

TOWARD DIFFERENTIATING BUSINESS ANALYTICS, BIG DATA ANALYTICS AND DATA SCIENCE: AN EMPIRICAL APPROACH

Mihaela MUNTEAN

Bucharest University of Economic Studies, Romania

mihaela.muntean@ie.ase.ro

Ana Ramona BOLOGA

Bucharest University of Economic Studies, Romania

ramona.bologa@ie.ase.ro

Abstract. *Business analytics, big data, or data science have become the terms we meet daily, and we get to use them, even if the meaning, scope and skills they involve are not entirely clear to us. This paper identifies and analyzes the main definitions of Business analytics, big data analytics, and data science to determine the similarities, differences and relationships between these terms. A comprehensive definition of the term Business Analytics is proposed and its current relationships with Big Data Analytics and Data Science are defined, discussing how they might evolve as a result of current trends.*

Keywords: business analytics, data science, big data analytics

JEL classification: J24, I21, I23

DOI: 10.12948/ie2019.04.05

1. Introduction

The rapid growth in the use of the term “analytics” has been accompanied by a proliferation in the way analytics is used, and terms such as “Big data analytics-BDA” or “Business analytics-BA” have become common in business. In the same area as Business Analytics, another term that has been more prominent in recent years is Data Science (DS), which is also related to data: data acquisition, data modeling, data gathering, etc. BA, BDA and DS have attracted attention from managers, academics, software vendors, and IT consultants. As the importance of these terms continue to increase so does the number of definitions for these concepts. At present, there is no unanimously accepted definition for BA, BDA and DS. Also, sometimes, the terms are used interchangeably, and this leads to confusion regarding their meaning and how they are related. For those concerned with the training of data specialists to fill the gap in the labor market, it is important to know how the meanings of these terms differ and overlap, what knowledge and skills each of them involves. In this context, the primary objective of this article is to identify and review the existing definitions of BA, BDA and DS to clarify the similarities, differences and relationships between them. Previous efforts in this direction are few [1], [2], [3] most do not venture into definitions, but provide descriptions of the methods and tools used or the skills that professionals in the field should have.

2. Business analytics versus Big data analytics and Data Science: a review of definitions

In the following, we will present the results of the analyzes carried out on the terms “Business analytics-BA”, “Big data analytics-BDA” and “Data science-DS” in order to establish the similarities, differences and relationships between these terms. We used the following tools: Google Trends, Google Ngram, WordItOut and SAS JMP Text explorer.

2.1. Analysis with Google trends and Google Ngram Viewer

Google Trends was used to create a context for keyword searches related to these terms. Figure 1 depicts the relative number of searches for the terms “Business Analytics”, “Big data analytics” and “Data science” between January 2004 and February 2019.

We observe that the term “Data science” is more popular than “Big data analytics” and “Business analytics”. In February 2019, “Data science” had a relative search score of 100 (that means that this term has a very high popularity) and “Big data analytics” with a score of only 9. Also, it is obvious that searches for all terms are on a steady increase. We observe that BDA has a score of 0 until 2011, meaning there was not enough data for this term until 2011.

Google Ngram Viewer was used then to see the frequency of using a word in printed text between 1500 and 2008. A Google Ngram search for the terms BA, BDA and DS for the period 1966–2008 shows the number of mentions for these terms in the corpus of books (Figure 2). The use of the term “Business analytics” started to rise in 1995 and a more detailed search highlights the first reference to this concept in [4] where BA was considered a quantitative method. “Big data analytics” is not found in corpus of books between. For Data science (or “the science of datalogy”) the first definition was given by Peter Naur, in 1966 [5].

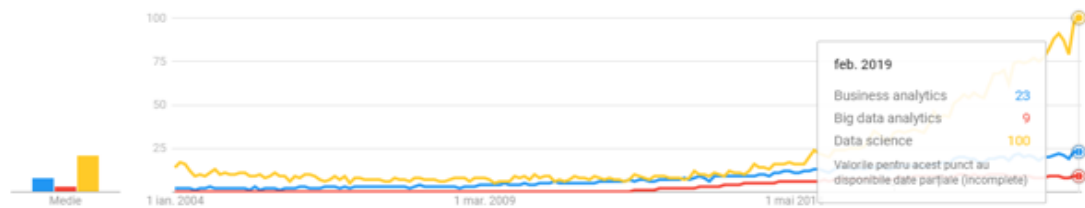


Figure 1. Interest overtime in key search terms -Google Trends.

