Identifying Essential Factors for Deriving Value from Big Data Analytics in Healthcare

Brenda Eschenbrenner(✉)

University of Nebraska at Kearney, Kearney, NE 68845, USA
eschenbrenbl@unk.edu

Abstract. Big data analytics is emerging in many industries and is a prominent undertaking in healthcare. The healthcare industry has more opportunities now to garner insights and advantages from data than ever before. Big data analytics has the potential to enhance many aspects of the industry from enhancing quality of patient care to revenue cycle improvements. However, successfully leveraging big data analytics poses challenges as well, such as data standardization and integrity. Therefore, it will be important to identify the essential factors that will facilitate the ability to derive the maximum value from big data analytics for such endeavors. This research proposes to identify these pivotal factors of big data analytics in healthcare by utilizing value-focused thinking (VFT). VFT will entail interviews with both healthcare data analysts and management to identify these important factors. The findings can provide guidance to practitioners when considering the essential factors for big data analytics success, as well as provide topics for future research.

Keywords: Big data analytics · Value-focused thinking · Healthcare analytics

1 Introduction

Big data analytics has the potential to provide increased, improved, and more salient insights that were previously unfathomable. With the increased usage of and continuous improvements in technology and its capabilities, utilizing massive amounts of data to improve healthcare services is now possible and being explored [1]. The healthcare industry has been exploring and using big data from a variety of sources such as electronic health records and health information exchanges [2]. Healthcare systems can leverage this data to gain new insights into reducing costs, generating new revenue streams, and, most importantly, improving quality of care and patient outcomes. Hence, utilizing data to assist healthcare and administrative professionals, is a significant focus [1]. Leveraging this data and applying big data analytics to derive value is a current topic of concern in healthcare.

With the escalating use of big data analytics in many industries, many opportunities as well as challenges emerge. Being able to utilize data analytics software applications, develop new insights into the data, having the skills and analytical prowess to extract the relevant data and model it in the most insightful manner, as well as develop and decipher the many different representations of data that are possible can be beneficial and problematic. Being able to garner the full potential of big data analytics will require

F. F.-H. Nah and K. Siau (Eds.): HCII 2019, LNCS 11589, pp. 189–198, 2019.
https://doi.org/10.1007/978-3-030-22338-0_15

more than just data, such as knowledge of and ability to utilize appropriate analytic techniques and tools [3]. Some organizations have not realized the benefits that were intended and are re-addressing their strategies, which makes identifying factors that will lend to deriving value essential.

Although the amount of data continues to grow, some healthcare organizations are not sure what the true potential is of big data analytics [4]. Also, organizations are trying to identify cost-effective and optimal ways to deploy their resources to gain the most benefit from data analytics. For example, leveraging predictive analytics to create definite business value has not been fully realized [5]. Factors that have been challenges in this endeavor include the volume and complexity of data, analysts who are not fluent with data analytic techniques, and being able to demonstrate a salient impact. Deriving value from big data analytics is a major undertaking in healthcare [1]. In order to derive the full value that big data analytics can offer to healthcare, it's essential that we identify the factors that will contribute to this endeavor. Therefore, the research question being pursued for this study is: What are the factors essential to deriving value from big data analytics in healthcare?

Shedding light on the perspectives of various individuals, at varying levels or areas of the organization, impacted by big data analytics will be important for realizing the true value from big data analytics in healthcare [1]. This research proposes to address this research question by interviewing data analysts and management using value-focused thinking to identify the factors essential to deriving value from big data analytics. The outcome from these interviews and subsequent analysis will be a means-ends objective network. This approach will allow the essential factors (i.e., means) needed to derive the value (i.e., ends) from big data analytics in healthcare to be identified.

2 Literature Review

Big data has been defined with various characteristics. Big data has been described with the attributes of volume, velocity, variety, veracity, value, variability, and valence [3, 6]. Volume refers to the amount of data. Velocity is the speed that data is being created and delivered. Variety is a multifarious gamut of structured and unstructured data. Veracity is the quality, uncertainty, and imprecision in the data. Value is the data's worth such as developing competitive advantages. Variability refers to continuous modification. Valence is connectedness or relational aspects of the data.

