility in Asia. (Emerald Publishing Limited), 47-55.
Cooper, C. L., Dewe, P. J., & O'Driscoll, M. P. (2001), Organizational Stress: A Review and Critique of Theory, Research and Application, Sage, Newbury Park, CA.
Currie, D. (2001). Managing Employee Well-Being. Oxford: Chandos Publishing (Oxford) Limited.
Davis, K. and Nestrom, J. W. (1985). Human Behavior at Work: Organizational Behavior. 7 Edition. New York: McGraw Hill.
Deuble, M. P., & de Dear, R. J. (2012). Green occupants for green buildings: The missing link. Buildings and Environment, 5, 21-27.
Doggart, J. (2006). "Future design-guidelines and tools," in Creating the Productive Workplace, ed. D. Clements-Croome, (E & FN Spon), 378-386.
Dole, C., & Schroeder, R. G. (2001). The impact of various factors on the personality, job satisfaction and turnover intentions of professional accountants. Managerial Auditing Journal, 16, 234-245.
Elnaklah, R., Fosas, D., & Natarajan, S. (2020). Indoor environment quality and work performance in "green" office buildings in the Middle East. Building Simulation, 13(5), 1043-1062.
EPA (2019). United States Environmental Protection Agency. Available at https://www.epa.gov/.
Feige, A., Wallbaum, H., Janser, M., & Windlinger, L. (2013). Impact of sustainable office buildings on occupant's comfort and productivity. Journal of Corporate Real Estate, 15(1), 7-34.
Fisk, W. J. (2000). Health and productivity gains from better indoor environments and their relationship with building energy efficiency. Annual Review of Energy Environment, 25, 537-566. doi: 10.1146/annurev.energy.25.1.537
Frontczak, M., Schiavon, S., Goins, J., Arens, E., Zhang, H., & Wargocki, P. (2012). Quantitative relationships between occupant satisfaction and satisfaction aspects of indoor environmental quality and building design. Indoor Air, 22, 119-131. doi:10.1111/j.1600-0668.2011.00745.x
George, J. M. and Jones, G. R. (2008). Understanding and Managing Organizational Behavior. Fifth Edition. New Jersey: Pearson/Prentice Hall.
Ghodrati, N., Samari, M., & Shafiei, M. W. M. (2012). Green buildings impacts on occupants' health and productivity. Journal of Applied Sciences Research, 8(8), 4235-4241.
Hackett, R. D., Bycio, P., & Hausdorf, P. A. (1994). Further assessments of Meyer and Allen's (1991) three-component model of organizational commitment. Journal of Applied Psychology, 79(1), 15-23.
Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: test of a theory. Organizational Behavior and Human Performance, 16, 250-279. doi: 10.1016/0030-5073(76)90016-7
Hoppock, R. (1935). Job Satisfaction: New York: Harper and Brothers.
Iverson, R. D., & Buttigieg, D. M. (1999). Affective, normative and continuance commitment: Can the `Right Kind' of Commitment Be Managed? Journal of Management Studies, 36(3),

307-333.
Jain, A. K., Giga, S. I., & Cooper, C. L. (2009). Employee wellbeing, control and organizational commitment. Leadership and Organization Development Journal, 30(3), 256-273.
Jex, S. M. (1998). Stress and Job Performance: Theory, Research, and Application. Newbury Park, CA: Sage.
Kaliski, B. S. (2007). Encyclopedia of Business and Finance. Second edition, Detroit: Thompson Gale.
Kaplan, S. (1983). A model of person-environment compatibility. Environment and Behavior, 15, 311-332. doi: 10.1177/0013916583153003
Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard--measures that drive performance. Harvard Business Review, January-February, 71-79.
Kim, J., & de Dear, R. (2013). Workspace satisfaction: The privacy-communication trade-off in open-plan offices. Journal of Environmental Psychology, 36, 18-26. doi:10. 1016/j. jenvp.2013.06.007
Kozusznik, M. W., Maricutoiu, L. P., Peiró, J. M., Vîrga, D.M., Soriano, A., & Mateo-Cecilia, C. (2019). Decoupling office energy efficiency from employees' well-being and performance: A systematic review. Frontiers in Psychology, 10, 1-24.
Le Fevre, M., Matheny, J., & Kolt, G. S. (2003). Eustress, distress, and interpretation in occupational stress. Journal of Managerial Psychology, 18(7), 726-744.
Leaman, A. (1995). Dissatisfaction and office productivity. Facilities, 13, 13-19.
Leder, S., Newsham, G. R., Veitch, J. A., Mancini, S., & Charles, K. E. (2016). Effects of office environment on employee satisfaction: A new analysis. Building Research and Information, 44, 34-50.
Liquna, G. and G. Yanqunb, 2011. Study on building materials and indoor pollution. International Conference on Green Buildings and Sustainable Cities. Procedia Engineering, 21, 789-794.
Mathieu, J. E., & Zajac, D. M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. Psychological Bulletin, 108(2), 171-194.
Meyer, J. P., & Allen, N. J. (1984). Testing the ``side-bet'' theory of organizational commitment: Some methodological considerations. Journal of Applied Psychology, 69, 372-378.
Meyer J. P., and Allen N.J. (1991). A three-component conceptualization of organizational commitment. Human Resources Management Review, 1(1), 61 -89.
Meyer, J. P., Paunonen, S. V., Gellatly, I. R., Goffin, R. D., & Jackson, D. N. (1989). Organizational commitment and job performance: it's the nature of the commitment that counts. Journal of Applied Psychology, 74, 152-156.
Meyer, J. P., Allen, N. J., & Gellatly, I. R. (1990). Affective and continuance commitment to the organization: evaluation of measures and analysis of concurrent and time-lagged relations. Journal of Applied Psychology, 75, 710-720.
Meyer, J. P., Allen, N. J., & Smith, C. A. (1993). Commitment to organizations and occupations: extension and test of a three-component conceptualization. Journal of Applied Psychology, 78, 4, 538-551.
Monfared, I. G., & Sharples, C. (2011). Occupants' perceptions and expectations of a green office building: A longitudinal case study. Architectural Science Review, 54, 355-344.

Morrow, P. C. (1983). Concept redundancy in organizational research: The case of work commitment. Academy of Management Review, 8(3), 486-500.
Nagy, Z., Yong, F. Y., Frei, M., & Schlueter, A. (2015). Occupant centered lighting control for comfort and energy efficient building operation. Energy and Buildings, 94, 100-108. doi: 10.1016/j.enbuild.2015.02.053
Newsham, G. R., Birt, B. J., Arsenault, C., Thompson, A. L., Veitch, J. A., Mancini, S., et al. (2013). Do 'green' buildings have better indoor environments? New evidence. Building Research and Information, 41, 415-434. doi: 10.1080/09613218.2013. 789951
Nurick, S., & Thatcher, A. (2021). The Relationship of Green Office Buildings to Occupant Productivity and Organizational Performance: A Literature Review. Journal of Real Estate Literature, 29(1), 18-42.
Ornetzeder, M., Wicher, M., & Suschek-berger, J. (2016). User satisfaction and well-being in energy efficient office buildings: Evidence from cutting-edge projects in Austria. Energy and Buildings, 118, 18-26. doi: 10.1016/j.enbuild.2016.02.036
Porter, L. W., Steers, R. M., & Boulian PV. (1974) Organizational commitment, job satisfaction and turnover among psychiatric technicians. Journal of Applied Psychology, 59, 603-609.
Povodör, A., Bodroghelyi, C., & Mouazan, E. (2010). European Green Office Handbook (K. A. f. S. E. (Hungary) Ed., Vol. 510859-LLP-1-2010-1-HU-LEONARDO-LNW.
Rauatmaa, H. (2011). Organisations' Satisfaction with the Green Office Environmental Management System By WWF Finland. Oulu University of Applied Sciences, Bacheor Thesis, Autumn, 1-74.
Reichers, A. E. (1985). A review and reconceptualization of organizational commitment. Academy of Management Review, 10, 465-476.
Schiavon, S., & Altomonte, S. (2014) Influence of factors unrelated to environmental quality on occupant satisfaction in LEED and non-LEED certified buildings. Building and Environment, 77, 148-159.

Seppa¨nen, O., & Fisk, W. J. (2006). "A procedure to estimate the cost-effectiveness of the indoor environment improvements in office work" in Creating the Productive Workplace, ed. D. Clements-Croome, (E & FN Spon), 407-433.
Shahzadi, I., Javed, A., Pirzada, S. S., Nasreen, S., & Khanam, F. (2014). Impact of employee motivation on employee performance. European Journal of Business and Management, 6(23), 159-166.
Sirviö, M. (2011). Green Office expert, WWF Finland. Email interview 14.6.2011 & 10.10.2011.
Statt, D. (2004). The Routledge Dictionary of Business Management. 3rd Edition. Detroit: Routledge Publishing.
Varanavicius, V., & Navikaite, A. (2015). Competitive advantage attainment via synergy in green offices. Journal of International Studies, 8(3), 141-151.
Vischer, J. C. (2007). The effects of the physical environment on job performance: towards a theoretical model of workspace stress. Stress Health, 23, 175-184. doi: 10.1002/smi.1134
Vischer, J. C. (2008). Towards an environmental psychology of workspace: How people are affected by environments for work. Architectural Science Review, 51(2), 97-108.
Weiner, Y. (1982). Commitment in organizations: A normative view. Academy of Management Review, 7, 418-428.

Wicker, A. W. (1973). Undermanning theory and research: Implications for the study of psychological and behavioral effects of excess populations. Representative Research in Social Psychology, 4(1), 185-206.

WWF Finland. (2011). Green Office GoNet. Internal source. http://www.wwf.fi/green_office/.

WWF International. (2011). Green Office. http://wwf.panda.org/what_ we_do/ where_we_work/greatermekong/our_solutions/ projects/green_ office/.

WWF Turkey, https://www.wwf.org.tr/sizneyapabilirsiniz1/isinize_dogayi_katin/yesilofisolun/ sistemnasilisliyor/. Retrieved Date: 2/23/2022.

Yang, H. (2008), Efficiency Wages and Subjective Performance Pay. Economic Inquiry, 46(2), 179-196.

Yazıcı, N. K. (2008). The Effect of Reward System Applications on Employee Performance in Service Sector. Marmara University, Institute of Social Sciences, Master Thesis.

Zou, P. X. W., & Couani, P. (2012). Managing risks in green building supply chain. Architectural Engineering and Design Management, 8(2), 143-158. //http://web.ebscohost.com/ehost/pdfviewer/ pdfviewer?vid=4&hid=19&sid= 27a3f53a-1b22-4329-9e1b-4d53a91489de%40sessionmgr10

17.

Transparency and Corporate Social Responsibility in English Soccer: A Case Study of Fulham FC

Muhammad Umar Farooq

New Mexico State University, Las Cruces, USA

Abstract

This paper uses a qualitative case study method to look at how a soccer team in the English Premier League, Fulham Football Club (Fulham FC) has utilized transparency, to showcase its past corporate social responsibility (CSR) performance and future goals, via its charitable foundation, the Fulham FC Foundation. I discuss the benefits of a focused corporate social responsibility strategy for a soccer team. This article draws on past research on professional sports teams' CSR practices and looks to contribute to the literature on CSR in sports, transparency, and strategy.

Keywords: Corporate social responsibility, transparency, English Premier League, Fulham FC, Fulham FC Foundation

Categories

Sustainability

44.

Sustainability Strategies: Uncovering Effective Institutional Pressures as Drivers of TBL Results

Alina Marculetiu [1], Cigdem Ataseven [2], Injazz Chen [3]

[1]Youngstown State University, Youngstown, USA. [2]Cleveland State University, Cleveland, USA. [3]Cleveland State Univeristy, Cleveland, USA

Abstract

Our research delves into the effect of institutional pressures on sustainability strategies, differentiating between symbolic and substantive strategies. We assess the direct impact of sustainability strategies on the triple bottom line (TBL) and the indirect effect of institutional pressures on the TBL. We use a unique sample of 388 responses from supply chain managers and executives and test our hypotheses using structural equation modeling. Our findings reveal that genuine substantive sustainability strategies are driven by norms embedded within organizations (i.e., normative pressures) and that such strategies fully mediate the relationships between normative institutional pressures and all TBL performance measures. In addition, this research provides potentially powerful organizations with information that can ripple persuasion upstream the supply chain and contribute to environmental and social stewardship. Overall, this study helps institutions identify successful sustainability strategies and provides insights for pressuring organizations to understand drivers that encourage participation in sustainable practices.

Categories

Sustainability

53.

Financial Assets, Inventory Levels, and the Impact of Disruptions

Jiangxia Liu [1], Sanjay Kumar [2]

[1]Valparaiso University, Valparaiso, USA. [2]VALPARAISO, VALPARAISO, USA

Abstract

Company managers maintain different levels of cash and inventory. Cash can help a company acquire critical resources needed to mitigate disruptions. Similarly, inventory can help full too. Extant literature indicates that buffer inventory and cash assets help to mitigate the effects of supply chain disruptions. We categorically examine the role of cash and inventory in the event of a supply chain disruption.

Categories

Supply Chain Resilience

60.

Economic and Reliability Analyses of Rainwater Harvesting Programs for rural cities.

Vinay Gonela [1], Raghavan Srinivasan [2]

[1]Texas A&M University – Central Texas, Killeen, USA. [2]Ball State University, Muncie, USA

Abstract

This research focuses on performing economic and reliability analyses of rainwater harvesting programs for rural cities. An integrated simulation and optimization-based approach is proposed in which: (1) local-level optimal household rain barrel capacity is determined by using Monte Carlo simulation; and (2) global city-level rainwater collection potential is assessed by using optimization model. A case study of a rural city in the Central Texas Region of US is used as an application of the proposed model and to gain important managerial insights. The results indicate that on an average, a reliability of 29.95% and a savings of $ 4545.30 per annum can be achieved for the entire city through rainwater harvesting. In addition, it is observed that monetary incentives are necessary when rainwater collection is primary goal compared to monetary savings from rainwater harvesting.

Categories

Sustainability

64.

Evaluating the long-term cost performance of Green Residential Buildings from Users' Points of View

Mohsen Goodarzi

Ball State University, Muncie, USA

Abstract

One of the most crucial aspects of long-term success is the cost performance from end-users' points of view as they are the only stakeholder encountering both short-term and long-term costs of the building. This study investigates the post-occupancy cost performance of green residential buildings from the users' perspective to determine the most influential economic determinants of users' satisfaction. Several economic aspects were selected from the existing literature, to be evaluated as the potential determinants of long-term success. A survey was conducted to collect data from residents of green buildings through a structured questionnaire. A total of 192 responses were collected and after conducting a confirmatory factor analysis (CFA) and validating the data, a multiple regression analysis was conducted to understand the relationships between the perceived post-occupancy cost performance of green buildings and the level of satisfaction of residents, as the determinant of long-term project success. The results showed that home value/rent has the highest significant influence on the satisfaction of the residents followed by utility bills and travel costs. On the other hand, the fees associated with living in green buildings such as HOA/Condo fees, property tax, maintenance fees, and other miscellaneous fees did not have any influence on the residential satisfaction of the green building residents. The findings of this study are beneficial for researchers and practitioners who focus on the development of sustainable residential buildings.