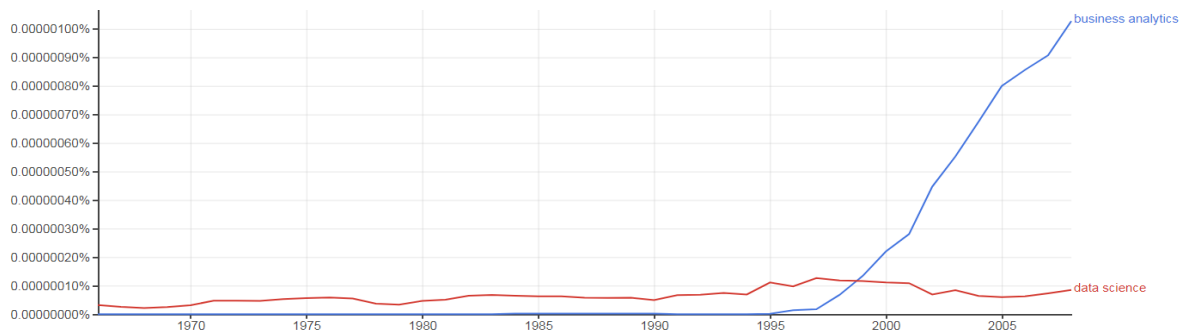


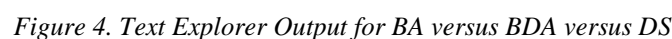
Figure 2. Relative interest in key terms Google Ngram Viewer

2.2. Text analyses of definitions

Until now, the terms BA, BDA and DS have been defined in many ways. Also, sometimes, DS and BDA or BA and BDA are used interchangeably, and this leads to confusion regarding their meaning and how they are related. We examined industry professionals’ blogs/websites and literature (only books) that deals with these terms to synthesize definitions and describe the relationships and differences between them. These definitions are grouped into two categories: the definitions provided by industry professionals and experts on their blogs/websites and the definitions provided by researcher’s books. The definitions found in the top 15 results of a Google Search have been processed and analyzed, identifying for each of them the focus and focus on data, technologies, analytical methods and skills. The results of this analysis were synthesized in Table 1. To understand the content of the 15 definitions, a Word Cloud was generated using the *Worditout* tool (<https://worditout.com/word-cloud/create>) as shown in Figure 3. The word cloud displays the words used in the definitions that have a minimum

[illegible]

To determine the most common terms and phrases in the definitions, we used the *JMP Text explorer*-SAS tool. We combined the related terms. Also, we eliminated the words that are not relevant for analysis. As shown in Figure 4 (a) there are 49 unique terms in 15 definitions of BA. They are listed below on the left side and sorted by frequency. The most common terms for BA are: “data”, “business” and “analytics”.



On the right side of Figure 4 (a), there is a list of phrases common to the definitions. The phrases “decision making” and “predictive analytics” have the most occurrences for BA. Also, the phrases “big data” and “data sets” have the most occurrences for BDA and the phrases “machine learning” and “big data” have the most occurrences for DS. The most common terms for BDA are: “data”, “analysis” and “big” and for DS are “data”, “information” and “business”. Also, we observe that the most used word in BA/BDA/DA definitions is “data”. Figure 5 display a Venn diagram that illustrates the similarities and the differences found between the definitions. What have in common these definitions? The common words of all definitions are data (big data), information, analysis, decision-making, analytics and business. Finally, Table 1 synthetically presents the main similarities and differences between the examined terms.