Healthcare information systems and big data analytics are prominent in the healthcare industry. Healthcare information systems can be classified as clinical information systems (CIS) and administrative information systems (AIS), although some systems combine both [1]. CIS support the service provider and delivery of the service (e.g., electronic health records). AIS support administrative functions (e.g., payroll systems). Types of big data analytics include descriptive, predictive, and prescriptive [6]. Descriptive analytics provide insights into what has occurred in the past and is currently taking place. Examples include dashboards and standardized reporting. Predictive analytics provide projections regarding what could occur. Examples include Monte Carlo simulations and data mining. Prescriptive analytics provide guidance and suggestions regarding recommended courses of actions. Examples include adaptive algorithms.

Big data analytics can be leveraged to provide valuable insights and support decision-making processes in organizations [6]. Specific applications, such as predictive analytics, have been utilized to refine business processes, discover unanticipated opportunities, and proactively identify problems [5]. For instance, predictive analytics are being used to assess the probabilities of patient readmissions and identify patients with higher risks of certain diseases. Insurance companies can create such predictive models based on medical claim data. Some argue that insurers' experience and history with predictive analytics, along with magnitude of the institution, provides an advantage over healthcare systems capabilities.

Also, sources of data used in analysis can be more expansive and diverse [5]. Analysis can include not only data seen by payers but include clinical as well as psycho-socioeconomic data. This can then be leveraged to identify at-risk patents. For example, if a patient is not filling a necessary prescription, their healthcare provider can be given this information who can then intervene by communicating with the patient about his or her treatment plan. Another possibility is to develop a treatment plan based on the patient's current status including levels of activity and health condition (e.g., weight).

The healthcare industry has utilized big data in areas of bioinformatics and exchanging healthcare information to improve patient care, for example disease diagnosis [2]. Opportunities for big data analytics in healthcare include improvements in medical diagnoses, predicting disease epidemics, improving quality of care, identifying discontinuity in the delivery of healthcare services, reducing costs, discovering trends associated with reactions to medicines and hospital readmissions, and deriving treatments or cures and doing so based on an individual's personal circumstances (e.g., previous treatments and reactions, medical history) [1, 6, 7]. Healthcare models are evolving to be more personalized as well as proactive versus reactive, such as identifying categories of individuals with similar biological basis of a disease for better treatment plans [8, 9]. Healthcare systems have leveraged clinical data to gain better insights into populations' concerns and improving accuracy in predictions [5].

In an accounting context, big data analytics has been used for endeavors such as health care charge capture, fraud detection, continuous auditing/monitoring, anomaly identification or gaining other insights during an audit, and vendor payment transactions [5, 7, 10–12]. The value of big data analytics is still emerging and will require users who can apply advanced technical and analytical skills making the factors needed for competent usage potentially unique in comparison to other system usage such as enterprise resource planning systems. Also, identifying new uses of the data that will benefit one's organization will require attributes, skills, and knowledge capabilities that expand many users' existing repertoires. Hence, existing frameworks may need adaptation to properly align with the needs of big data analytics.

Big data analytics in healthcare is replete with challenges. This includes issues such as security as well as privacy [2, 7]. Other potential issues some have posed include the use of big data analytics over healthcare professional services. Greater reliance may be placed on big data analytics with less reliance on healthcare professionals' opinions and diagnoses, creating lesser need for their services [7].

Data governance issues have emerged such as fragmented data sets that could be integrated to decrease costs and improve integrity, accessibility, reliability, standardization, sharing, integration, transmission, processing, storage, and security [5, 6]. Data, such as electronic health record data, can be in both unstructured and structured formats [9]. Unstructured data, such as clinical notes, can be problematic to analyze because formats may vary, unnatural grammar or misspellings may be included, and domain-specific acronyms may be used. The frequency of data collection also varies. Some tests only occur once (e.g., genome), while others can be several times in one day (e.g., labs) or continuous (e.g., respiration). Data can also be generated from sensor devices, self-reporting data, videos, and medical images [6].

Data may be incomplete (e.g., pharmacy data for an entire population), substantial costs may be incurred, and extensive amounts of time to generate a return on investment can all be considered barriers to overcome [5, 6]. Data can have a significant amount of dimensions and highly heterogenous, and the quality of the data can be problematic [6, 9]. For instance, inadequate signal-to-noise ratios, errors in data entry, and gaps in data [9]. Effective utilization of predictive analytical models can be problematic due to data analysts' lack of medical knowledge creating inabilities to provide recommendations [5]. Also, healthcare professionals are deficient on time to participate in predictive model development and need to integrate these predictive models into existing work routines.