Categories

Sustainability

Teaching, Innovation, Curriculum & Cases

21. CONTEXTUALIZING OPERATIONS MANAGEMENT CONCEPTS THROUGH WEB-BASED SIMULATIONS

Esmail Mohebbi

University of West Florida, Pensacola, USA

ABSTRACT

Operations management courses are known for covering a broad spectrum of complex and interrelated quantitative and quantitative concepts and technics that are applicable to various decisions involved in designing, planning, executing, and controlling of manufacturing or service operations. The vast diversity of topics coupled with various degrees of intricacy and sophistication associated with the technics covered by courses of this nature necessitates supplementary instructional tools to facilitate students learning. Simulation-based instructional tools are increasingly incorporated into operations management courses as integral components of teaching strategies aimed at enhancing student learning outcomes. This paper presents some take-aways from an experiment involving the application of web-based simulation tool in an undergraduate OM course.

Keywords: Operations Management; Instructional Methods; Simulation Games, Teaching and Learning

INTRODUCTION

Teaching operations management (OM) regardless of what name it is identified as a course in a business degree curriculum is indeed an interesting and challenging task. The course typically occupies a unique position in most business degree programs by serving as the main (and often only) conduit for delivering a mélange of interrelated qualitative and quantitative concepts and technics with various degrees of intricacies and sophistications to fulfill one component of what is commonly known as the ternion of content requirements for business core curriculum: namely, accounting/finance, marketing, and operations. Accordingly, students often themselves befuddled by the extent, variety, and complexity of seemingly “practical” – but not easily connectable – problem-solving and decision-making methodologies that are encapsulated in the course content. This poses the instructor with the challenge of cultivating a learning environment in which students are provided with the tools that they need to not only gain competency in applying individual analytical technics, but to also develop the ability to contextualize, interpret, and coalesce such techniques in real-life business settings, thereby elevating their motivation and enhancing the relevance and utility of their educational experience. It should be noted, however, that while the difficulty of meting this challenge may vary from one course to another depending on the subject, the concerns about the consequences of leaving it unanswered are uniformly strong and attention worthy across all business courses.

In relationship to the above, business education is known for facing the challenge of keeping pace with the rapidly evolving world of modern business practices on an ongoing basis. Accordingly, driven by employers' demand, business schools continuously strive to produce work-force ready graduates by infusing real or quasi-real world business context into their curriculum. To that end, the simulation approach toward teaching business subjects promotes utilizing "representations of some aspects of business life to which students are asked to react." (Able, 1980). Over the year, both quantitative and nonquantitative simulations have been incorporated into various business courses as instructional and learning tools. The dependencies of such tools on technology ranges widely from low-tech role-playing exercises in the classroom to high-tech computer simulations and web-based applications.

Business games are popular simulations that have long been recognized as means of promoting experiential learning by engaging students with artificial and simplified (or abstract) forms of real-word business-like environments and providing them with the opportunity to acquire and strengthen problem solving, critical thinking and decision-making skills while acting as business practitioners as realistically as possible (Wolfe, 1993). The scope of business games may vary widely from functional simulations involving operational decision making using quantitative techniques in a structured manner to total enterprise simulations involving unstructured strategic management decision making using mostly qualitative methods. While functional business simulations are germane to cultivating and enhancing students' technical problem-solving and decision-making skills in specific disciplines form an operational standpoint, enterprise simulations gravitate toward immersing students in multi-disciplinary and cross-functional qualitative decision-making processes. (Gibbons et al., 2022). Engagement, interaction, imagination, sense of achievement and fun are commonly mentioned among the important characteristics of effective business simulation games in the literature (Goi, 2019).

This paper presents several takeaways form an experiment involving the incorporation of a web-based functional simulation platform into an OM class offered to undergraduate business students in an AACSB accredited business school in online and blended formats.

BACKGROUND

Historically, the development and adoption of business simulation games as training and educational tools seems to have started in the early 1930s in Europe (Faria et al., 2009) and spread rapidly among the business schools in North America and around the globe (Tanner et al. 2012). On that note, modern simulation games have evolved extensively since their introduction by the American Management Association in the 1950s (Cohen & Rhenman, 1961) owing to computers and ensuing advancements in computing, graphics, and information technologies– which has opened a wide spectrum of possibilities for simulation tools in traditional, online, or blended classroom environments.

Early applications of computer simulation models as a teaching tool in the business education literature reflect comprehensive budgeting, banking, and profit planning exercises. The computerized simulation models were viewed as an efficient enabler for capturing the complexity of relationships between the decision variables involved in such endeavors albeit in approximate manner and help students understand microeconomic relationships and concepts in managerial economics (Sale, 1972, Breen and Boyd, 1976;

Rives 1976). Examples of more recent applications of computer simulations in related fields include Chen & Yur-Austin (2013), Chulkov & Wang (2020), Neriz et al. (2020), Sierra (2020), among others. International business (e.g., McGuinness, 2004; Farrell, 2005), human resource management (e.g., Elliott et al., 2002; DeGroot et al., 2009; Rompho, 2011), ethics (e.g., Teach et al. 2005), and management information systems (e.g., Larson 2013; Saraswat et al., 2014; Hwang, 2019; Henkel and Bider, 2020) are also among the areas of business education where applications of simulation-based teaching tools have been reported in the literature.

Strategic management is one area of business education that has been historically more attuned to using simulation games. Capstone courses in business curriculum usually require students to formulate and execute business strategies as if they were making decisions for a real firm; hence, reports of adopting simulation games in such courses are more observable in the literature (e.g., Keyes, 1997; Jenner et al. 2010; Reid et al., 2012; Capelo et al., 2015; Capelo et al., 2021; García and Cabañas, 2022).

Simulation-based instructional and learning tools have been incorporated into operations, logistics, and supply chain management (SCM) course designs and textbooks over the years. To that end, perhaps the most famous manifestation of how simulation games can be incorporated into OM and similar courses ensued using the popular Beer Distribution Game (BDG) to exhibit the interlinked operations of a multiechelon production-inventory system, otherwise known as a supply chain, and highlight the cascading effects of decisions made by individual players regarding their operating policies on the end results. For example, motivated by the BDG concept, Anderson & Morrice (2000) presented a computerized simulation of a service-oriented supply chain with no finished goods inventory where players must manage order backlogs by adjusting their capacity. Sparling (2002) described an SCM teaching strategy which calls for creating instructional activities based on modified BDG simulations using Excel spreadsheets. Based on personal experience, the author found the simulation to work well in both undergraduate and graduate courses with participants ranging form15 to 50 students. Similarly, drawing upon their experiences with BDG and a few other games, Merkuryev et al. (2009) presented a computer simulation model designed to convey the concept of cyclical versus noncyclical inventory replenishment polices in supply chains.

Among other published reports, Ellis et al. (2014) described a physical simulation exercise involving the supply chain for producing paper airplanes to facilitate students' understanding of lean six sigma concepts. Based on survey results gathered from a group of MBA students, they concluded that exercise proved to be an effective learning tool by inducing participation and teamwork among students. Likewise, Webb et al. (2014) presented an in-class simulation exercise based on the SCOR (Supply Chain Operations Reference) model in which students assumed the position of various links in supply chain producing a consumer good. The authors reported positive feedback from participants suggesting a "holistic understanding" of SCM as the main takeaway of students from the experience. A five-point Likert survey conducted by Nguyen (2015) suggested that students playing a web-based simulation game in an operations management course believed they had experienced a deeper learning whereas those who did not participated in the activity felt more motivated and competent in their learning efforts. Tiger et al. (2019) presented a spreadsheet simulation tool in Microsoft Excel for teaching the concepts of risk pooling and safety-stock reduction in inventory management. Feedback from undergraduate and MBA students appeared to be encouraging.

More recently, Chuang (2020) provided an overview of an off-the-shelf web based SCM simulation game and shared teaching tips based on classroom experiences with the software. By the same token, Kacprzak et al. (2020) presented an overview of four computer simulation games available in the market for teaching production management. On a related topic, Thürer et al. (2020) described a physical simulation game intended for teaching production control techniques in a non-repetitive context. The proposed game involved an extension of the match-and-dice game to capture the underlying information feedback looping structure of card-based production control systems. Games scores and self-assessment survey results were used to verify the effectiveness of the teaching instrument. Song et al. (2021) introduced an online simulation game for teaching OM and SCM courses. The game is intended to help students understand supply shortages and product ratioing in supply chains that face emergency situations. The authors used comparative test scores to verify that students who played the game performed better than those than their peers who did not.

It is important to note that regardless of their physical or computerized nature, the common theme of these simulation tools has been to immerse students in an interactive, dynamic, and rapidly changing fictitious manufacturing, service, or other business operations environment to which they can map on specific topics and techniques covered in the course while emulating various practitioners' roles as closely as possible. However, from a historical standpoint, presentation of computerized simulation games as instructional tools for teaching OM/SCM concepts has received considerably more attention in the literature than examining the impact of such tools on students' learning outcome. In other words, due to scarcity of off-the-shelf interactive simulation games capable of addressing the instructional needs of OM/SCM courses in the past, developing simulation-based tools that could capture the intricacies and characteristics of specific operational or supply chain managerial settings and/or presenting tips on how to incorporate them into course constructs appears to have taken precedence over assessing the effectiveness of incorporating such tools in course designs. Not surprisingly, the growing availability of customizable interactive off-the-shelf OM/SCM-related computer simulation platforms designed for educational purposes coupled with their ease of access and potential incorporation into various web-based learning management systems (LMS) has invigorated a line of research inquiries into the fruitfulness rather than sheer development or application of such tools in more recent years. The present report reflects an attempt in pursuing this line of inquiry.

EXPERIMENTAL SETTING

An undergraduate OM course taught in multiple sections each semester provides the setting for an ongoing pedagogical experiment involving incorporation of computer simulation exercises into the course content. The experiment is intended to gain insights into if and how such exercises could facilitate the students' perceived as well as demonstrated functional knowledge of OM concepts and techniques, thereby devising improved strategies to enhance their overall learning experience in the course.

The OM course in focus is part of the multi-course core program of study required for all students pursuing business degrees in an accredited business school in the US. Students usually take the course in their junior or senior years after successfully completing a prerequisite in elementary statistics. The course in turn serves as a prerequisite for the capstone course in the college core curriculum. While not all sections are taught by the same instructor, the experiment and the ensuing take ways reported in this article are based only on the sections taught by the author using the same textbook and LMS and covering the same content regardless of their modality (i.e., traditional, blended, online). The textbook is embedded in a web-based instructional platform supported by the publisher which enables the instructors to choose from various builtin supplementary constructs as well as adaptive and interactive tools to create customized contents. To that end, the web-based platform offers a la carte menu of interactive scenario-based simulation exercises devised to emulate the plethora and urgency of the challenges that OM practitioners in real-world business operations. Each exercise is highly interactive and dynamic by design and caters to both qualitative and quantitative (when applicable) aspects of specific functional topics of interest (e.g., inventory control, project management, forecasting, etc.). Furthermore, each scenario hinges on synthesizing the interconnectivity and nontrivial interactions among factors influencing OM decisions and their expected outcomes. Awareness, engagement, and diligence as well as agility are among the key qualities that players must demonstrate to accomplish the tasks required in each exercise. Simulations are self-contained. An orientation module embedded in each exercise illustrates how the simulation works, what tasks are to be performed, what constraints are to me met, and what objectives are to be achieved. Hints are provided upon request. After completing an attempt at the simulation, students are assigned a score along with a report containing meaningful feedback about their performance in each portion of the simulation. The report also provides students with instructional tips for improving their performance in subsequent attempts if applicable.

The features outlined above, coupled with the ability to seamlessly blend the simulation exercises into the LMS used to deliver the course, led to adopting the web-based platform as the experiential tool for exploring the fruitfulness of using OM simulations in a controlled experiment. Comparative analysis of the students' perceived as well as actual (i.e., performance-based) learning outcomes between multiple treatment and control groups is designated as the primary means of drawing inferences based a quasi-experimental design. To that end, the experiment has the distinctive characteristic of placing an emphasis on examining the impact (if any) of the simulation-based activities on students' ability to contextualize OM concepts and techniques.

PRELIMINARY TAKEAWAYS

Inquiries into both subjective and objective assessments of the impacts of OM simulations on students' learning outcomes are in progress. Preliminary takeaways based on student' feedback and instructor's observations concerning this ongoing experiment include:

- The self-paced guided structure of simulation exercises appears to be very appealing to the students. Furthermore, they the ability to pause the simulation to revisit the tutorial or review course topics, when authorized in the settings, before resuming a run has been reported to be very useful in mapping specific topics onto the simulated operations; thereby helping students with conceptualization of concepts.
- Allowing multiple attempts coupled with generating a score sheet containing meaningful performance-based feedback and hints for improvement after each run has proven to be critical in motivating and helping students meet the targeted competency levels.
- When multiple attempts are allowed at each simulation, students appear to be more likely to rely on trails and errors or improvised heuristics rather than utilizing more elaborate methods covered in the course for achieving performance targets.
- Replications of similar scenarios with different data sets appears to be as positive influencers of students' self-efficacy and facilitators of learning outcomes. Furthermore, students' abilities to retain and reapply the concepts and skills covered by each scenario appear to have been noticeably improved accordingly.
- Students seem to find the simulations more beneficial toward improving their interpretative and system-thinking abilities compared to enhancing their quantitative skills.

- Students' motivation and engagement toward other coursework activities appear to be improving over time as simulation exercises are introduced throughout the semester.

CONCLUSIONS

Computer simulation is commonly described as an integral component of the system thinking approach that enables the analyst to emulate and examine the key dynamics and inter-complexities of real-world systems through elaborate (albeit approximate) modeling and experimentation. Business educators have been known for developing or adopting computer simulation games in various disciplines with varying degrees of interest and intensity over the years. To that end, web-based operations and supply chain simulations have gained popularity in online or blended courses in recent years – owing to advancement in technology on the one hand and the ever-growing complexity of OM/SCM problems in practice on the other. This article served as a conduit for sharing some preliminary observations based on ongoing experiment involving the incorporation of a computer simulation platform into an undergraduate OM course. More research is needed to draw inferences based on quantitative and qualitative assessments of performance and perception-based data-driven results.