Table 1. Comparative analysis of definitions

	Business analytics-BA (decision centric)		Big data analytics-BDA (data centric)	Data science-DA (data centric)
is:	- “a process of transforming data (sometimes big data) into insight;” [6], [7], [9], [12], [13], [14]		- “process of collecting, organizing, transforming and analyzing big data” [6], [8], [10], [11], [15], [16], [17], [18]	- “deals with the creation of insights or data products from a given set of data files (usually in unstructured form), using analytics methodologies;” [19] - “a field of Big Data geared toward providing meaningful information based on large amounts of complex data;” [20]
	- “the collection of skills, technologies, applications and systems that used by organizations to improve business performance and business decision-making;” [10], [15], [21], [23]		- “is the collection of analytics technologies to collect, organize and analyze big data;” [24] - “systems that enable end-user access to and analysis of data;” [25]	- “lies at the intersection of hacking skills; math and statistics knowledge; substantive expertise;” [26] - “refers to the collective processes, theories, concepts, tools and technologies that enable the review, analysis and extraction of valuable knowledge;” [8] - “is a multidisciplinary blend of data inference, algorithm development, and technology in order to solve analytically complex problems;” [27] - “a multi-disciplinary field that: uses scientific methods, processes, algorithms and systems to extract knowledge and insights from data;” [22], [19]
	- “methods and techniques pursued during the process of exploration data;” [8], [29], [30]		- “include predictive and prescriptive analytics, advanced analytics techniques that operate on big data;” [22], [28], [31], [32]	- “combines machine learning, statistics, advanced analysis, and programming;” [22] - “combines social sciences, advanced statistics, and computer engineering skills to acquire, store, organize, and analyze information across a variety of sources;” [28] - “is creating insight-driven manufacturing;” [11] - “is the business application of machine learning, artificial intelligence and other quantitative fields;” [33] - “is the discipline of making data useful;” [34] - “the science of dealing with data;” [5] - “is the art and the science of acquiring knowledge through data;” [35]
focus on	data	Yes (sometimes refers to big data)	Yes (big data)	Yes (structured or unstructured-big data)
	analytical methods	- statistical methods - descriptive analytics (reporting) - predictive analytics (data mining) - prescriptive analytics (optimization techniques)	- statistics - descriptive (visualization) - predictive analytics (text analytics, machine learning, data mining, natural language processing, forecasting, operations research, econometrics)	- statistics - descriptive (visualization) - predictive (machine learning, artificial intelligence, mining) - mathematics

	Business analytics-BA (decision centric)		Big data analytics-BDA (data centric)	Data science-DA (data centric)
		- quantitative methods and mathematical models	- prescriptive (optimization, simulation)	
	technologies	No	Yes (Hadoop, MapReduce and NoSQL databases)	No
	skills	very little	No	-programming skills, computer engineering skills, hacking skills
Goal	- to improve business performance		- to discover hidden patterns, unknown correlations in data	- review, analysis and extraction of valuable knowledge and information from raw data
	- to support business decision-making;		- to help organizations make informed business decisions;	- generate insights that analysts and business users translate into tangible business value;
The most used words (Word cloud)	business, analytics, data		data/big data, analytics	data, information, business, science
The most common terms by frequency (JMP)	data, business, analytics		data, analytics, big	data, information, business, insight, knowledge
The most common phrase (JMP)	-decision making - predictive analytics		- big data - data sets	- machine learning - big data
The common words (Venn)	analysis, analytics, big data, information, process, statistics, decision, insights, business			

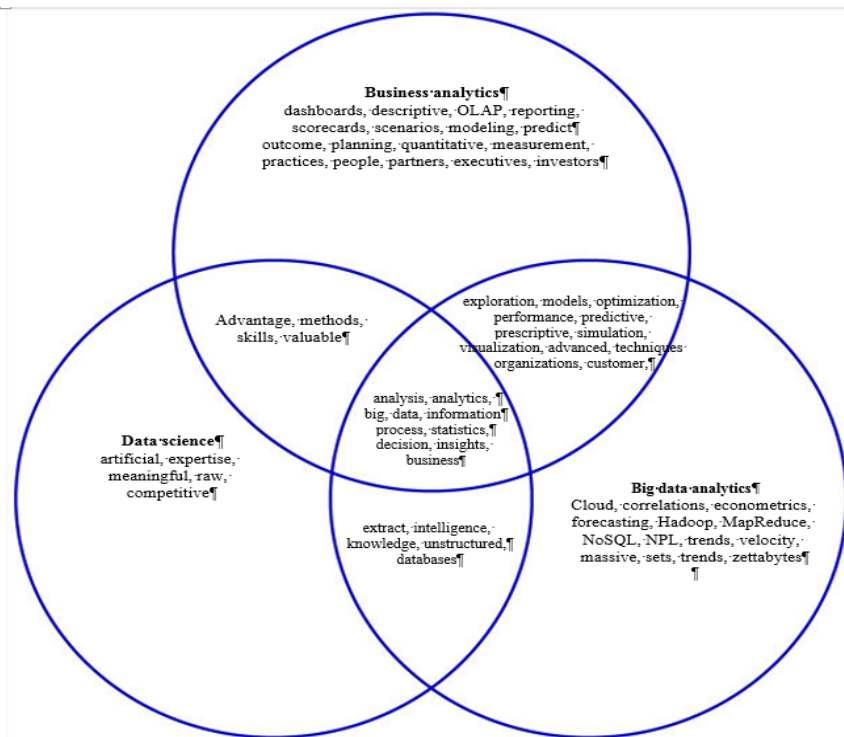


Figure 5. Venn diagram

3. Discussion and conclusions

An examination of BA/BDA/DS definitions reveals that terms can mean different things to different people, depending on their business focus and their perspective. For example, BA has been defined in terms of types of methods and techniques pursued during the process of exploration data [8], [29], [30]. The methods mentioned include statistics, predictive analytics, prescriptive analytics, descriptive analytics, optimization, etc. Also, BA has been defined by some [10], [23], [15] as including “the collection of skills, technologies, applications and systems that used by organizations to improve business performance and business decision-making”. Finally, BA has been defined by some ([6], [11], [9], [7], [12], [13]) as a process of

