# Exploring Business Intelligence

David A. Wagle

School of Business, Northcentral University TIM7020: Databases & Business Intelligence Dr. Marty Crossland

March 21, 2021

# Exploring Business Intelligence Introduction

The field of Business Intelligence (BI) is a broad and diverse field covering many arenas. BI has no single definition and numerous researchers propose various capabilities that should be considered either within or excluded from the BI label. López-Robles et al. [(2019),](#_bookmark10) in a bibliometric study, offers an organizational focus, suggesting that BI is “the collection, analysis, interpretation, and dissemination of strategic information at the right time for use in the decision-making process.” This definition, while largely well-received, can be problematic for at least two reasons. First, the inclusion of specifically “strategic information” leaves out BI systems that are built to enhance tactical responsiveness, such as help-desk ticketing systems. While the overall effectiveness and efficiency of a company’s help-desk ticketing system is a strategic concern, the utilizing BI tool output to address a single ticket is not. Thus, this definition appears to be excessively narrow.

Turban [(2011)](#_bookmark16) addresses the difficulty by providing an even broader, but more IT-centric response. For Turban, BI is a broad concept that covers the entire set of archectures, tools, data systems, and methodologies that are utilized to provide

information in support of decision making in a business context. This definition likewise is problematic in some respects as it allows relatively simplistic relational database query

ad-hoc reporting to be considered a BI tool, a characterization that other BI researchers might not accept. However, this definition has the advantage of recognizing that the rapidly shifting world of BI is a capability built on a collection of individual tools an systems that allow data to be transformed into human decisions, while avoiding any attempt to limit the scope or level of those decisions within the business organization.

Chen et al. [(2012)](#_bookmark4) prefers to speak about BI in terms of an evolution of capabilities used to support business decision making. For Chen et al., BI is properly understood in stages, starting with Business Intelligence and Analytics (BI&A) 1.0, which is utilizing relational data models to produce dashboards, scorecards, and basic statistical analysis and

moving through BI&A 3.0, the current environment. This arena is utilizing highly

person-centered, context and location aware analysis to providing information and insights about individual interactions of systems and people. Each successive generation, of course, builds on the work of the previous.

It is with this context of BI&A that Eggert and Alberts [(2020)](#_bookmark6) performed a recent taxonomic review of the literature and defined 39 characteristics and 7 dimensions along which to measure BI&A literature. Within this taxonomy, six specific “emerging research areas” were defined. These are: human-computer interaction, data science foundations, text analysis, social analytics, network analytics, and mobile analytics.

For the purposes of this paper, the examination of BI&A will be limited to the emerging space of social analytics. This theory is chosen specifically because it presents both numerous technical difficulties to overcome as well as providing enormous potential for businesses that wish to embrace this challenge.

# Social Analytics in Theory

The current state of Social Analytics in BI&A is relatively advanced. For example,

J. Chen et al. [(2017)](#_bookmark5) built and demonstrated a system to use social media analytics data combined with physical sensors to both predict and to provide public health decision support for smog disasters in 8 cities in China.

Meanwhile, Gkatziaki et al. [(2017)](#_bookmark8) built a system that tracks mobile social media data to establish community boundaries and dynamics in an urban environment. This allowed the researchers to dynamically segment the city based on the behaviors of different demographic groups. This can help entities from individual citizens to public services to businesses better understand the actual workings of a city on a moment-by-moment basis. This would allow public health recommendations to be tailored to specific dynamic behaviors, for example, or for advertising firms to purchase time on dynamic display signage based on the demographic groups most likely to view the sign at a particular time.

A controversial, but highly interesting area of research is focused on leveraging social analytics against financial information to improve the accuracy of financial decisions, for example by enhancing credit scores (Chao & Chang, [2020](#_bookmark3)). This research noted not only the ability to leverage social analytics to enhance credit scores, but also was also able to identify existing deficiencies in the existing credit score models based on that same data.

Of course, the ethical and privacy implications of turning this research into an actual in-use method need to be explored further.

This concern about the ethics and privacy of such research is not limited to social analytics data, of course. It is a major area of discussion throughout big data analytics (Polonetsky, [2012;](#_bookmark11) Richards & King, [2014).](#_bookmark12)

These techniques have shown greater growth in non-business context in recent years.

For example, some researchers have been able to use social analytics systems to help college students mediate their drinking behavior and their perception of peer drinking behaviors (Russell et al., [2021).](#_bookmark13) Researchers have also used these techniques to help identify issues of depression and academic performance in adolescents (Badri et al., [2021](#_bookmark1); Wentzel et al., [2021).](#_bookmark17)

# Social Analytics in Practice

In understanding how little social analytics is impacting business today, it is important to understand the distinctions between social analytics and other similar areas. Social analytics is specifically about the analyzing individual behavioral characteristics that can be garnered from large real-time data sets (Eggert & Alberts, [2020)](#_bookmark6) and shouldn’t be concerned with related arenas such as geo-location analysis, mobile device interaction analysis, web analytics, opinion mining, social media network analysis, or social media analytics. It does, in some ways, ecompase elements of all of those fields, but it is in trying to build a holistic and dynamic model of the social behaviors of the person that social analytics differs.