REFERENCES

1. Able, W.W. Simulations: an applied approach to teaching business. *Technological Horizons in Education*, 7(4), 45-46, 1980.
2. Anderson, E.G. and Morrice, D.J. A simulation model for teaching service-oriented supply chain management: does information sharing help managers with service capacity decisions? *Productions and Operations Management*, 9(1), 40-55, 2000.
3. Breen, W., and Boyd, J. Classroom simulation as a pedagogical device in teaching money and banking. *The Journal of Financial and Quantitative Analysis*, 11(4), 595-606, 1976.
4. Capelo, C., Lopes, A., and Mata, A. A Simulation-based approach for teaching the systems perspective of strategic performance management. *Accounting Education: An International Journal*, 241, 1–26, 2015.
5. Capelo, C., Pereira, R., and Ferreira Dias, J. Teaching the dynamics of the growth of a business venture. *The International Journal of Management Education*, 19(3),100549, 2021.
6. Chen, X., and Yur-Austin, J. Exploring international investment through a classroom
7. portfolio simulation project. *Journal of Teaching in International Business*, 24(1), 21–30, 2013
8. Chuang, M. A web-based simulation game for teaching supply chain management. *Management Teaching Review*, 5(3) 265–274, 2020.
9. Chulkov, D., and Wang, X. The educational value of simulation as a teaching strategy in a finance course. *e-Journal of Business Education & Scholarship of Teaching*, 14(1), 40-56, 2020
10. Cohen, K., and Rhenman, E. The role of management games in education and research. *Management Science*, 7(2), 131-166, 2006.
11. DeGroot, T., Stambaugh, J., and Owen, J., An empirical examination of the use of a simulation in teaching human resource management. *Journal of Human Resources Education*, 3(3), 1-12, 2009.
12. Elliott, M., Kaufman, S., Gardner, R., and Burgess, G. Teaching conflict assessment and frame analysis through interactive web-based simulations. *The International Journal of Conflict Management,* 13 (4), 320-340, 2002.
13. Ellis, S., Goldsby, T.J., Bailey, A.M., and Oh, J. Teaching lean six sigma within a supply chain context: the airplane supply chain simulation. *Decision Sciences Journal of Innovative Education*, 12(4), 287-319, 2014.
14. Faria, A.J., Hutchinson, D., Wellington, W., and Gold, S. Developments in business gaming – a review of the past 40 years. *Simulation & Gaming*, 40(4), 464-487, 2009.
15. Farrell, C. Perceived effectiveness of simulations in international business pedagogy: an exploratory analysis. *Journal of Teaching in International Business*, 16(3), 71-88, 2005.
16. García, L.D.G, and Cabañas, M.A. Integrating the external enablers of new venture creation theory into practice: A simulation approach for teaching entrepreneurship. Journal of Education for Business, 1-12, 2022 (DOI: 10.1080/08832323.2022.2138248).

17. Gibbons, B., Fernando, M., and Spedding, T. Innovation through developing a total enterprise computer simulation: teaching responsible decision making. *Journal of Management Education*, 46(1), 16–42, 2022.
18. Goi, C. The use of business simulation games in teaching and learning. *Journal of Education for Business*, 94(5), 342–349, 2019.
19. Henkel, M., and Bider, I. Creating apprenticeship simulations for teaching enterprise modelling – principles and lessons learned. *Innovations in Education and Teaching International*, 57(4), 413423, 2020.
20. Hwang, M. ERP simulation games in asynchronous online classes, *e-Journal of Business Education & Scholarship of Teaching*, 13 (3), 37-45, 2019.
21. Jenner, S., Zhao, M., and Foote, T.H., Teamwork and team performance in online simulations: the business strategy game. *Journal of Online Learning and Teaching*, 6(2), 416-422, 2010.
22. Kacprzak, K., Pietraszewska, A., and Wasążnik, M. Computer simulation games in teaching production management. *Zarządzanie*, 47(4), 7-19, 2020.
23. Knott, U.S., and Keys, J.B., Teaching strategic management with a business game. Simulation & Gaming, 28(4), 377-394, 1997.
24. Larson, E.C. Utilizing classroom simulation to convey key concepts in IT portfolio management, *Journal of Information Systems Education*, 24(2), 99-104, 2013.
25. McGuinness, M. A simulation Game for an introductory course in international business. *Journal of Teaching in International Business*, 15(4), 47-66, 2004.
26. Merkuryev, Y., Merkuryeva, G., Bikovska, J., Hatem, J., and Desmet, B. Business simulation game for teaching multi-echelon supply chain management. *International Journal of Simulation and Process Modelling*, 5(4), 89–299, 2009.
27. Neriz, L., Núñez, A., Fuentes-Caceres, V., Ramis, F., and Jerez, O. Simulation-based training as a teaching and learning tool for management education. *Innovations in Education and Teaching International*, 57 (6), 701–713, 2020.
28. Nguyen, T.N. Motivational effect of web-based simulation game in teaching operations management, *Journal of Education and Training Studies*, 3(2), 9-15, 2015.
29. Reid, M., Brown, S., and Tabibzadeh, K. Capstone teaching models: combining simulation,analytical intuitive learning processes, history and effectiveness. *Journal of Education for Business*, 87(3), 178-184, 2012.
30. Rives, J.M. Teaching managerial economics with a computer simulation model. *The Journal of Economic Education*, 7(2), 133-134, 1976.
31. Rompho, N. Using Simulation games and case studies in teaching performance measurement and management, *Journal of Higher Education Theory and Practice*, 11(3), 43-50 2011.
32. Sale, J.T. Using computerized budget simulation models as a teaching device. *The Accounting Review*, 47(4), 836-839, 1972.
33. Saraswat, S.P., Anderson, D.M., and Chircu, A.M. Teaching business process management with simulation in graduate business programs: an integrative approach. *Journal of Information Systems Education*, 25(3), 221-232, 2014.
34. Sierra, J. The importance of simulation in teaching and learning economics: The students' perspective. *Innovations in Education and Teaching International*, 57 (5), 521–531, 2020.

35. Song, J.M., Park, A., and Zhao, Y. The hunger chain: A competitive simulation for teaching supply chain management. *Decision Sciences Journal of Innovative Education*, 19(3), 204-213, 2021.
36. Sparling, D. Simulations and supply chains: strategies for teaching supply chain management. Supply Chain Management, 7(5), 334-342, 2002.
37. Tanner, J., Stewart, G., Totaro, M., and Hargrave, M. Business simulation games: effective teaching tools or Window dressing? *American Journal Of Business Education*, 5(2), 115-128, 2012.
38. Teach, D., Christensen, S.L., and Schwartz, R.G. Teaching business ethics: integrity. *Simulation & Gaming*, 36(3), 383-387, 2005.
39. Thürer, M., Cole, R., Hanna, M.D., and Protzman, C.W. Classroom simulations for teaching production control in nonrepetitive contexts: insights for theory and practice. *Decision Sciences Journal of Innovative Education*, 18(4), 568-588, 2020.
40. Tiger, A.A., Loucks, J., and Burns, C. Spreadsheet-based supply chain simulation for teaching risk pooling combined with facility location. *Independent Journal of Management & Production*, 10(6), 1932-1951, 2019.
41. Webb, G.S., Thomas, S.P., and Liao-Troth, S. Teaching supply chain management complexities:
a SCOR model based classroom simulation. *Decision Sciences Journal of Innovative Education*, 12(3), 181-198, 2014.
42. Wolfe, J. A history of business teaching games in English-speaking and post-socialist countries:
the origination and diffusion of a management education and development technology. *Simulation & Gaming*, 24(4), 446-463, 1993.

27. Validating Core Skills in IS Data Analytics Curriculum: A Knowledge Graph Approach

Yong Seog Kim [1], Emma Cherrington [2]

[1]Utah State University, Logan, USA. [2]Revere Health, Salt Lake, USA

Abstract

In this study, we compared core skills geared toward future data analysts and data scientists from two data sources: 50 Management Information Systems (MIS), IS or Data Analytics course descriptions of business colleges in the Intermountain West Region and 11,225 job descriptions of data scientists. According to word clouds from the two data sources, curriculum source presents more general terms related to theories while job description sources illustrates more specific terms related to data usage, data types and analytical skills. In addition, from multiple knowledge graphs (KGs) based textual course and job descriptions, we found that job descriptions source presents the importance of application of specific skills to various business settings while curriculum source focuses on the conceptual and theoretical skills along with fundamental skills such as SQL. Based on our findings, we suggest that courses should address not only technical skills, but also soft skills to make informed decisions and turn them into business decisions through real application scenarios.

Keywords

Knowledge graph, Word cloud, Data analytics, Data science, IS curriculum

Introduction

By its nature closely tied to the rapidly evolving Information and Communication Technology (ICT), the Management Information Systems (MIS), IS or Data Analytics discipline (denoted as DAIS from now on for notational convenience) has responded to program sustainability issues and core curriculum change issues due to technology disruption where newer technologies replace old legacy technology platforms and associated business paradigms over the last 50 years (Case et al. 2019). However, within DAIS discipline, there have been different views on the right direction of the discipline that would determine core teaching curriculum and research themes to overcome IS identity crisis (Müller et al. 2016). For example, one group of DAIS scholars claims that DAIS research community should devote more times and efforts to research and teaching on phenomena intimately associated with IT-based systems (Benbasat and Zmud 2003). In contrast, another stream of IS scholars insists that IS research community should stay on more macro studies of IT to capture the transformational power of the technology (Agarwal and Lucas 2005). While there are different views on detailed ideas of core DAIS teaching curriculum and research theme for future, it is very clear that DAIS discipline should adopt technology disruption as an inevitable phenomenon and develop core teaching curriculum centered around emerging digital technologies so that students master appropriate technical skill sets and prepare for new business models, products, and services upon their graduation (Case et al. 2019).

In particular, the pervasive expansion of data volume due to the proliferation of the web, social media, and mobile devices brings topics such as data mining to IS educators as being the next big thing (Jafar et al. 2017). In fact, during the past two decades, data mining courses in IS curriculum has evolved from appearing as an elective course to a minor area of study, a co/dual-major, or even as a fully-independent degree program under various names of data mining, data science, data analytics, business analytics or business intelligence (Jafar et al. 2017). While contents among these courses are very similar, business analytics or business intelligence courses tend to emphasize the application of data analysis to solve business problems by revealing patterns and extracting insights from the data, other courses focus on statistical, mathematical or algorithmic properties of data analysis (Jafar et al. 2017). However, the authors believe that the innovation of DAIS curriculum should be evaluated not only by measuring whether new technologies and principles are incorporated into a new DAIS curriculum but also by measuring whether newly incorporated contents in a DAIS curriculum are appropriate in terms of depth and breadth for practical application of such skills.

In this paper, assuming that almost all academic programs try to educate undergraduate and master students for their future careers in organizations, we would like to measure whether newly incorporated contents in data analytics curriculum are aligned with highly demanded skills from organizations. Specifically, we aim to identify core analytical skills that current DAIS curriculum intends to offer for future data scientists or data analysts from course descriptions. At the same time, we also like to identify core analytical skills that industrial employers identify critical for positions as data scientist or data analyst from job descriptions. Once such skill sets are identified, we would like to compare them to see if there exists mutual consensus or disparity between skill sets obtained from course and job descriptions and make recommendations for changes in DAIS curriculum to bridge the gaps if necessary.

Our paper is organized as follows: First, we will start with brief literature review on data analytics curriculum in IS and knowledge graphs (KGs). Then we will briefly introduce two data sets used for our analysis along with data engineering steps used to pre-process them. Then, we present outcomes including core hard and soft analytic skills from course and job descriptions at two different levels of granularity. At word or phrase level, we will create and contrast word clouds to visualize the most prominent words and phrases from a set of course and job descriptions. Then, we will extend analysis to the sentence/paragraph level at which semantic and syntactic roles of words or phrases within sentence are considered. To this end, we will identify a fact in a triple of the form from each sentence in course (or job) descriptions and present all facts in a knowledge graph (KG). Note that a KG breaks down textual information into a set of sentences, each of which will be restructured as a fact with three main components as a triple of the form which includes the source entity, relation, and target entity. We expect to find multiple facts such as “The candidate will need SQL skill sets.” Once we obtain KGs from course and job descriptions, we will compare them to see whether data analytic skills from two data sets are consistent. We will conclude this paper with future research direction.

Literature Review on Knowledge Graphs

While there is no single clear definition of Knowledge Graphs (KGs) (Ehrlinger and Wöß 2016), they are often defined as large networks of real world entities and their relationships with entities' semantic types and properties over various topical domains (Paulheim 2017). In essence, a KG is a multi-relational graph composed of entities (nodes) and relations (different types of edges), where each edge is represented as a triple of the form (source entity, relation, target entity) called a fact. Its notion has garnered a great deal of attention since Google introduced its own version of a KG to improve the value of search results in 2012. Since then, a number of KGs including Yahoo's Spark, Microsoft's Satori and Facebook's entity graph have been applied to voice assistants (Paulheim 2017; Singh et al. 2018), information integration and extraction (Schulz et al. 2009; Pfaff et al. 2018), named entity disambiguation (Lan et al. 2016), and locating content files (Yahya et al. 2012). Other notable applications of KGs include cases to integrate textual medical knowledge with health data (Shi et al. 2017) and evaluate the stock news sentiment market (Liu et al. 2019) with some success.

KGs have several useful properties. First of all, by representing entities with their semantic types and properties, KGs help AI applications understand the context and the meaning behind various concepts in data (e.g., a person with title, role on a project and expertise areas). At the same time, KGs are also very useful for data integration tasks because they need to aggregate all information in disparate data sources with different data sizes and levels of heterogeneity throughout organization on a specific topic, person, project, or product (Ehrlinger and Wöß 2016). However, one of most useful properties of KGs is that since they are intended to identify hidden facts and relationships between entities, they often uncover facts from seemingly unrelated events or entities. For example, relationships identified by a KG from two seemingly unrelated crimes in two different places in the UK revealed that these criminals knew each other and were affiliated to the same crime group (Ehrlinger and Wöß 2016). However, it is important to note that maintaining data sources complete and consistent without conflicting information is highly critical for the quality of KGs (Bordes and Gabrilovich 2014; Paulheim and Bizer 2014; Rashid et al. 2019).

Recently, Enterprise KGs provide a general platform that allow companies to build their own network of knowledge in their business domain. Therefore, with the help of Enterprise KGs, organizations can integrate heterogeneous data sources within their existing IT systems and make decision making process agile and efficient. For this reason, KGs has significantly impacted traditional decision support system (DSS) by presenting information and knowledge in a more dynamic, scalable and domain independent form (Elnagar and Weistroffer 2019). That is, KGs become an essential part of enterprise DSSs by generating scalable real-time knowledge representation with integrity from unstructured data for easy search and retrieval tasks (Paulheim 2017; Gomez-Perez et al. 2017; Zaveri et al. 2016).