Some healthcare professionals may be averse to change [1]. Recommendations to effectively leverage data analytics include understanding the impact desired, consolidating data into a single location, integrating data analytics when deciding upon patient care and treatment, and utilizing qualitative as well as quantitative data [5]. Also, achieving the benefits that big data has to offer is dependent upon the correct assembly of technology, methods, and skill sets [1]. Organizations will need to consider effective approaches to integrate big data analytics in their current IT infrastructures. Also, for those who have historically used analytics, they will need to integrate big data analytics in a complementary fashion.

Previous research has addressed some aspects of leveraging information technology to achieve its full potential and the greatest return on investment. For instance, the abilities of the users. Previous literature has conceptualized information systems user competency or the ability of some users to obtain the greatest benefits possible from IS usage [13]. Also, previous research has explored critical success factors in domains such as ERP implementations (e.g., [14]). However, users of big data analytics can have unique challenges because of the inherent and potentially unstructured nature of the tasks to accomplish. For example, the objective may not be clearly defined and entail more exploratory, innovative, and unplanned discovery. Big data analytics is unique in that it can also entail statistics and machine learning [6]. The importance of proficient big data analytics usage needs to permeate the organization and not be centralized to just a few. In addition, this research focuses on ongoing usage of big data analytics versus its implementation.

Previous research has explored aspects of big data analytics in healthcare such as improving models, enhancing the richness of the data used, and identifying more optimal methods of processing data to improve hospital readmission predictive analytics [15]. Research has explored the use of variegated -omic and electronic health record data to facilitate precision medicine [9]. Research has put forth frameworks for studying big data and business-IT alignment from a social dynamics perspective [1]. However, previous research has not identified the elements needed to realize the desired value of big data analytics in healthcare. In other words, a gap in the literature can be filled by focusing on identifying the essential factors needed to derive the desired value from big data analytics in healthcare.

3 Research Methodology

This study proposes to utilize value-focused thinking (VFT) to identify essential factors associated with deriving value from big data analytics in healthcare. VFT can be more readily understood using Means-Ends Chain (MEC) Theory and a laddering interviewing technique [16]. According to MEC, product or service attributes are linked to the values that individuals are trying to achieve through the consequences associated with consumption [16, 17]. These values are pivotal drivers in decision making [17]. When individuals identify items (such as product or service attributes) that have the potential to fulfill what one values, these items can be categorized as such. Items or attributes used to create these categories are derived from one's values and originate from distinctions which "are dichotomies that represent *the end points of dimensions along which objectives may be compared*" [17, p. 63]. In other words, item or attribute categories are created based on their association with or ability to achieve a desired value or end state.

Consequences are the outcomes derived from an activity, whether they are experienced immediately or in the future, and are evaluated based on situational circumstances [17]. These consequences of engaging in an activity are associated with that activity in one's memory. Consequences can be directly derived from the activity, or indirectly from external sources (e.g., other individuals' reactions) after the activity occurs. The basic tenet is that decision making among various activities is based on achieving consequences that are deemed favorable or desirable and reducing unfavorable or undesirable outcomes. In some circumstances, trade-offs can occur in that achieving certain consequences means having to tolerate unfavorable consequences or sacrificing some other favorable consequences.

Values are arranged in one's value system by their importance [17]. Accordingly, consequences associated with more cardinal values or desired ends are more salient. Whereas, consequences associated with less vital values are less salient. Therefore, its important to understand the correlations between attributes and the consequences that can be derived in order for a decision to be made.

Attributes are considered *means* and the lowest level of the hierarchy of one's memory [16, 17]. The next level is considered *means* as well and consists of the consequences generated from consuming the product or service. The consequences can be physical, psychological, or social. The final level is the *ends* or the value that is derived from the consequences. Hence, in the decision-making process, individuals identify the means (or attributes and consequences) that ultimately provide the desired ends (or values). This chain can then shed light on the means that are important to achieving the ends.