transforming data (sometimes big data) into insight. All definitions relate to the data, the analytical methods, the technologies used, and the required skills. Some industry professionals offer definitions for all terms (for example [6] and [8]). Some definitions are similar (for example, [9] and [18] for BA and BDA).

While all the BA and BDA definitions share the same goal of supporting business decision making and improving business performance, DS aim at extracting valuable knowledge from raw data and to solve analytically complex problems. Is BA a subfield of DS?

Both BA, BDA and DS focus on data analysis, all types of data for DS, and only business data for BA. DS can answer questions that are related to business decisions, but usually, they answer more general questions. Even though the BA mainly uses structured data, and DS and BDA use both structured and unstructured data, we believe that this difference will no longer exist in the future. Also, according to Table 1, DS, BA and BDA use analytical methods (descriptive, predictive, prescriptive) and mathematical models. We can say that the BA definition is very similar to that of BDA. Both BA and BDA focus on data and analytical methods and share the same goal. The only difference is that BDA refer also, to the technologies used. Also, according to the studied definitions, DS involves advanced programming skills, advanced mathematical modelling skills, but less business domain expertise, while BA does not involve programming skills, but domain expertise. However, we consider that a BA specialist must also have programming skills (such as experience with R or Python).

On the basis of the above considerations, a comprehensive definition of the term Business Analytics is the following: *BA refers to apply advanced analytical methods/tools (descriptive, predictive, prescriptive) and mathematical models on large data sets (structured, semi-structured and unstructured) to derive insights that will be used to improve business performance and business decision making. Also, BA involves both soft skills, analytical skills and technical skills (programming skills and experience with big data).*

Although the definitions of the analyzed terms have changed over time, it is expected that their scope will also evolve along with the development of technology. The clear understanding of BA, BDA and DS concepts and as well as the relevant knowledge and skills for these fields are important for the academic programs that teach BA/BDA and DS, but also, for organizations looking for professionals in these fields. We believe that our approach makes an important contribution in this direction, by analyzing how these related areas were defined up to now, and clarifying the inherent confusions and overlaps that arise.

References

- [1] D. J. Power, C. Heavin, J. McDermott, M. Daly, "Defining business analytics: an empirical approach," *Journal of Business Analytics* [Online], vol. 1, no. 1, 2018, pp. 40-53. Available at: <https://www.tandfonline.com/doi/full/10.1080/2573234X.2018.1507605>
- [2] M. Bichler, A. Heinzl, "Business Analytics and Data Science: Once Again?" *Business & Information Systems Engineering* [Online], vol. 59, no. 2, 2017, pp. 77-79. Available at: <https://link.springer.com/article/10.1007/s12599-016-0461-1>
- [3] C. Longbing, "Data science: A comprehensive overview," *ACM Computing Surveys*, [Online]. vol. 50, no. 3, 2017, pp. 1-42. Available at: <http://203.170.84.89/~idawis33/DataScienceLab/publication/CSUR-Datascience-overview-published.pdf>
- [4] Harvard Business School Bulletin, vol. 55-56, Harvard Business School, 1979, pg. 7.
- [5] P. Naur, "The science of datalogy," *Communications of the ACM*, vol. 9, no. 7, 1966, pg. 485.
- [6] Techtartget Definitions. Available at: <https://whatis.techtarget.com>, [Jan 20, 2019]
- [7] Techadvisory. Available at: <https://www.techadvisory.org/2014/01/business-analytics-defined/>, 2014, [Feb 10, 2019]