Data Sets and Data Engineering

Two data sets were collected, cleansed, and engineered for this study. The first data set is a small sample of undergraduate and graduate course descriptions offered by several colleges in Intermountain West Region. The second data set is a compiled set of 11,225 job descriptions advertised for data scientists and data analysts across the United States. We provide detailed descriptions of the two data sets in the following sections.

Course Description Data

This data set was obtained by collecting course descriptions of 50 DAIS classes of colleges in the Intermountain West Region that the authors believe provide career foundations for data analysts and data scientists through providing contents in programming, databases, data warehousing, data mining, and data analytics. Other core DAIS classes such as project management, telecommunication, security, business strategies, and behavioral/organizational studies were not considered in this study because they are not directly related to data analytics positions. For each course, we collected 'CourseNumber', 'CourseName', 'CollegeName', and 'CourseDescription' that describe core skills and deliverables of each class in different styles, formats, and lengths. For examples, many courses describe core skills and objectives under various headings or sections such as 'Objectives', 'Outcomes' or 'Subjects' either in full complete sentences or multiple short phrases. To process original course descriptions for KG analysis, we used Python libraries to remove any extra spacing and lines around these headings or sections. In addition, we also removed headings and sections themselves so that they are not considered valid entities as a part of course description. Otherwise, they may be recognized in our entities in KGs, which will give unexpected results. Note that KGs can be built based only on sentences that include at least two entities (mostly represented by subject and object) and a relationship (most represented by a verb).

Job Description Data

The secondary data set of job descriptions was acquired by utilizing Web scraping library in Python 3.75 from Indeed's website (www.indeed.com). The scraping was completed on November 22nd 2019 for data scientist or data analyst positions in US, resulting 11,225 jobs along with self-explanatory input variables such as 'JobID', 'Company', 'JobDescription', 'JobLocation', 'JobWebsite', 'Salary', and 'JobTitle.' Other additional input variables include 'JobPosted' (job posting days from web scraping date) and 'Summary' (brief summary of job description), and 'Rating' (company rating based on 1-5 scales). While it is possible to use 'Summary' variable, we used 'JobDescription' information that typically include multiple paragraphs of information on the job to build a comprehensive set of KGs. While this data set was relatively clean, some basic cleaning needs to be done using Python library such as removing unnecessary line breaks, white noise characters (e.g., tabs), and job listing number in job descriptions.

Results and Discussion

Word Cloud Results

We first visualized the most prominent words or phrases in both course and job descriptions using word clouds, which visually present textual information weighted by their frequencies or importance using different font sizes or colors. To this end, JobDescription and CourseDescription variables from two data sets were loaded into Python using the WordCloud package. We presented two outputs of word cloud analysis in Figures 1 (a) and (b).

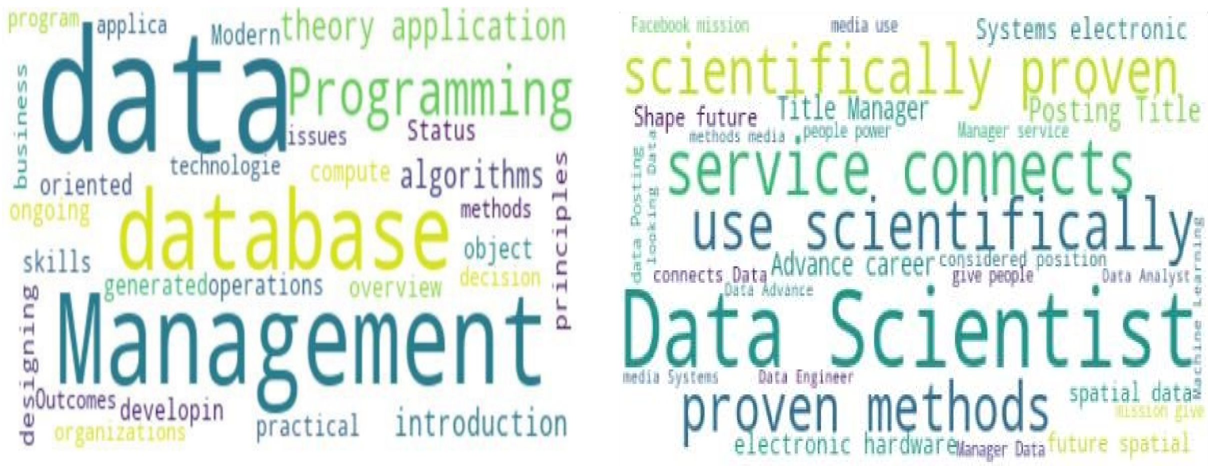


Figure 1 (a): *Class Description Word Cloud* Figure 1 (b): *Job Description Word Cloud*

The first word cloud in Figure 1 (a) display most prominent words with bigger font sizes for more prevalent words in CourseDescription variable. Note that we also had to remove several stop words (e.g., 'course', 'covers', 'teaches', and 'concepts') that appeared frequently in the course descriptions but do not deliver useful insights. According to Figure 1 (a), 'Management', 'data', and 'database' were found to be the most prevalent words in course descriptions. We attribute this finding to the fact that many data analytical courses in DAIS program are still built around the usage of SQL in database management systems. Other frequent words reflect objectives of many courses toward integrating theories (e.g., 'theory' and 'principles') and applications (e.g., 'business', 'decision', 'practical' and 'application') in STEM degrees like DAIS. However, specific analytical skills that each course aim to deliver are not shown in this word cloud except general programming related words such as 'object', 'oriented', 'programming' and 'algorithm'.

Figure 1 (b) presents the most prevalent words from the 11,225 job descriptions after removing stop words such as 'together', 'brings', 'mutual', 'job', 'seen', and 'free'. Note that some stop words like 'together' and 'mutual' may reflect cultural and societal goals of job hosting companies. However, we decided to remove them from our word cloud mainly because our primary goal is to identify key analytical and technical skill mentioned in job descriptions for data scientists. As expected, the most prevalent single words are 'data scientist' with other associated words like data usage terms ('use scientifically', 'scientifically proven', 'data posting', 'looking data', 'connects Data', and 'Data Advance'). We also found words related to specific types of data ('spatial data', 'methods media', 'media Systems') and a specific analytical skill such as ('Machine Learning') although we could not find other analytical skill terms such as Python, Analysis, SQL and Artificial Intelligence. Few other prevalent words are mainly related to titles ('Posting Title', 'Manager service', 'Title Manager'), and careers ('Advance career', 'Shape future', 'future spatial').

Knowledge Graph from Course Description Data Set

While word clouds are useful for identifying prominent words, they are limited in that they do not consider semantic meaning and syntactic structure of words in sentences. To this end, we created a large number of facts or relationships between entities using all sentences in course and job descriptions while paying attention to the relationship between the subject and object of each sentence. To achieve this, several Python libraries including Spacy, Networkx, and Tokenize packages were utilized. In our implementation for creating KGs of course descriptions, all sentences in each course description were extracted and then broken into their grammatical parts (i.e., subject, object, and verb (or predicate)) identified using Spacy and Tokenize packages. In the end, each sentence with identified grammatical parts is represented as a triple (<subject, predicate, object>) and it is added to a KG. Specifically, subject and object are represented as source and target nodes, respectively, and predicate is represented as a directional link from subject to object. Note that since there could be multiple sentences with different predicates for the same pair of subject and object, a KG in our analysis contains possible multiple directional links for the same pair of source and target nodes, making it a multi directional graph. In a similar way, KGs of job descriptions were also created.

We first present in Figure 2 the overall structure of the KG that was created from the course description data set. This KG has a total of 79 nodes (i.e., source and target nodes) that were connected through 68 directional edges (or relations). As expected, many relations started from few core nodes like “course”, “students” and “topics”. For example, the KG in Figure 2 shows that many DAIS courses and their topics (as source nodes) deliver or lead to (through relations) theories or skill sets (as target nodes) such as database systems, SQL focus, data analytics, mining business technologies, Java programming language, database theory, data mining, data visualization, and predictive information. We found that this is very encouraging because many DAIS courses indeed target skill sets that are foundational or critical for data scientists or data analysts.

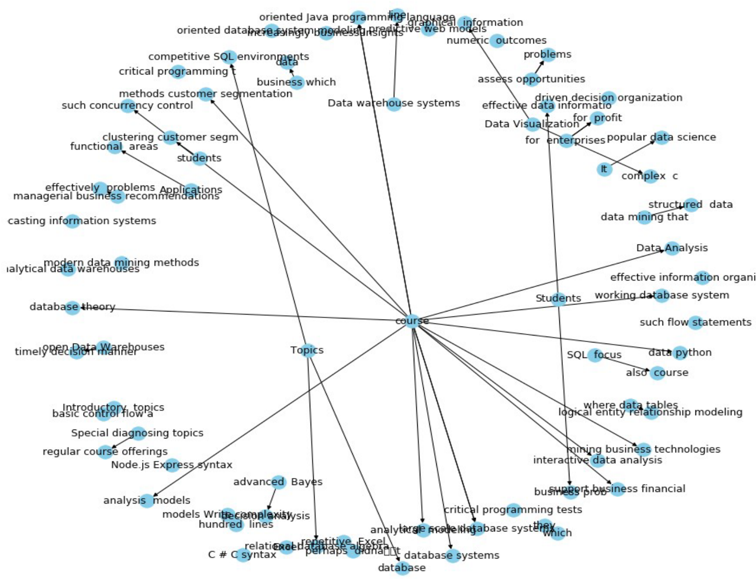


Figure 2: *Course Description Knowledge Graph*

Noting that the KG is a collection of multiple facts or relations between source and target nodes, we decided to take a closer look on specific relations that we are interested in. For example, we paid attention to a relation called 'develop' mainly because we were eager to understand what specific analytical skills courses 'develop' or help students 'develop'. The first set of source and target nodes associated with the 'develop' relation is 'course' and 'analysis models', which can be interpreted as "the course 'develop' anaylsis models". The second set of source and target nodes is 'students' and 'real world business problems', which completes the following fact: "students develop (skills for) real world business problems." From these two facts, we infered that at least some DAIS courses (we verified from our data set that these courses are indeed courses related to data mining and data sciences) aim to help students develop analysis models so that as future data scientists they can apply such analysis models to tackle real world business problems.

Relation Knowledge Subgraphs from Job Description Data Set

We also created a complete KG that includes all relations between source and target nodes based on job descriptions. However, since this complete KG has a total of 11,633 unique directional relations between 3,129 nodes, it is almost impossible to understand or interpret due to its complex structure and size. Therefore, we decided to analyze multiple subgraphs from this complete KG based on relations between source and target nodes that we are interested in. The first relation that we wanted to analyze was the 'Looking For' relation and we present in Figure 3 a subgraph of the complete KG that includes all pairs of source and target nodes connected through 'Looking For' relation. Our intuition of selecting this specific relation and focusing on the resulting subgraph was that companies are most likely to list analytical skills in their job description that they were looking for from eligible candidates for their job postings.

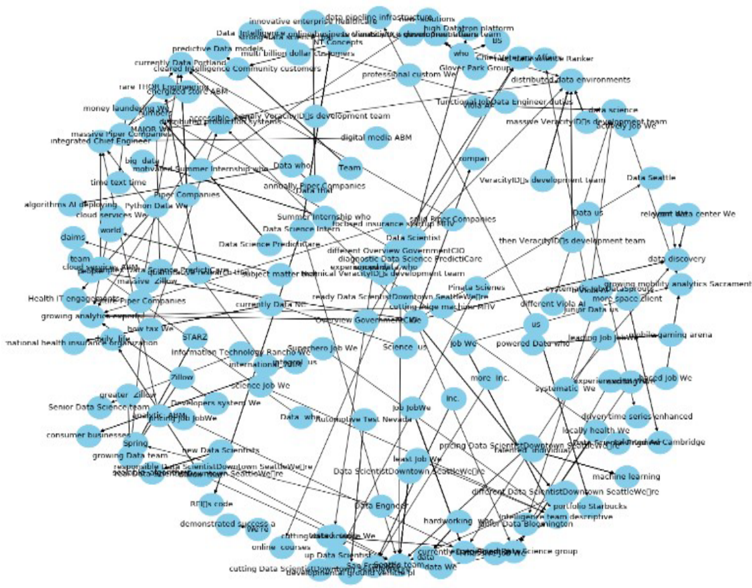


Figure 3: Knowledge Subgraph with 'Looking For' Relation

Figure 3 presents 'Looking For' relations between source and target nodes. Visual inspections on the subgraph followed simple Python commands for network analysis revealed that target nodes such as 'Algorithms AI deploying', 'Machine Learning', 'Data Science Engineer', 'Data Discovery', and 'Growing analytics Experts' have the largest number of 'Looking For' relations from source nodes like '7 State Providence Company', 'Handworker who', 'Glover Park Group', and 'Data Engineer duties'. Thus, this knowledge subgraph confirms our initial conjecture that many companies are looking for candidates with broad ranges of data analytic skills such as 'Algorithms AI deploying', 'Machine Learning', 'Data Science Engineer', and 'Data Discovery.' However, we found that this knowledge subgraph does not identify target nodes that reflect specific analytic skills like 'Python Programming', 'R Programming', or 'Statistical Analysis' for a chosen 'Looking For' relation.

The second knowledge subgraph was identified using the 'Is responsible' relation, anticipating that the knowledge subgraph would find multi-facts reading like 'the candidate is responsible for a task'. We show this knowledge subgraph with 'Is responsible' relations between source and target nodes in Figure 4. Careful investigations on this subgraph revealed that target nodes such as 'innovative business solutions', 'analytical business operations', 'complex business', and 'wide university marketing' have 'is responsible' relations from source nodes such as 'Data Scientist', 'Position Division' or 'Staff Data Scientist'. We found this observation extremely interesting considering the fact that almost all data scientist and analysts job position are supposed to be responsible for creating business [or university] solutions or operations by utilizing their skills and knowledges. This implies that many companies emphasize in their job description not only hard skills that allow analysts or developers to analyze data sets but also soft skills that utilize information and knowledge extracted from data sets to ultimately generate business values. We also found that data scientists are also responsible for specific missions such as ‘companies data systems’, ‘analytical team support’, ‘enterprise reports’ or ‘testing program.’