Laddering allows deeper probing into cognitive structures and has been used with techniques such as Repertory Grid [18, 19]. This is accomplished by first prompting the interviewee to identify attributes that differentiate products or services [16]. Then, to reveal the ladder or means-ends chain (i.e., attribute, consequence, and end value), the interviewee is asked questions about the attribute elicited such as "Why is that important to you?" [16, p. 29]. The answers provided represent levels of abstraction and are subsequently used for additional questioning to further understand the chain or ladder until the highest level of abstraction is revealed. In the context of means-ends chains, this technique provides a mechanism to expose the relationship between the means/attributes, consequences, and values. Content analysis can then be utilized to identify pivotal elements of means/attributes, consequences, and values elicited from interviewees. Then, a cognitive map is created by the salient relationships among the elements.

VFT provides a mechanism to illuminate what individuals consider important to them and the manner to successfully attain it [16]. In other words, VFT focuses on the means-ends value chain. Values are identified by stated objectives, or something one aspires to accomplish. These objectives include "a decision context, an object, and a direction of preference" [16, p. 30]. Two types of objectives can be elicited – fundamental and means objectives. Fundamental objectives represent the values or end state one desires [16, 20, 21]. The means objectives are the methods to accomplish or bring to fruition these end states. The goal of VFT is to identify both of these objectives as well as their relationships, with the result being a means-ends objective network that is representative of decision-making in a given context.

Considering that VFT explicitly elicits the methods of obtaining a desired end state, this research method is considered appropriate considering the focus of this research is identifying those factors (i.e., methods of achieving the end state) associated with value from big data analytics in healthcare. VFT also places no restrictions on the objectives (i.e., means and ends) that participants can generate and facilitates a comprehensive approach to obtaining these objectives [22]. VFT is considered most ideal for this study because other methods, such as surveys, may not shed light on the relationships between means and ends.

VFT has been utilized previously in contexts such as strategic management and decision making, mobile technology in education, mobile technology's strategic impact, information system security, terrorism, system architecture assessment, and emerging technologies [16, 20–26]. The VFT approach allows the uncovering of what an individual values, or a desired objective, and the means with which it can be achieved. In the context of big data analytics, this study can utilize VFT to identify the values that data analysts have and the means to achieve these values. This study can also use VFT to identify management's values and the means with which they believe that they can be achieved. These means represent factors necessary to derive value from big data analytics in healthcare, and this study intends to identify these factors both from analysts' and management's perspectives to provide a more complete understanding of the relevant factors.

The VFT procedures are noted in Table 1 [16, 25, 27]:

Interviews with both data analysts and management will be recorded and transcribed. Research participant interviews will continue until the point of saturation is

Table 1. Value-focused thinking procedures

Step	Procedure
1	Research participants will be interviewed and asked to identify what they value in big data analytics or the value they see in big data analytics. This will include identifying issues, goals, and potential benefits. Participants will include both data analysts and management who are most familiar with these aspects of big data analytics (e.g., issues, desired value) so a more inclusive view of essential factors can be identified. Also, these factors will be identified by those who are interacting with, or responsible for the analysis, and are more likely directly impacted by or are familiar with objectives associated with big data analytics. Others, such as medical professionals, may be considered end users which can be addressed in future research studies. Specific questions to identify values will include: 1. What are the goals of using big data analytics in healthcare? 2. What are the potential benefits that can be derived from using big data analytics in healthcare? 3. If there were no limitations with big data analytics, what value could be derived? 4. What are the issues with effectively using big data analytics? After participants have identified all values they can generate, additional questions will be asked to identify additional values. For example, the pros and cons of big data analytics as well as the use of big data analytics in specific contexts. Also, questions will be posed to identify potential issues such as cost versus benefits and expertise. All participants' value lists will then be aggregated. Objectives will be reviewed by two individuals and consistent themes or meanings will be grouped together
2	The identified values are then transformed into objectives which will then be classified as either fundamental or means objectives. As mentioned previously, an objective is considered "something one wants to strive towards" and "has three features: a decision context, an object, and a direction" [27, p. 535]. Discerning these objectives can be accomplished by asking probing questions about the importance of each objective. More specifically, during the interviews, participants will be asked "Why is that important" in order to clearly distinguish means from fundamental objectives. Fundamental objectives are goals that the individual would like to achieve. Means objectives are methods of achieving other objectives. Hence, fundamental objectives will be identified as those that provide end values states and additional objectives cannot be derived
3	A means-ends objective network will then be constructed from the findings in the previous step in which the relationships between the means and fundamental objectives are depicted. The network will include any sub-objectives as components of the overarching fundamental objective