- [8] Techopedia Dictionary. Available at: <https://www.techopedia.com/dictionary>, [Jan 20, 2019]
- [9] M. Galetto, "What is BA?" Available at: <https://www.ngdata.com/what-is-business-analytics/>2016, 2016, [Feb 16, 2019]
- [10] Webopedia. Term Definitions. Available at: <https://www.webopedia.com/TERM>, 2018, [Feb 20, 2019]
- [11] Maruti Techlabs Follow. Available at: <https://www.marutitech.com>, 2017 [Feb 16, 2019]
- [12] J. Evans, "Business Analytics. Methods, models and decisions," Pearson Education, 2012, pp. 27-63.
- [13] D. Anderson, J. Camm, J. Cochran, "Essentials of Business Analytics," Cengage Learning, 2015, pg. 5.
- [14] M. J. Schniederjans, D. G. Schniederjans, C. M. Starkey, "Business Analytics Principles, Concepts, and Applications: What, Why, and How," Pearson Education, 2014, pp. 1-15.
- [15] CIO White papers review. Technical Terms and Definitions. Available at: <https://whatis.ciowhitepapersreview.com/definition/business-analytics/>, [Feb 11, 2019]
- [16] Informatica Glossary of terms. Available at: https://www.informatica.com/services-and-training/glossary-of-terms/big-data-analytics-definition.html#fbid=_45_1RvBZJF, [Jan 20, 2019]
- [17] SAS. "Big data analytics. What it is and why it matters?" Available at: https://www.sas.com/en_us/insights/analytics/big-data-analytics.html, 2018, [Feb 18, 2019]
- [18] A. Venkat, "Big Data Analytics," Packt Publishing, 2016, pp. 1-16.
- [19] Z. Voulgaris, "Data Science: Mindset, Methodologies, and Misconceptions," Technics Publications, 2017, pp. 13-24.
- [20] J. Frankenfield, "What is Data Science," Available at: <https://www.investopedia.com/terms/d/data-science.asp>, 2018 [Feb 15, 2019]
- [21] IBM Annual Report 2013. "What will we make of this moment?" Available at: https://www.ibm.com/annualreport/2013/bin/assets/2013_ibm_annual.pdf, pp. 30.
- [22] IBM Analytics. Available at: <https://www.ibm.com/analytics/>, [Feb 15, 2019]
- [23] Gartner IT Glossary. Available at: <https://www.gartner.com/it-glossary/business-analytics>, 2018, [Feb 20, 2019]
- [24] Z. Sun, H. Zou, K. Strang, "Big data analytics as a service for business intelligence" in Proc. Conference on e-Business, e-Services and e-Society, Open and Big Data Management and Innovation, LNCS, vol. 9373, 2015, pp. 200-211. Available at: <https://hal.inria.fr/hal-01448039/document>
- [25] Dresner Advisory Services. LLC, Wisdom of Crowds® Series. "Big Data Analytics Market Study," Available at: https://www.microstrategy.com/getmedia/cd052225-be60-49fd-ab1c-4984ebc3cde9/Dresner-Report-Big_Data_Analytic_Market_Study-WisdomofCrowdsSeries-2017.pdf, [Feb 10, 2019]
- [26] G. De Joel, "Data Science from Scratch: First Principles with Python," O'Reilly Media, 2015, pp. 1-13.
- [27] Datajobs. "What is Data Science?", Available at: <https://datajobs.com/what-is-data-science>, [Feb 15, 2019]
- [28] TDWI research. TDWI Glossary. Available at: <https://tdwi.org/Portals/TDWI-Glossary.aspx>, [Feb 10, 2019]
- [29] E. Stubbs. "The value of business analytics: Identifying the path to profitability," SAS Institute Inc., 2011, Available at: <https://pdfs.semanticscholar.org/bfce/dcd276d618bcad842df76c8b31ad2c392862.pdf>, [Feb 18, 2019]

- [30] A. Taber, S. Samaddar, R. Bordley, R. Musa, M. Smith, F. Stein, C. Truxillo and Z. Waltz, "Certified Analytics Professional (Cap®) Examination Study Guide," 2014, Available at: <https://www.informs.org>, [Dec 16, 2018].
- [31] P. B. Goes, "Big Data and IS Research", MIS Quarterly, vol. 38, no. 3, 2014, pp. 3-8.
- [32] M. C. Minelli, D. Ambiga, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for today's Businesses," John Wiley & Sons, Inc., 2013, pp. 1-18.
- [33] V. Kotu, B. Deshpande, "Data Science: Concepts and Practice," Morgan Kaufmann, 2018, pp. 1-18.
- [34] C. Kozyrkov, "What on earth is data science?" Available at: <https://www.kdnuggets.com/2018/09/what-is-data-science.html>, 2018, [Feb 15, 2019]
- [35] S. Ozdemir, "Principles of Data Science," Packt Publishing, 2016, chapter 1, pp. 1-20.