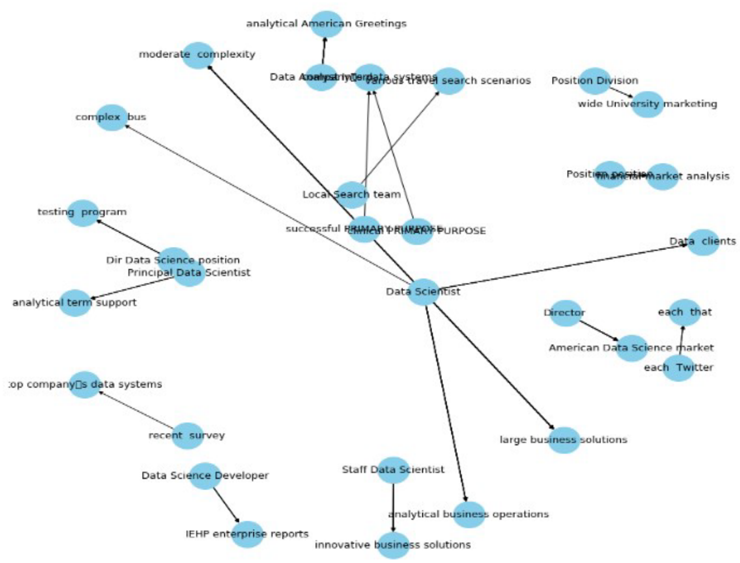


Figure 4: *Knowledge Subgraph with ‘Is Responsible’ Relation*

Another knowledge subgraph was identified using the ‘encompasses’ relation, anticipating that the knowledge subgraph would find multi-facts reading like ‘the position encompasses this task or skill’. This knowledge subgraph shown in Figure 5 is very unique in that, within this subgraph, all source nodes have ‘encompasses’ relation with single target node, ‘enquiry business decision making’ and each of the subject nodes includes ‘data science’ as a part of job title. This subgraph ultimately illustrates that data science jobs must encompass business decision making. While this finding may seem to be obvious, its implication is significant for students who desire to be a data scientist. In essence, data scientists and analysts must take their technical skills one step further to be able to utilize the information they have gained to make business decisions.

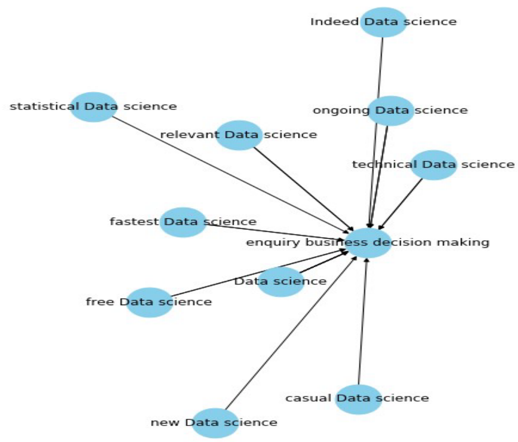


Figure 5: *Knowledge Subgraph with 'Encompasses' Relation*

Knowledge Subgraphs with Chosen Nodes

In this section, we created multiple knowledge subgraphs for source or target nodes that are associated with specific skills, which are different from knowledge subgraphs for specific relations. In particular, we created such knowledge subgraphs using 'Database Systems' for the course descriptions and 'looking' for the job descriptions as source or target nodes. We chose these subgraphs because database related skills have been regarded as one of most prominent analytical skills in DAIS curriculum, and we wanted to see if companies are looking for candidates with database skills. To create this kind of knowledge subgraphs, we have to first convert the complete KG into a non-directed graph and then use subgraph commands with chosen node names (e.g., 'Database Systems' or 'Databases').

Figure 6: *Course Description 'Database Systems' Subgraph*

We first represent the knowledge subgraph in Figure 6 that contain representative nodes connected to the ‘Database Systems’ node from the complete KG built from the course description data. We first noted that ‘Database Systems’ node has most connections and hence the thickest edge lines with ‘course’ node mainly because ‘(this) course is about Database Systems’. Other well-connected nodes to the ‘course’ node include database theory related nodes (e.g., ‘database theory’, ‘large scale database systems’ and ‘concurrency control’) and data analysis related nodes (e.g., ‘analytical modeling’, ‘(interactive) data analysis’, and ‘analysis models’,). However, we do not believe that ‘Database Systems’ courses directly deliver all the identified contents shown in Figure 6. Other skills related to ‘Database Systems’ node are most likely to be identified in this subgraph because there are many sentences relating a database course as a pre-requisite course for students to more advanced data analytic courses that focus on utilizing more advanced skills like ‘(object) oriented Java programming language’, ‘(data) mining business technologies’ and ‘data (analysis with) python’.

In the same way, a knowledge subgraph for the chosen node, ‘Looking’, from the complete KG of the job description data was also created. However, resulting knowledge subgraphs with chosen nodes from job description data is still very large due to the large number of associated nodes and edges. Therefore, we selected at maximum top 50 nodes determined by betweenness centrality measure that represents the number of shortest paths passing through the considered node. Note that the betweenness centrality measure is one of the most popular measures to determine the importance of nodes in graph theory and network analysis by measuring how important the node is to the flow of information through a network (Brandes 2001). We also varied the size of nodes based on values of betweenness centrality, making nodes larger as more shortest paths between other nodes pass them. The width of edges between nodes is also adjusted using the number of other nodes connected to the considered node. We present the final version of the knowledge subgraph in Figure 7.

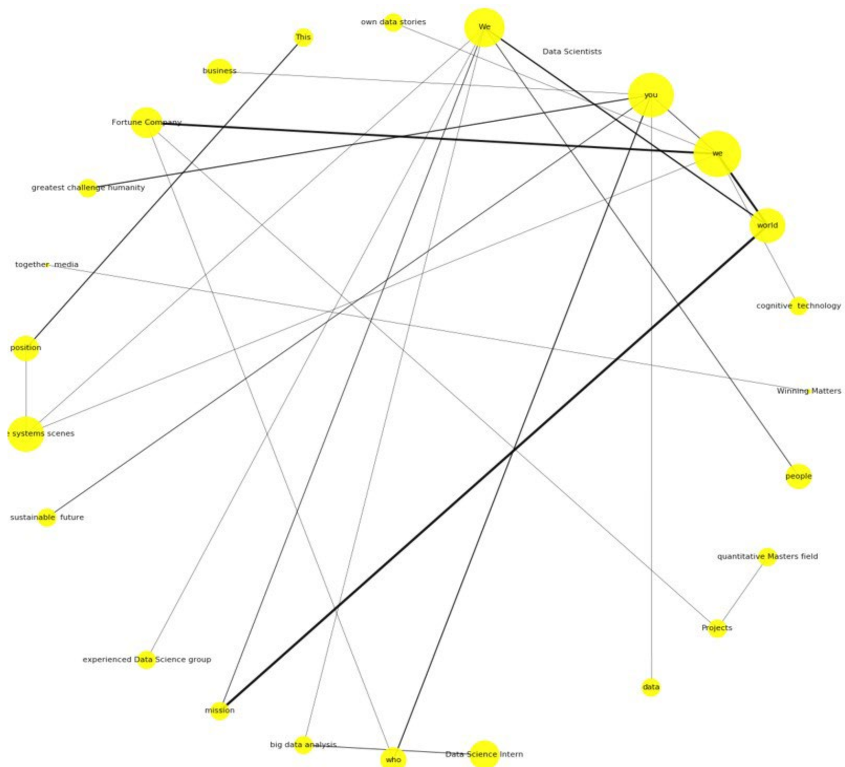


Figure 7: *Job Description 'Looking' Subgraph*

Unlike the course description subgraph, these nodes in Figure 7 are more interconnected with varying sizes of relation edges. Some of the connected nodes include pronouns such as 'you', 'we', 'Fortune company' as source entity (or subject) are linked to destination entity (or object) of 'data', 'data science intern', 'big data analysis', 'projects', and 'data scientist'. While many of the nodes would include work with a database, we were surprised to see that a node including the word 'database' was not displayed in the subgraph. However, like the course description subgraph, this shows that many skills will be connected, and data scientist need to be able to use all them together.

Limitations and Future Work

In this study, we compared core skills geared toward future data analysts and data scientists from two data sources: 50 DAIS course descriptions of colleges in the Intermountain West Region and 11,225 job descriptions of data scientists. According to word clouds from the two data sources, DAIS course descriptions present more general terms related to theories while job description sources illustrates more specific terms related to data usage, data types and analytical skills. In addition, from multiple KGs based textual course and job descriptions, we found that the job descriptions source presents the importance of application of specific skills to various business settings while the curriculum source focuses on the conceptual and theoretical skills along with fundamental skills such as SQL. Therefore, we suggest that if DAIS courses truly want to prepare students for being data scientists, they should include not only technical skills, but also soft skills to make informed decisions and turn them into business decisions through real application scenarios.

We admit that our course description data is limited, and therefore one of our immediate future research directions is to validate our findings based on more comprehensive DAIS curriculum data from multiple regions.

REFERENCES

Agarwal, R. & Lucas, H. C., Jr. 2005. “The Information Systems Identity Crises: Focusing on High-Visibility and High-Impact Research,” *MIS Quarterly* (29:3), pp. 381-398.

Benbasat, I., and Zmud, R. W. 2003. “The Identity Crisis within the IS Discipline: Defining and Communicating the Discipline’s Core Properties,” *MIS Quarterly* (27:2), pp. 183-194.

Bordes, A., and Gabrilovich, E. 2014. “Constructing and Mining Web-scale Knowledge Graphs: KDD 2014 Tutorial,” in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1967. https://doi.org/10.1145/2623330.2630803

Brandes, U. 2001. “A Faster Algorithm for Betweenness Centrality,” *Journal of Mathematical Sociology* (25:2), pp. 162-177.

Case, T., Dick, G., Granger, M. J., and Akbulut, A. Y. 2019. “Teaching Information Systems in the Age of Digital Disruption,” *Journal of Information Systems Education* (30:4), pp. 287-297.

Ehrlinger, L., and Wöß, W. 2016. “Towards a Definition of Knowledge Graphs,” In *Proceedings of the Posters and Demos Track of 12th International Conference on Semantic Systems – SEMANTiCS 2016.*

Elnagar, S., and Weistroffer, H.R. 2019. “Introducing Knowledge Graphs to Decision Support Systems Design,” In: Wrycza S., Maślankowski J. (eds) *Information Systems: Research, Development, Applications, Education*. SIGSAND/PLAIS 2019. *Lecture Notes in Business Information Processing*, vol 359. Springer, Cham.

Gomez-Perez, J. M., Pan, J. Z., Vetere, G., and Wu, H. 2017. “Enterprise Knowledge Graph: An introduction,” In: Pan, J., Vetere, G., Gomez-Perez, J., Wu, H. (eds.) *Exploiting Linked Data and Knowledge Graphs in Large Organisations*, pp. 1–14. Springer, Cham.

Jafar, M. J., Babb, J., and Abdullat, A. 2017. “Emergence of Data Analytics in the Information Systems Curriculum,” *Information Systems Education Journal* (15:5), pp. 22-36.

Lan, M., Xu, J., and Gao, W. 2016. “Ontology Feature Extraction via Vector Learning Algorithm and Applied to Similarity Measuring and Ontology Mapping,” *IAENG International Journal of Computer Science* (43:1), pp. 10–19.

Liu, J., Lu, Z., and Du, W. 2019. “Combining Enterprise Knowledge Graph and News Sentiment Analysis for Stock Price Prediction,” In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*, pp. 1247-1255.

Müller, O., Junglas, I., Brocke, J., and Debortoli, S. 2016. “Utilizing Big Data Analytics for Information Systems Research: Challenges, Promises and Guidelines,” *European Journal of Information Systems* (25:4), pp. 289-302.

Paulheim, H. 2016. “Knowledge Graph Refinement: A Survey of Approaches and Evaluation Methods,” *Semantic Web* (8:3), pp. 489–508.

Paulheim, H., and Bizer, C. 2014. “Improving the Quality of Linked Data Using Statistical Distributions,” *International Journal on Semantic Web and Information Systems* (10:2), pp. 63–86.

Pfaff, M., Neubig, S., and Krcmar, H. 2018. “Ontology for Semantic Data Integration in the Domain of IT Benchmarking,” *Journal of Data Semantics* (7;1), pp. 29–49.

Rashid, M.R.A., Rizzo, G., Torchiano, M., Mihindukulasooriya, N., Ccorcho, O., and Carcia-Castro, R. 2019. “Completeness and Consistency Analysis for Evolving Knowledge Bases,” *Web Semantics: Science, Services and Agents on the World Wide Web* (54), pp. 48–71.

Singh, M., Stenzhorn, H., Boeker, M., and Smith, B. 2009. “Strengths and Limitations of Formal Ontologies in the Biomedical Domain,” *Electronic Journal of Communication, Information and Innovation in Health* (3:1), pp. 31-45.

Singh, M. P., Agarwal, P., Chaudhary, A., Shroff, G., Khurana, P., Patidar, M., Bisht, V., Bansal, R., Sachan, P., and Kuma, R. 2018. “KNADIA: Enterprise KNowledge Assisted DIAlogue Systems using Deep Learning” in *Proceedings of 2018 IEEE 34th International Conference on Data Engineering (ICDE)*, pp. 1423-1434.

Yahya, M., Berberich, K., Elbassuoni, S., Ramanath, M., Tresp, V., and Weikum, G. 2012. “Natural Language Questions for the Web of Data,” In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Jeju Island, Korea, pp. 379–390.

Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., and Auer, S. 2016. “Quality Assessment for Linked Data: A Survey,” *Semantic Web* (7:1), pp. 63–93.

33.

Fostering a Mastery Mindset: Today's Statistics Students

Austin Swack

Hawkes Learning, Columbus, USA

Abstract

With an application-based curriculum, instructional simulations, technology instructions and relatable chapter projects, Hawkes empowers students to master complex concepts and make real-world connections. Learn how the intelligent courseware instills a mastery mindset with intelligent, error-specific feedback and step-by-step tutorials. Win one of THREE $25 Amazon gift cards!

Categories

Teaching, Innovation, Curriculum and Cases

34. Accepting, Rejecting, or Ignoring? Higher Education Institutes' Responses to ChatGPT

Jun He

University of Michigan-Dearborn, Northville, USA

Abstract

ChatGPT, due to its remarkable ability to engage in natural-sounding dialogue and carry out a wide range of tasks, presents a challenge to education. Given the widespread awareness and fast adoption of the new technology among students, higher education institutions must respond carefully, promptly, and decisively to the challenge. This study aims to examine how higher education institutions are responding to the challenge. The websites of 143 public universities in the Midwest of USA were examined to identify their institutional responses to ChatGPT. The findings will provide an overview of different types of responses from universities, and hopefully offer insights into how higher education institutions are addressing the emergence of a potentially disruptive new technology.

Introduction

ChatGPT, an artificial intelligence chatbot developed by OpenAI, has garnered significant attention and enthusiasm from the public and media alike. Within just two months of its late November 2022 launch, ChatGPT reached 100 million users, making it the fastest-growing consumer application ever (Hu, 2023). With its remarkable ability to engage in natural-sounding dialogue and carry out a wide range of tasks, ChatGPT has been lauded as a ground-breaking natural language processing technology, and the mainstream media has hailed it as the start of the artificial intelligence revolution. Capable of conducting tasks from online searches to data analysis to even composing new music (Marr, 2023), ChatGPT has the potential to profoundly impact our daily lives (Jackson, 2023).