reached, or no new objectives emerge. Two individuals will code the results and identify means and fundamental objectives, and inter-rater reliability will be assessed using Cohen's Kappa coefficient [28]. Yin's [29] Principles of Data Collection will be utilized to enhance the reliability and validity of this study. This will include utilizing various sources of evidence, a database, and chains of evidence.

Utilizing various sources of evidence will be fulfilled by triangulation of the data in which two individuals will independently code the data and inter-rater reliability assessments made using Cohen's Kappa coefficient. Coefficients that are above .65 will be considered acceptable based on previous recommendations [30]. Also, both data analysts and management will be interviewed. To address the principles regarding creating and maintaining a database and chains of evidence, a database of all transcripts and notes taken during the interviews will be created. Also, all coding and the results obtained through each iteration will also be maintained in a database, and done so separately so a chain of evidence is created.

Based on the research participants responses to the probing question "Why is that important" and the objectives identified, the means-ends objective network will be derived. The ladders or relationships between the objectives will be included in the network if at least four participants have generated it, based on previous recommendations [16].

4 Expected Contributions, Implications, and Conclusion

Big data analytics is being heavily adopted in healthcare. However, deriving the full value it has to offer is a challenge for many organizations. The findings from this research study can potentially improve outcomes in the healthcare industry by identifying the pivotal factors needed to successfully leverage big data analytics. This model of means and fundamental objectives can provide insights into the most important factors (i.e., means objectives) needed to realize the greatest benefits from big data analytics in healthcare.

Practitioners can utilize the findings from this study to address the factors needed to realize the greatest return on their investment in big data analytics, as well as achieve successful performance outcomes. This can provide guidance regarding mechanisms that need greater investment or development to derive the desired values. Organizations may have already identified some of these mechanisms and not others. Hence, organizations can receive confirmation that current investments may be justified, and new investments may be needed. Also, current investments in mechanisms that are not identified in the network may need to be re-addressed in regard to the contribution they are making to achieve the end goal, i.e., the desired value from big data analytics.

Leveraging Means-Ends Chain (MEC) Theory and a laddering interviewing technique, value-focused thinking uncovers the methods or means to achieve the desired end or value. In this context, the means to achieve the desired value from big data analytics in healthcare (i.e., ends) will be identified. The research results from this study will provide a means-ends objective network that can provide future research directions. For example, the importance of means or essential factors for deriving value from big data analytics in healthcare can be addressed. Also, the findings can be explored or used in future research studies such as experiments entailing the assessment of identified means or mechanisms for deriving value from big data analytics. The network could be applied in action research studies of big data analytics in a healthcare institution as well, for example. Also, the findings can be assessed for generalizability to other domains, such as the financial services industry. Overall, big data analytics has the potential to make a significant contribution to improvements in healthcare, and identifying essential factors to derive its value is pivotal to this endeavor.