There are numerous products out there for social media and social network analytics (Sivarajah et al., [2020).](#_bookmark14) But true social analytics is, as Eggert and Alberts called out, an emerging area of research. While there are numerous ideas on how social analytics can be used by businesses, such as making better human resource decisions (Galli, [2019)](#_bookmark7) to the aforementioned use of dynamic city social maps helping advertisers, at present there are no truly notable examples of social analytics in “the wild.”

# Mind the Gap

The reason for this research/practice gap is multifaced. First, there are serious ethical and privacy concerns that are being grappled with as user-data analytics become more and more capable and omnipresent.

Even in seemingly benign applications, such as library services at universities using BI&A to help predict students who can benefit from interventions and assistance to ensure better academic outcomes, privacy issues are limiting adoption (Jones, [2019).](#_bookmark9)

When the idea is carried out to larger scales, such as creating smart cities, or dynamically mapping individuals based on social analytics, the concerns only increase (Chang, [2021).](#_bookmark2) At issue, is the problem of consent (Solove, [2012).](#_bookmark15) Users are rightly concerned about how their personal information is being used, and the ecosystem does not yet have a regulatory framework in place to adjudicate what is reasonable extensions of existing norms and what is ethical abuse.

A second consideration is simply the amount of data and computing power such analysis requires. Big Data is defined by the three-V model: volume, velocity, an variety (Wu et al., [2014).](#_bookmark18) Volume is the amount of data that is being considered, velocity is how fast that data is created, and variety is the number of sources and disparate structure(s) of the data. Other researchers add in “value” and ”veracity” to create the five-v model.

However, it should be noted that data by itself is intrinsically without value, it is only when the data can be used to create decisionable information that business value can be

realized. To that end, it is impossible to know if data is valuable or not prior to value being realized. Similarly, the veracity of the data is both relative (some margin of error in numerical empirical measurements is always allowed for example) and unknowable up front. For this reason, the three-v model will be used.

The general consensus is that an entity is dealing with big-data when either the commodity scale-up or commodity scale-out of the workload is no longer possible and speciality systems must be obtained in order to handle the workload (Appuswamy et al., [2013).](#_bookmark0) As the Microsoft research team showed, big-data is much, much rarer than most people presume. Indeed, since that 2013 research, the situation has only improved from a commodity perspective.

To do true large-scale social analytics requires an enormous amount of data acquired at a remarkable velocity. Consider doing dynamic social analytical mapping of an urban area with only 500,000 people. If the system is taking samples of a dozen data points from 10% of the population every 15 minutes, and each data point and associated metadata is 1 kilobytes, the system will collect .5 terabytes of data per day, and just shy of 170 terabytes over a year.

Utilizing to Amazon’s AWS calculator and choosing the best pricing model for a reserved Aurora DB, caching, and load balancing to handle 2Pb of writes per year, the annual cost is almost $900 thousand as shown in figure [1.](#_bookmark19) This is prior to providing for any CPU or GPU time for analytical services, network utilization reserve or other components. This is the cost purely for reading the collected data into a database and storing it.

This is a real barrier to adoption for emerging BI&A efforts. Big data is not inexpensive and without demonstrated returns on the investment, most organizations are not going to spend the resources to pursue what is still theoretical advantages.

Even very large organizations are going to focus primarily on more proven technologies rather than pursuing new areas of primary research.

# Overcoming the Gap

Like previous areas of BI&A, adoption will start with the largest companies when they are presented with supported products from reliable vendors. There will be a typical adoption curve, with some companies seeking an advantage by being an early adopter, and others taking a more “wait-and-see” approach. Once the vendors have mastered building and delivering the technology, it will become more of a commodity product and offerings will appear aimed at smaller-to-mid-sized companies. However, the SMC adopters will likely not enjoy the success of social analytics that larger companies will enjoy due to the cost of data storage for truely large data sets.

However, there likely will arise a market for big-data service providers. For example, a company that was performing dynamic social analytics mapping of a metro area could reasonably expect to sell the results of that analysis to any number of marketing firms, thus permitting an economically reasonable ecosystem to arise around the technology, with the resulting analysis provided as a purchased service.

# References

Appuswamy, R., Gkantsidis, C., Narayanan, D., Hodson, O., & Rowstron, A. (2013).

Nobody ever got fired for buying a cluster, 1–12.