While ChatGPT has generated excitement in the media, many educators view it with skepticism and concern about its potential impact on education. School districts across the United States, from the Los Angeles Unified School District to the New York City Department of Education, have blocked ChatGPT from their networks and devices, citing worries about plagiarism and misinformation (Jimenez, 2023). University professors have even found that ChatGPT is capable of passing graduate-level exams at prestigious institutions such as the law school of the University of Minnesota, Wharton School of Business at the University of Pennsylvania, and Stanford Medical School, leading some elite universities like Cambridge and Oxford University (Burnett, 2023), as well as Sciences Po in France (Reuters, 2023) and the University of Hong Kong (Yau and Chan, 2023) to ban its use out of fear of academic misconduct. However, other experts argue for ChatGPT's potential as a valuable teaching tool. Lipman and Distler (2023) question schools' decision of blocking ChatGPT, auguring that in today's technology-driven environment, it is important for students to learn how to use it ethically. Shields (2023), a senior high school English teacher, called for a shift towards teaching with ChatGPT rather than banning it altogether. Villasenor (2023), a law professor at the University of California Los Angeles, incorporated ChatGPT in his teaching and found it beneficial for improving students' writing skills. Notably, Princeton University has taken a proactive approach and provides clear guidance to faculty on the acceptable use of ChatGPT in the classroom (Spike, 2023). Young people are quick to adopt new technology (Morris and Venkatesh, 2000). In less than two months since its launch, ChatGPT had gained a significant following among college students, with 46% of them reporting awareness of the technology, and 30% having used it on written assignments. (Intelligent, 2023). This widespread awareness and fast adoption of the new technology among students highlights the need for higher education institutions to respond carefully, promptly, and decisively, given its potentially profound impact on education.

This study aims to examine how higher education institutions are responding to ChatGPT. To achieve this goal, we conducted a survey of university websites among public universities in the Midwest region of the United States. The findings will provide an overview of different responses to ChatGPT from universities, and hopefully offer insights into how higher education institutions are addressing the emergence of a potentially disruptive new technology.

Research Method and Results

To identify potential participants for our survey, we conducted a Google search of public universities in the Midwestern United States and found 143 institutions across 12 states, including both 4-year and 2-year colleges. These universities have a combined student population of more than 1.26 million as reported for the 2020-2021 school year. Table 1 shows the coverage of the surveyed universities. While it cannot be claimed that every Midwest public university was identified in the search, the sampled universities are likely representative of the main institutions in the region. Data compiled by the National Center for Educational Statistics (2022) show that during the 2020-2021 school year Midwest public universities reported a total of 1,590,736 enrolled full-time undergraduate students.

State	University Type	University Count	Student Size
IA	2-year	9	56,867
	4-year	2	30,113
IL	2-year	20	81,120
	4-year	10	117,727
IN	4-year	5	86,706
KS	2-year	8	30,415
	4-year	5	49,892
MI	2-year	4	39,275
	4-year	14	198,711
MN	2-year	1	680
	4-year	12	84,173
MO	2-year	5	20,901

	4-year	8	64,403
ND	2-year	1	2,783
	4-year	4	22,989
NE	2-year	1	3,021
	4-year	4	38,616
OH	2-year	2	17,378
	4-year	12	190,942
SD	4-year	6	22,293
WI	2-year	1	3,955
	4-year	9	102,727
Grand Total		143	1,265,687

Table 1. Sampled Midwest Public Universities

A keyword search for "ChatGPT" was performed on the websites of the sampled universities. The search results were updated as of February 20, 2023. In the analysis, university-provided content was reviewed to determine institutional reactions to the technology; other content, such as individual blogs or course syllabi, was excluded. The resulting reactions are summarized in Table 2.

	Findings	Counts	Percentage	Total
University Resources	Dedicated webpage, portal, or archive	43	30.1%	100%
	No Resources Provided	100	69.9%	
University Policy	Ban	3	2.1%	100%
	Allow	10	7.0%	
	Not Taking Sides	16	11.2%	
	No policy	114	79.7%	
University Policy Developed By	Administration (the Provost office)	1	0.7%	20.3%
	HR/Faculty Development	4	2.8%	
	Teaching Center	20	14.0%	
	Writing Center	1	0.7%	
	Library	3	2.1%	
University-Sponsored Faculty Discussion	Workshop/Panel Discussion/Webinar	32	22.4%	100%
	No Open Discussion	111	77.6%	

Overall	Reaction from the University	44	30.8%	100%
	No Reaction from the University	99	69.2%	

Table 2: University Reactions to ChatGPT

Out of the 143 Midwest public universities sampled from 12 states, 43 have dedicated resources in the form of a special webpage, portal, or archive for ChatGPT. The other 100 universities do not provide any university resources on the topic, leaving individual faculty members to learn about ChatGPT on their own without institutional support.

In terms of university policy on ChatGPT, three universities have explicitly banned its use, with one labeling its use as plagiarism. Ten universities encourage faculty members to consider incorporating ChatGPT in their teaching, and offer guidance on syllabus templates and assignment design. Sixteen universities have taken no stance on the issue and leave it up to individual faculty members to decide whether or not to allow the use of the technology in their classes. Shockingly, 114 out of the 143 sampled universities (or close to 80%) provide no policy nor guidance on the technology. The inconsistent and troubling reactions among the sampled universities confirm a common observation that "Many educators are struggling with the question of whether or not to incorporate ChatGPT into their curricula or to ban it outright" (Intelligent, 2023).

Out of the 29 universities with policies or guidance on the use of ChatGPT, the majority of the policies are aimed at assisting faculty members in incorporating the technology into their teaching. The policies are primarily developed by the university's teaching center (20 out of 29). A few policies are formulated by the university administration, HR department, or faculty career development program (5 out of 29). Only 4 universities have framed their policies under the writing center or university library system, which makes the policies accessible to students.

Additionally, the search found that 32 universities had taken measures to educate faculty members about ChatGPT and encouraged open discussions among them, including seminars, workshops, discussion panels, or online webinars. For the other 111 universities, however, no university-sponsored avenues were provided for sharing information or discussing the topic of ChatGPT.

Overall, out of the 143 sampled universities, only 44 provided any reaction or guidance regarding ChatGPT, while the remaining 99 universities did not appear to address the challenge posed by the technology. Searches on their websites generated either "No Results" or led to personal webpages, blogs, or course syllabi.

It could be argued that some universities, with limited resources, may not have the means to address this emerging and unexpected challenge. Therefore, it would be intriguing to explore the potential relationship between a university's available resources and its reaction to ChatGPT.

Student body size may serve as an indicator of a university's resources. College Board (2023) groups universities into three categories based on size: small (under 5,000 students), medium (between 5,000 and 15,000 students), and large (over 15,000 students). To ensure consistency across all sampled universities, which include both 4-year and 2-year institutions, we use the size of enrolled undergraduate students during the 2020-2021 school year as our metric. Based on the distribution of actual student sizes, we further divided the sampled universities into six categories, as outlined below:

Very Small	Small	Medium	Medium Large	Large	Very Large
<1,500	1,500 – 5,000	5,000-10,000	10,000-15,000	15,000-25,000	>25,000

Table 3: University Size by the Size of Undergraduate Student Body

To examine the relationship between university size and their reactions to ChatGPT, Table 4 breaks down the sampled universities' responses based on their student body size categories.

University Size	Counts	Reaction Types					Overall	
		Providing Resources	Policy: Ban	Policy: Allow	Policy: No Sides	Discussion	Reaction	No Reaction
Very Small	23	0	0	0	0	2	2 (9%)	21 (91%)
Small	38	0	0	0	0	4	4 (11%)	34 (89%)
Medium	37	7	1	0	6	7	11 (30%)	26 (70%)
Medium Large	18	4	0	2	2	5	5 (28%)	13 (72%)
Large	18	12	2	6	4	7	14 (78%)	4 (22%)
Very Large	9	6	0	2	4	7	8 (89%)	1 (11%)
Total	143	29	3	10	16	32	44 (31%)	99 (69%)

Note: numbers in parentheses are percentages of the according category

Table 4: Pattern of University Responses to ChatGPT

Table 4 demonstrates a clear correlation between the size of a university and its response to ChatGPT: small universities are less likely to take action, whereas larger universities tend to allocate resources and provide assistance for faculty to address the challenge. The correlation coefficient between the actual student size of a university and its response to ChatGPT was 0.515 ($p<0.001$), indicating that university resources play a significant role in determining whether to respond to this challenge.

Given that universities vary in size, merely counting the number of institutions may not provide a comprehensive understanding of how higher education institutions are responding to ChatGPT. We should also examine the total student population enrolled in universities with and without reactions. Of the 1,265,678 students included in the sample, 582,996 (46.1%) were studying in universities that had not reacted to the ChatGPT challenge, while 682,691 students (53.9%) were studying at universities that had taken some form of action to address it.

Discussion

This study aims to examine how higher education institutes respond to the emerging challenge of ChatGPT. A total of 143 public universities in the Midwestern United States were sampled. Their websites were searched with the keyword “ChatGPT” for university responses.

Overall, the search found that many public universities in the Midwest, especially that small to medium in size, had not responded to the ChatGPT challenge. Given the profound impact of ChatGPT on education and the increasing number of students using the technology, the response of higher education institutes, at least in the Midwest, is falling behind. What is more concerning is that a recent survey has revealed that most students (75%) who use ChatGPT for written assignments believe it to be cheating, yet continue to use it regardless. Without proper guidance or control from educators, the fairness of education is at risk. It is crucial for educators to respond to this emerging challenge and develop strategies to guide students in the proper use of AI technology both responsibly and ethically.

Another finding is that, when universities take action in response to ChatGPT, their main focus is on supporting faculty, with students often being overlooked. Table 2 reflects this trend, as the majority of allocated resources are provided by the university teaching center. The negligence of students must be addressed. As more students begin to use ChatGPT, they must be clearly informed of any university policy on the new technology and be guided on the properly handlings. To safeguard academic integrity, universities should take a proactive approach to assure that ChatGPT and similar AI tools are used responsibly and ethically.

In addition to examining institutional responses to ChatGPT, this study has also investigated the relationship between university size and these reactions. The findings suggest that small universities are less likely to take proactive measures in response to the emergence of this powerful and disruptive technology. This is concerning. While limited resources may constrain universities' ability to provide support to faculty and students, it is imperative that they acknowledge and address this challenge. One strategy is to promote open discussions among faculty to exchange information on the technology, share experiences, and develop effective practices that can be adapted to their classes. By fostering an open and collaborative environment, universities can better equip themselves to manage the impacts of ChatGPT and other emerging technologies.

References

Burnett T (2023) Cambridge University among elite universities to ban ChatGPT due to plagiarism fears. Available at: https://www.cambridge-news.co.uk/news/cambridge-news/cambridge-university-among-elite-universities-26361465 (accessed 3 March 2023).

College Board (2023) Understand college campus and student body size. Available at: https://bigfuture.collegeboard.org/plan-for-college/college-basics/types-of-colleges/understand-college-campus-student-body-size (accessed 3 March 2023).

Hu K (2023) ChatGPT sets record for fastest-growing user base - analyst note. Reuters. Available at: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/ (accessed 3 March 2023).

Intelligent (2023) Nearly 1 in 3 college students have used ChatGPT on written assignments. Available at: https://www.intelligent.com/nearly-1-in-3-college-students-have-used-chatgpt-on-written-assignments/ (accessed 3 March 2023).

Jackson L (2023) How A.I. can help. The New York Times. Available at: https://www.nytimes.com/2023/03/02/briefing/chatgpt-ai.html (accessed 3 March 2023).

Jimenez K (2023) 'This shouldn't be a surprise' The education community shares mixed reactions to ChatGPT. USA Today. Available at: https://www.usatoday.com/story/news/education/2023/01/30/chatgpt-going-banned-teachers-sound-alarm-new-ai-tech/11069593002/ (accessed 3 March 2023).

Lipman J and Distler R (2023) Schools shouldn't ban access to ChatGPT. Time. Available at: https://time.com/6246574/schools-shouldnt-ban-access-to-chatgpt/ (accessed 3 March 2023).

Marr B (2023) The best examples of what you can do with ChatGPT. Forbes. Available at: https://www.forbes.com/sites/bernardmarr/2023/03/01/the-best-examples-of-what-you-can-do-with-chatgpt/?sh=6868094cdf11 (accessed 3 March 2023).

Morris MG and Venkatesh V (2000) Age differences in technology adoption decisions: Implications for a changing work force. *Personal Psychology* 53(2): 375-403.

National Center for Education Statistics (2022). Undergraduate Enrollment. Condition of Education. U.S. Department of Education, Institute of Education Sciences. Data retrieved on March 5, 2023 from https://nces.ed.gov/programs/coe/indicator/cha.

Reuters (2023) Top French university bans use of ChatGPT to prevent plagiarism. Available at: https://www.reuters.com/technology/top-french-university-bans-use-chatgpt-prevent-plagiarism-2023-01-27/ (accessed 3 March 2023).

Shields C (2023) Don't ban ChatGPT. Use it as a teaching tool. Education Week. Available at: https://www.edweek.org/technology/opinion-dont-ban-chatgpt-use-it-as-a-teaching-tool/2023/01 (accessed 3 March 2023).

Spike C (2023) Princeton memo directs faculty on acceptable use of ChatGPT and other AI. Princeton Alumni Weekly. Available at: https://paw.princeton.edu/article/princeton-memo-directs-faculty-acceptable-use-chatgpt-and-other-ai (accessed 3 March 2023).

Villasenor J (2023) How ChatGPT can improve education, not threaten it. Scientific American. Available at: https://www.scientificamerican.com/article/how-chatgpt-can-improve-education-not-threaten-it/ (accessed 3 March 2023).

Yau C and Chan K (2023) University of Hong Kong temporarily bans students from using ChatGPT, other AI-based tools for coursework. South China Morning Post. Available at: https://www.scmp.com/news/hong-kong/education/article/3210650/university-hong-kong-temporarily-bans-students-using-chatgpt-other-ai-based-tools-coursework (accessed 3 March 2023).

37. Cybersecurity measures from new technology users' perspective

In the cloud computing environment, cyber vulnerability is a critical issue that poses a threat to organizations worldwide. Modern economies and businesses rely heavily on new technology for daily operations, making the need for effective cybersecurity measures more critical than ever before.