References

1. Weerasinghe, K., Pauleen, D., Scahill, S., Taskin, N.: Development of a theoretical framework to investigate alignment of big data in healthcare through a social representation lens. Australas. J. Inf. Syst. **22**, 1–23 (2018)
2. Guha, S., Kumar, S.: Emergence of big data research in operations management, information systems, and healthcare: past contributions and future roadmap. Prod. Oper. Manag. **27**(9), 1724–1735 (2018)
3. Grover, V., Chiang, R.H.L., Liang, T.-P., Zhang, D.: Creating strategic business value from big data analytics: a research framework. J. Manag. Inf. Syst. **35**(2), 388–423 (2018)
4. Angood, P.B.: Data and information critical for health care's future. Phys. Executive **40**(4), 4–5 (2014)
5. Appold, K.: Turn data into insight: how predictive analytics can capture revenue. Managed Healthc. Executive **27**(7), 16–21 (2017)
6. Saggi, M.K., Jain, S.: A survey towards an integration of big data analytics to big insights for value-creation. Inf. Process. Manag. **54**(5), 758–790 (2018)
7. Alexandru, A.G., Radu, I.M., Bizon, M.-L.: Big data in healthcare – opportunities and challenges. Informatica Economica **22**(2), 43–54 (2018)
8. Whelan, D.: Big data analytics, digital health platforms, and precision medicine tools: moving towards individualized and personalized care practices. Am. J. Med. Res. **5**(2), 28–33 (2018)
9. Wu, P.-Y., Cheng, C.-W., Kaddi, C.D., Venugopalan, J., Hoffman, R., Wang, M.D.: Omic and electronic health record big data analytics for precision medicine. IEEE Trans. Biomed. Eng. **64**(2), 263–273 (2017)
10. Raphael, J.: Rethinking the audit. J. Account. **223**(4), 29–32 (2017)
11. Tschakert, N., Kokina, J., Kozlowski, S., Vasarhelyi, M.: The next frontier in data analytics. J. Account. **222**(2), 58–63 (2016)
12. Tysiac, K.: How financial statement audits deliver key business insights: increased use of data analytics is helping auditors find more information that could be helpful to clients. J. Account. **223**(6), 44–45 (2017)
13. Eschenbrenner, B., Nah, F.F.-H.: Information systems user competency: a conceptual foundation. Commun. Assoc. Inf. Syst. **34**(1), 1363–1378 (2014)
14. Reitsma, E., Hilletofth, P.: Critical success factors for ERP system implementation: a user perspective. Eur. Bus. Rev. **30**(3), 285–310 (2018)
15. Zolbanin, H.M., Delen, D.: Processing electronic medical records to improve predictive analytics outcomes for hospital readmissions. Decis. Support Syst. **112**, 98–110 (2018)
16. Sheng, H., Siau, K., Nah, F.F.-H.: Understanding the values of mobile technology in education: a value-focused thinking approach. DATA BASE Adv. Inf. Syst. **41**(2), 25–44 (2010)
17. Gutman, J.A.: A means-end chain model based on consumer categorization processes. J. Mark. **46**(2), 60–72 (1982)
18. Kelly, G.A.: A Theory of Personality. W.W. Norton, New York (1963)
19. Kelly, G.A.: The Psychology of Personal Constructs. W.W. Norton, New York (1955)
20. Kunz, R.E., Siebert, J., Mütterlein, J.: Combining value-focused thinking and balanced scorecard to improve decision-making in strategic management. J. Multi-Criteria Decis. Anal. **23**(5–6), 225–241 (2016)
21. Kunz, R.E., Siebert, J., Mütterlein, J.: Structuring objectives of media companies: a case study based on value-focused thinking and the balanced scorecard. J. Media Bus. Stud. **13**(4), 257–275 (2016)

22. Dhillon, G., Torkzadeh, G.: Value-focused assessment of information system security in organizations. Inf. Syst. J. **16**(3), 293–314 (2006)
23. Keeney, G.L., von Winterfeldt, D.: Identifying and structuring the objectives of terrorists. Risk Anal. **30**(12), 1803–1816 (2010)
24. Lesinski, G.: Application of value focused thinking and fuzzy systems to assess system architecture. Procedia Comput. Sci. **61**, 168–175 (2015)
25. Sheng, H., Nah, F.F.-H., Siau, K.: Value-focused thinking and its application in MIS research. J. Database Manag. **18**(3), I–V (2007)
26. Sheng, H., Nah, F.F.-H., Siau, K.: Strategic implications of mobile technology: a case study using value-focused thinking. J. Strateg. Inf. Syst. **14**(3), 269–290 (2005)
27. Keeney, R.L.: The value of internet commerce to the customer. Manag. Sci. **45**(4), 533–542 (1999)
28. Cohen, J.: A coefficient of agreement for nominal scales. Educ. Psychol. Measur. **20**(1), 37–46 (1960)
29. Yin, R.K.: Case Study Research: Designs and Methods, 2nd edn. Sage Publications Inc, Thousand Oaks (1994)
30. Sun, H., Zhang, P.: Adaptive IT use: conceptualization and measurement. In: Proceedings of the Fifth Annual Workshop on HCI Research in MIS, Milwaukee, WI, pp. 65–69 (2006)