Badri, M., Khaili, M. A., Bahar, M. A., Yang, G., Reynhout, G., & Rashdi, A. A. (2021). Social connection and self-perceived depression among adolescents: A path analytic model for abu dhabi. *Journal of Child and Family Studies*, *30* (1), 146.<https://doi.org/10.1007/s10826-020-01891-2>

Chang, V. (2021). An ethical framework for big data and smart cities. *Technological* *Forecasting & Social Change*, *165*. <https://doi.org/10.1016/j.techfore.2020.120559>

Chao, C.-M., & Chang, W.-T. (2020, December). Application of social analytics to credit rating in financial technology, In *2020 international computer symposium (ics)*. 2020 International Computer Symposium (ICS), IEEE.<https://doi.org/10.1109/ICS51289.2020.00111>

Chen, Chiang, & Storey. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36* (4)jstor 10.2307/41703503, 1165.<https://doi.org/10.2307/41703503>

Chen, J., Chen, H., Wu, Z., Hu, D., & Pan, J. Z. (2017). Forecasting smog-related health hazard based on social media and physical sensor. *Information Systems*, *64*, 281–291. <https://doi.org/10.1016/j.is.2016.03.011>

Eggert, M., & Alberts, J. (2020). Frontiers of business intelligence and analytics 3.0: A taxonomy-based literature review and research agenda. *Business Research*, *13* (2), 685–739. <https://doi.org/10.1007/s40685-020-00108-y>

Galli, L. (2019). How organizations are using behavioral science to make better people decisions. *People & Strategy*, *42* (1), 1–6. Retrieved March 20, 2021, from [http://proxy1.ncu.edu/login?url=https:](http://proxy1.ncu.edu/login?url=https)

//search.ebscohost.com/login.aspx?direct=true&db=ofs&AN=134950513&site=eds- live

Gkatziaki, V., Giatsoglou, M., Chatzakou, D., & Vakali, A. (2017). Dynamicity: Revealing city dynamics from citizens social media broadcasts. *Information Systems*, *71*, 90–102. <https://doi.org/10.1016/j.is.2017.07.007>

Jones, K. M. L. (2019). "just because you can doesn’t mean you should": Practitioner perceptions of learning analytics ethics. *portal: Libraries and the Academy*, *19* (3), 407–428. Retrieved March 20, 2021, from

http://proxy1.ncu.edu/login?url=https://search.ebscohost.com/login.aspx?direct= true&db=eric&AN=EJ1222099&site=eds-live

López-Robles, J., Otegi-Olaso, J., Porto Gómez, I., & Cobo, M. (2019). 30 years of intelligence models in management and business: A bibliometric review.

*International Journal of Information Management*, *48*, 22–38.

<https://doi.org/10.1016/j.ijinfomgt.2019.01.013> Polonetsky, O. T.

bibinitperiod J. (2012). Privacy in the age of big data: A time for big decisions. *Stanford Law Review Online*, *64*, 63. <http://www.stanfordlawreview.org/online/privacy-paradox/big-> data$%5C%5Cbackslash$nhttp:

[//www.stanfordla](http://www.stanfordlawreview.org/sites/default/files/online/topics/64-SLRO-63_1.pdf)wr[eview.org/sites/default/files/online/topics/64-SLRO-63\_1.pdf](http://www.stanfordlawreview.org/sites/default/files/online/topics/64-SLRO-63_1.pdf) Richards, N. M., & King, J. H. (2014). Big data ethics. *Wake Forest Law Review*.

<https://doi.org/10.1177/2053951714559253>

Russell, A. M., Patterson, M. S., & Barry, A. E. (2021). College students’ perceptions of peer alcohol use: A social network analytic approach. *Substance Use & Misuse*, *56* (1), 46–53. Retrieved March 20, 2021, from http://proxy1.ncu.edu/login?url=https://search.ebscohost.com/login.aspx?direct= true&db=s3h&AN=148039141&site=eds-live

Sivarajah, U., Irani, Z., Gupta, S., & Mahroof, K. (2020). Role of big data and social media analytics for business to business sustainability: A participatory web context.

*Industrial Marketing Management*, *86*, 163–179.<https://doi.org/10.1016/j.indmarman.2019.04.005>

Solove, D. J. (2012). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review*, *126*, 1880. https://heinonline.org/HOL/Page?handle=hein. journals/hlr126&id=1910&div=&collection=

Turban, E. (Ed.). (2011). *Business intelligence: A managerial approach* (2nd ed). Boston, Prentice Hall

OCLC: ocn611966305.

Wentzel, K. R., Jablansky, S., & Scalise, N. R. (2021). Peer social acceptance and academic achievement: A meta-analytic study. *Journal of Educational Psychology*, *113* (1), 157–180. <https://doi.org/10.1037/edu0000468>

Wu, X., Zhu, X., Wu, G.-Q., & Ding, W. (2014). Data mining with big data. *Knowledge and Data Engineering, IEEE Transactions on*, *26*, 97–107. <https://doi.org/10.1109/TKDE.2013.109>

# Figure 1

*AWS Estimate*