To protect information and systems from cyber-attacks, businesses allocate a significant portion of their annual technology budgets to security measures. An example is measures that deal with software vulnerability, flaws within a software product that can lead to work contrary to its documented design, which can be exploited to cause the system to violate its documented security policy. Studies have shown that software vulnerabilities can have a widespread impact, resulting in significant economic and noneconomic damage to firms.

The annual costs of all malicious code attacks to businesses worldwide run into billions of dollars and continue to increase each year. These costs include additional elements such as lost business revenues, cleanup costs, productivity losses, overtime costs, and damage to reputation and trust.

In addition to software, users play a critical role in ensuring cybersecurity. As more users become more dependent on network-enabled services for communications, social networking, and web-based transactions, the number of cyber-attacks and the ease with which they occur have increased, putting more systems at greater risk. Therefore, prioritizing cybersecurity measures to help users navigate the new technological landscape is crucial for securing today’s assets. This paper aims to examine the human aspects of cybersecurity, including user characteristics and user perception of cybersecurity measures.

MIDWEST DECISION SCIENCES INSTITUTE

61. Bolt Socks – A Class Project on Process Mining and Robotic Process Automation

Cody Baldwin

University of Wisconsin-Madison, Madison, USA

ABSTRACT

The following paper outlines a class project that teaches university students how to understand and improve business processes using (1) process mapping, (2) process mining, and (3) robotic process automation. During the project, students imagine they own a consulting company, and that they have been hired by Bolt Socks, a company that manufactures and sells wholesale socks to retail stores. Students are tasked with understanding and improving their "order-tocash" process, which is critical in many businesses. Additionally, while working on the project, students work with popular software tools and a real-world dataset. The primary educational theory underpinning this project is that of constructionism, which states that students learn best when they create something and share it with others.

KEYWORDS: Process Mining, Robotic Process Automation, Business Analytics, Project-Based Learning

INTRODUCTION

This paper outlines a class project that teaches students to understand and improve business processes using (1) process mapping, (2) process mining, and (3) robotic process automation. The project is used with business analytics students at the University of Wisconsin-Madison, but it could be used across a number of programs and courses – such as engineering, accounting, operations, data science, lean six sigma, etc.

As context, *processes* are repetitive work in our business. They are different from *projects*, which are temporary endeavors. Every business has hundreds – sometimes thousands – of processes. Examples include hiring employees, paying suppliers, repairing equipment, and handling returns. With any process, two primary goals should always be to *understand* and *improve* them. In other words, we seek to understand their "as is" state and get them to a better "to be" version. Process mapping and process mining are tools for gaining understanding and robotic process automation is a tool for making improvements.

The following list provides learning objectives for this student project:

- Create a process map to understand and communicate steps in a process.
- Mine process event logs for insights, issues, and improvement opportunities.
- Automate several process steps using robotic process automation (RPA).
- Communicate process insights and recommendations to business stakeholders.

SUPPORTING THEORIES

The research underpinning this student project includes that of constructionism, as outlined by Seymour Papert, and the taxonomy of educational learning objectives, as initially created by Benjamin Bloom (often referred to as Bloom's taxonomy).

First, according to constructionism, student learning increases when they are involved in building something. (In this project, students create process maps and process automations.)

Papert states that this learning approach is especially successful when the "learner is consciously engaged in constructing a public entity" (Papert, 1991). Thus, it is a social process, and students should work together on their creations and share them in a public way. (Within this project, students work as teams to create their process map. Also, at the conclusion of the project, students summarize and present their findings.) Ultimately, Papert was interested in how social interactions and "conversations [about student creations] boost self-directed learning" (Ackerman, 2001). Thus, the teacher is no longer responsible for "pushing" knowledge to the student. The student learns (and desires to learn) on their own as they make something that interests them and share it with their classmates. Papert believed, "students best learn when engaged over long periods of time in the construction of personally meaningful products - or products they truly care about" (Ackerman, 2010). (During the project, students have plenty of freedom as they investigate issues and opportunities. Ultimately, they get to determine areas of focus. Process improvement recommendations often vary across students.) Finally, the following quote, which has been credited to Confucius (450BC), sums up the theory of constructionism in a different way: "Tell me, and I will forget. Show me, and I may remember. Involve me, and I will understand."

Second, Bloom, in his now famous "taxonomy of educational objectives," classifies student thinking (Bloom, 1956). It is often visualized as a pyramid, where higher levels indicate higher levels of thinking. In a revised version of Bloom's taxonomy, the peak of learning objectives is to "create" (Anderson, 2001). This assignment would meet that objective. It is more than an analysis. Students create recommendations for the business as they map, mine, and automate the process. Those recommendations are function of several inputs – an interview, process data, external research, their own experience, etc.

PROJECT BACKGROUND

On the project, students imagine they own a consulting company, and that they have been hired by Bolt Socks, a company that manufactures and sells socks to retail stores. Students are tasked with understanding and improving their "order-to-cash" process, which is critical in many businesses. The company wants to address cash flow challenges they face. In other words, they hope to get paid faster.

There are four parts to the project, which are outlined in Table 1 along with the software tools that are used during each part.

Table 1: *Project Activities and Software Tools*

#	Project Activity	Software Tool
1	Process Mapping	Lucidchart
2	Process Mining	Celonis
3	Process Automation	UiPath
4	Summary & Recommendations	Microsoft PowerPoint

For each software tool listed, free licenses are made available to our students. Those licenses provide sufficient access to complete all activities in the project. Also, as a note, the chosen tools were selected based on several factors: (1) market share, (2) market research (e.g. Gartner magic quadrants), (3) student and employer feedback, and (4) the availability of free licenses and vendor-provided training materials.

BUSINESS PROCESS CONTEXT

During the project, students investigate the "order-to-cash" process at Bolt Socks. The process begins when a customer places an order and finishes when cash is paid for that order. It is a critical process in many businesses. They need cash to operate and invest in the future. Prior to beginning the project, students learn more about an order-to-cash process more generally and explore why it's important in a business. For example, they answer questions such as: *Why is cash critical in business? Why would we extend credit?* This general understanding improves the quality of their investigation later. It also helps students compare this process at Bolt Socks with others across industry.

Process Interview

As one source of process context, students are provided with an interview from Jake Lendale (shown below), the Founder of Bolt Socks. It is his recollection of the process. However, on a typical consulting project, students would interview several parties to better understand the process. It is best to get a variety of perspectives – especially from those who use the process.

> *"At Bolt Socks, we make really cool socks. That's basically all we do. We have great designs – and a high-quality product. The socks have added support in the toes – so they don't rip easily. They also stay up, so people don't have to pull at them all the time. And we sell our socks directly to businesses – mainly retail chains. Sometimes they buy 100 pairs; other times they buy 10,000. Overall, it's been a successful business for us.*
>
> *"Now, the reason why we brought you in was to review our order-to-cash process. We've had some cash flow problems recently and need to identify the issues. So, here's the best way I could describe that process… When a customer needs socks, they go to our website to make an order. They could also contact a sales rep directly, but I don't think that happens very much anymore. (We're living in a digital age!) Once the order is made, we send the customer a sales order that includes the terms for the credit. But, if they're a new customer, we typically perform a credit assessment prior to that, and that's done by an outside agency that we pay. Then, our warehouse staff pick and pack the order. From what I understand, that typically takes 24 hours, but I'm not positive. (I hope it doesn't take longer than that. We know that our customers care a lot about timeliness.) The next step is to ship the order. That's done by US Express. They stop by the warehouse each day at 4PM. (In the future, my hope is that we can have other shippers at our disposal. I think we could get better rates.) Now, finally, once we confirm that the delivery was made, we send an invoice. Typically, we want them to pay within 14 days, but I'm not actually sure how often than happens. So, there you have it. That's the process. Thanks for looking into this for us. We need your help!"*

Process Data

As another source of process context, students are given a process event log by Jake Lendale. At Bolt Socks and many other companies, the order-to-cash process (among others) is tracked using an enterprise resource planning tool. In this case, it is SAP. The tool produces event logs that have at least these three data points about each process step: (1) a case identifier, which uniquely identifies an order; (2) an activity, which is an individual step in the process; and (3) a timestamp, which shows when the activity occurred. In other words, it shows the *who*, *what*, and *when* of the process. Process mining requires this level of data. Table 2 below shows a sample of the event log that is provided to students.

Table 2: Process Event Log (Sample)

case_id	Activity	timestamp
2208767001	Sales Order Created	12/09/17 00:12:46
2208767001	Picking Done	12/11/17 11:02:22
2208767001	Confirmed Delivery Date	12/14/17 08:03:54
2208767001	Customer Pick-Up	12/14/17 15:40:38
2208767001	Shipment Sent	12/14/17 15:52:42
2208767001	Delivery Completed	12/14/17 21:45:18
2208767001	Invoice Created	12/15/17 23:42:47
2208767001	Payment Received	12/17/17 08:22:39

PROCESS MAPPING

In the first part of the project, students create an "as is" (or current state) process map using the interview with Jake Lendale. Process mapping is a key first step in understanding a process. It shows what the participants – in this case, Jake Lendale – *think* is occurring in the process. This may differ from what is *actually* occurring. (This will become clearer in part two of the project – process mining. It uses data to automatically generate the process map.)

Students create two versions of the process map – first, a paper version (sticky notes on a wall), and then, a digital version in Lucidchart. When creating the paper version, students work together with teammates. (This team activity aligns with constructionism, and its emphasis on working together on a creation. Additionally, within an actual business, involving those who participate in the process in the mapping is ideal, since they are most familiar with it. It also helps them feel involved in your process improvement efforts, which can make the less resistant to changes later.) This paper version of the process map provides students with flexibility to move steps around without too much difficulty. Afterwards, students work independently to create the digital version of their process map in Lucidchart.

When creating the process map, students are instructed on common process mapping symbols as shown in Figure 1. Students are also taught several best practices for process maps including, but not limited to, the following: (1) activities should begin with a verb; (2) activities should have at least one arrow in and only one arrow out; (3) decisions should have "yes" and "no" outputs; and (4) the process should have "start" and "end" terminators.

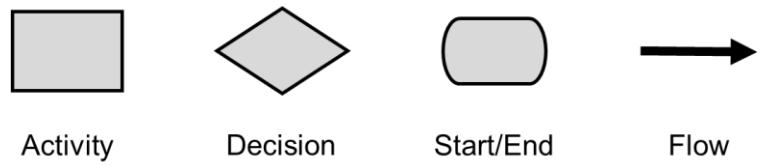


Figure 1: *Common Process Mapping Symbols*

Figure 2 below shows what a completed process map might look like for this project. During class, students have an opportunity to share their completed, digital version with their classmates. (This again aligns with the principles of constructionism.) As a note, when creating the process map, students might be tempted to propose improvements. Those ideas should be written down somewhere, but they should not inform the "as is" version of the process. It is important to have a clear, accurate "as is" process map, so students understand the gap between the "as is" and the "to be" versions of the process.

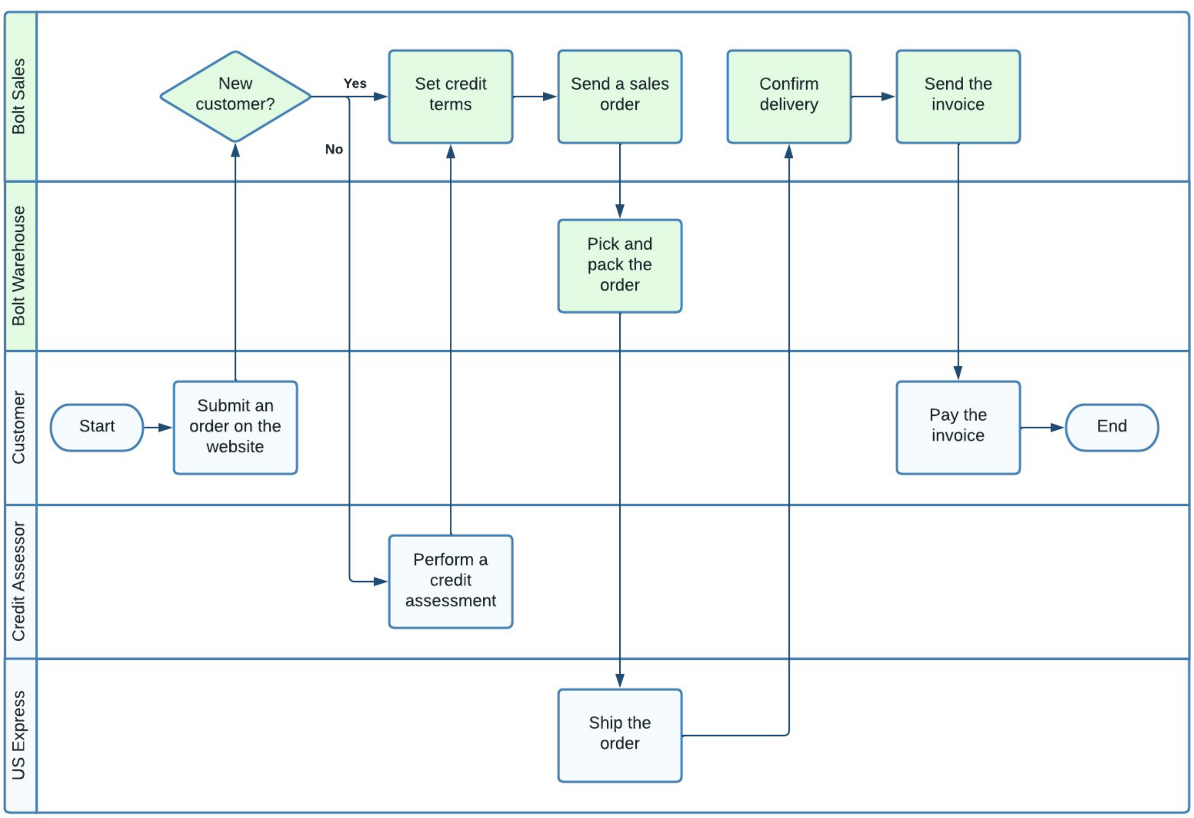


Figure 2: Completed Process Map (Lucidchart)

PROCESS MINING

In the second part of the project, students mine the process using Celonis, a popular tool in the marketplace. Process mining is another important step – in addition to traditional process mapping – for understanding the process and identifying opportunities for improvement. Process mining tools use event logs, which include data about each step in a process. (See the sample event log shown previously in this paper.) Process mining tools take those event logs and automatically generate process maps, which can be explored and investigated for insights. Figure 3 below summarizes the inputs and outputs of process mining.

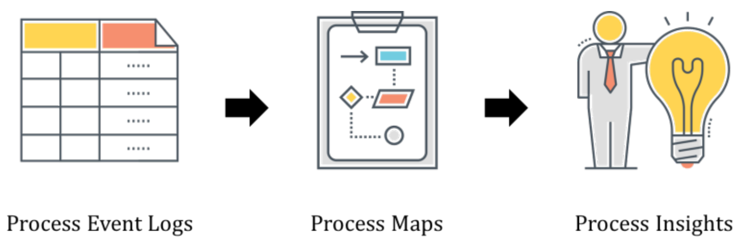


Figure 3: *Process Mining – Inputs and Outputs*

Process mining is an objective, data-driven activity. It can show what is *actually* occurring in a process versus what someone *thinks* is occurring. After students use Celonis to generate the process map, they compare their process map from Lucidchart with the process map generated in Celonis. Almost immediately, students can see that the actual process is different than how Jake Lendale understands it.

Figure 4 is the initial process map this is generated in Celonis using the order-to-cash data. By default, it only includes the most common variations ("variants") of the process, but students can add more to understand the complexity of the process. Once the map is generated, students have countless ways to drill-down into the map to gather process insights.

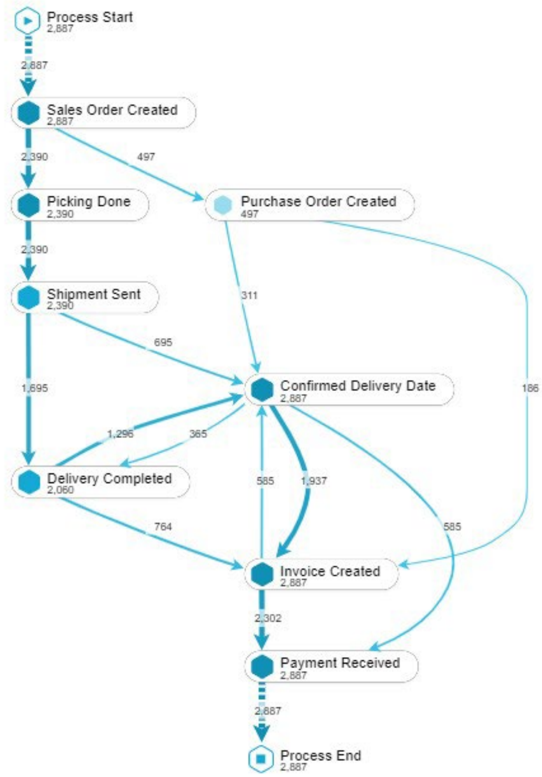


Figure 4: *Auto-Generated Process Map (Celonis)*

Important outputs of process mining are insights that can be used to understand and improve the process. In a way, students act like a detective. They are provided with a few questions to spark their investigation, but they are only intended to spur thought not to provide a clear "trail of breadcrumbs." Some of the insights could be related to (1) throughput, (2) complexity, (3) rework, and (4) deviations. Example insights in these categories are provided in the following sub-sections. However, the insights are not limited to these categories.

Process Insight – Throughput

Students can calculate process duration and explore steps that are taking longer or shorter than expected. Also, as shown in Figure 5 below, using the data from the order-to-cash process, you can see the distribution of process times. The median throughput time is 24 days (567 hours). Additionally, the fastest order was 1 day, and the slowest order was 232 days.

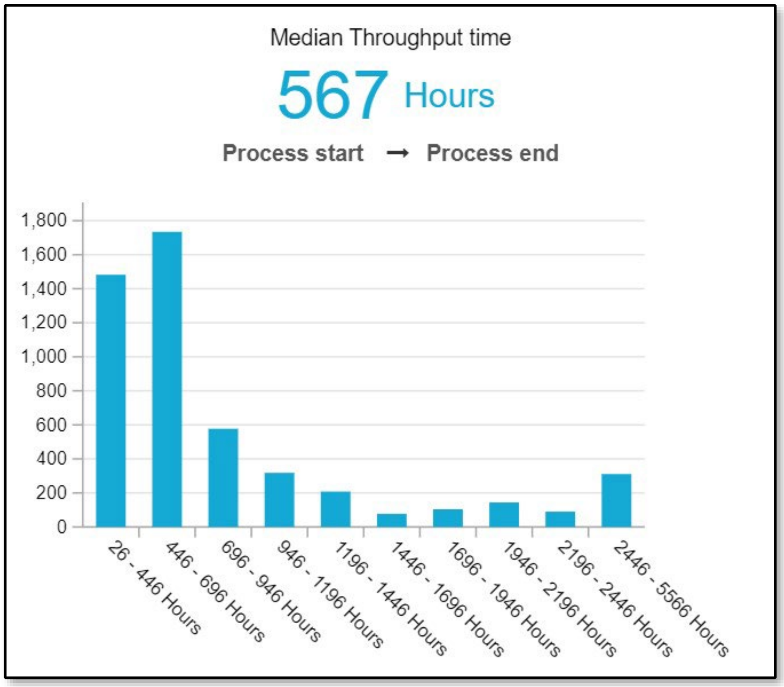


Figure 5: Process Throughput Metrics and Distribution (Celonis)

Process Insight – Complexity

Students can explore complexity in the process and examine all variants. Figure 6 below from Celonis shows all the process variants, not just the most common ones as represented previously. In this order-to-cash process, there are 68 different variants, but the most common one represents 26% of orders and the top 5 represent 54% of orders. This complexity is an interesting insight, since, by nature, we tend to think processes are simple. They are often much more complex. Given this, there could be a few interpretations of a process – what we *think* it is, what it *actually* is, and what it *should* be.

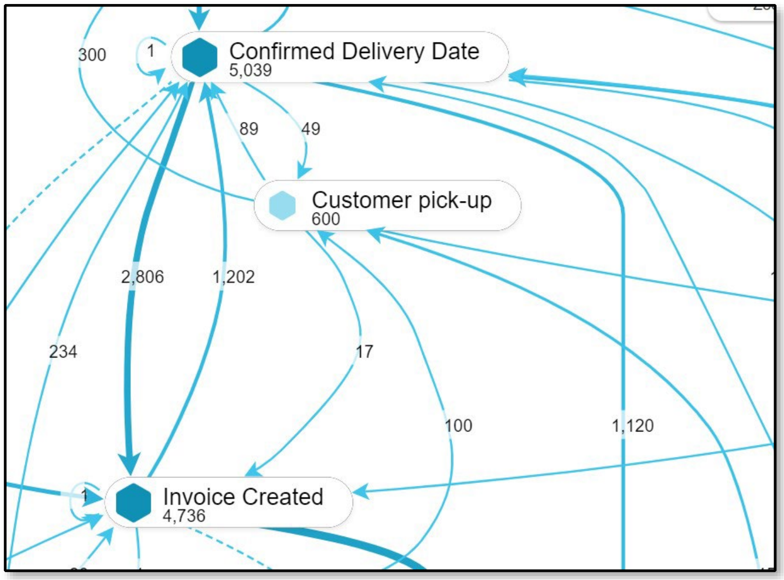


Figure 6: Process Variants (Celonis)

Process Insight – Rework

Students can evaluate rework or unnecessary revisions, which cause delays and increase costs. In other words, it can thwart efforts to get paid faster. In the project data, 12% of orders have an activity that was executed more than once. If we dive deeper into reworked activities as seen in Figure 7, we find that hundreds of orders had to be picked more than once.

Case ID	# Activities	# Rework Activities
67798752017	14	7
87288982017	14	7
70701922017	12	6
10050693006	11	5
48555671001	9	1
48555671002	9	1
48555671003	9	1
48555671004	9	1
48555671005	9	1
48555671006	9	1

Figure 7: Process Rework Activities (Celonis)

Process Insight – Deviations

Students can monitor deviations from the standard process (as defined by the company) and report on "violations" of that standard. As an example, which is shown in Figure 8, 12% of orders include customer pick-up at the warehouse. This could complicate order processing, if part of the order is shipped and part is picked up. It also requires more staff at the warehouse. Another example is that "confirmed delivery date" precedes "shipment sent", which could potentially a data quality issue. (The company should not be confirming a delivery that has not yet been sent.)

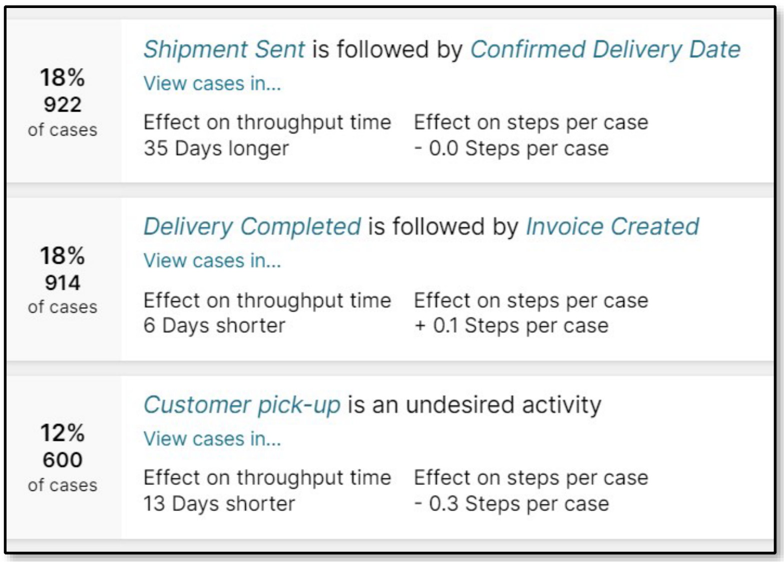


Figure 8: *Process Deviations and Violations (Celonis)*

Following their investigation, students summarize key insights in one presentation slide for Jake Lendale. Most of the listed insights should be related to the business question – the need to get paid faster. This presentation slide allows students to show progress to Bolt Socks. Additionally, their research helps them identify automation opportunities (part three of the project) and serves a starting part for their summary presentation to Bolt Socks (part four).

ROBOTIC PROCESS AUTOMATION

In the third part of the project, students automate one or more steps in the order-to-cash process. Automation opportunities are often identified during process mapping (part one) and process mining (part two). As background, robotic process automation (RPA) involves automating repetitive tasks that users perform in a graphical user interface (GUI). It uses software "bots" and not human robots. It also typically involves little or no coding. Example tasks that could be automated include opening e-mail and attachments, filling in forms, moving files and folders, logging into web applications, and so on. These may seem like trivial tasks, and to some extent, that is the point. We can automate mundane tasks and offer people more meaningful work, which can increase employee satisfaction and reduce turnover.

As an important note, RPA is one of many possible improvement methods. Lean six sigma outlines a number of additional techniques. Additionally, students are advised that it would be preferable to eliminate process steps rather than to automate them. (On occasion, they seek to dive directly into automating steps.)

Within the project, students are asked to consider the following opportunity, which was identified while process mining: it takes more than a day after a completed delivery to create and send an invoice. Bolt Socks could automate this process and, hopefully, get paid a day faster on average. This may not seem like a dramatic change, but Bolt Socks may have thousands of orders per year and additional tasks might be easily automated beyond these initial steps. Figure 10 summarizes the activities to be automated. Students build this automation in UiPath.

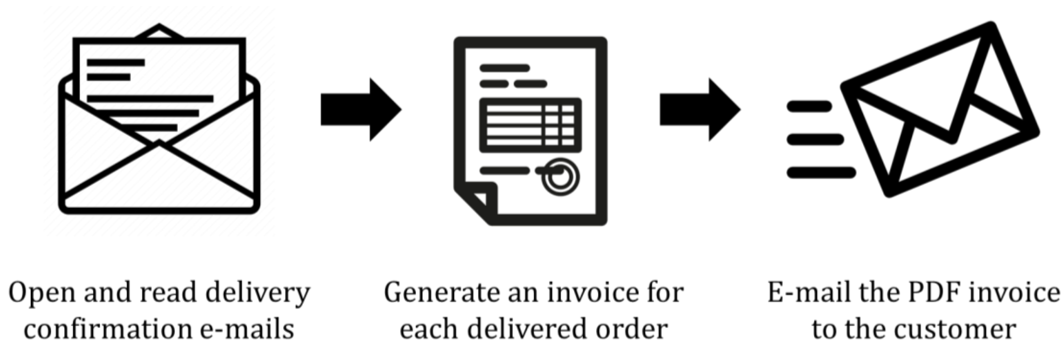


Figure 10: Invoice Creation Steps

SUMMARY AND RECOMMENDATIONS

In the final portion of the project, students create a summary presentation for Bolt Socks. This is arguably the most important part of the project work. Students could have incredible ideas, but if they can't articulate the value of those ideas, their efforts may be futile. Their suggestions may never be read or implemented. As students develop their presentations, they seek to achieve the following objectives:

- The presentation must show students clearly understand the business problem.
- The presentation must include their recommendations for addressing that problem.
- The presentation must provide data that supports their recommendations.
- The presentation must explain how to manage process changes in the company.

Within the presentation, students are directed to prioritize recommendations that are "lowhanging fruit" (low effort, high value). These small, fast projects can build momentum and excitement. Additionally, students are encouraged to perform additional outside research – beyond what was seen in the interview or process data. This can provide meaningful comparisons to other companies in their industry.

CONCLUSION AND NEXT STEPS

Up to this point, student feedback on the project has been excellent. However, the feedback has largely been qualitative. More quantitative analysis of student outcomes is needed. This analysis will be the subject of future experiments and publications. Additionally, another future consideration is to provide students with data that requires more cleaning and transformation. This is a common, time-consuming challenge in the real world.

REFERENCES

Ackermann, E. (2010), “Constructivism(s): shared roots, crossed paths, multiple legacies,” *Proceedings of Constructionism 2010*.

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.6201&rep=rep1&type=p df. Last accessed on 13 March 2023.

Ackermann, E. (2001), “Piaget’s constructivism, Papert’s constructionism: what’s the difference?” *Future of Learning Group Publication*, 5(3), 438. http://learning.media.mit.edu/content/publications/EA.Piaget%20_%20Papert.pdf. Last accessed on 13 March 2023.

Anderson, L. and Krathwohl, D. (2001), “A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives,” Allyn and Bacon.

Bloom, S.; Engelhart, M.; Furst, E.; Hill, W.; and Krathwohl, D. (1956), “Taxonomy of educational objectives: The classification of educational goals.” *Handbook I: Cognitive Domain*, New York, David McKay Company.

Papert, S. and Harel, I. (1991), “Preface, Situating Constructionism in Harel & S. Papert (Eds),” *Constructionism, Research Reports and Essays,* 1985-1990, New Jersey, pp. 1. http://www.papert.org/articles/SituatingConstructionism.html. Last accessed on 13 March 2023.

Index

Midwest Decision and Sciences Institute

Conference Sponsors

BALL STATE UNIVERSITY
Miller College of Business

PURDUE UNIVERSITY
Mitchell E. Daniels, Jr.
School of Business

YOUNGSTOWN STATE UNIVERSITY
Williamson College of Business Administration

HAWKES LEARNING

BALL STATE UNIVERSITY
WE FLY