

## How to measure economic feasibility of big data?

Ricardo B. Verschuere  
Department of Computer Science  
The University of Gloucestershire  
The Park, Cheltenham.G152 6BX England UK

**Abstract**— This research investigates how economic feasibility of big data solutions can be measured. After reviewing the literature, no existing papers are found within this area of research which focus on the economic or financial aspect. The research looks at some evaluations of feasibility studies and studies that highlight the benefits of big data solutions. Based on these together with a framework for accessing big data solution and the cost benefit analysis approach a method is presented for determining the economic feasibility of big data solutions.

**Keywords:** Big-Data, Economic feasibility, Cost benefit analysis, BigDAF.

### INTRODUCTION

The concept of Big data has increasingly taken off and has not found widespread application throughout numerous industries and application. However, the question remains, does it make business sense to adopt big data technologies? For corporations like Facebook, it is often asked: “What makes them worth so many billions?”. Facebook [1] its value lies in its accumulation of data and large user base. Its demographic data has enabled a marketing heaven for advertisers online. However, Mark Zuckerberg himself states that the platform wasn’t built for this purpose, IE generating money, and they only provide this service such that they can build better services.

While most companies produce standard management intelligence (MI) reports, the question lies in what underlying data is used to produce those reports. And whether true big data is being applied. This raised questions on which data

should be stored, and if so does it make business sense to store all this data? In most cases the data in itself has little intrinsic value, however, joining different data sets through specific computational methods can create MI value for the organisation. Because the value of this MI can’t be quantified in monetary terms, if it can be quantified at all, it is considered a very complex matter to access the economic feasibility of big data.

This paper seeks to understand the underlying concepts of what big data means, what the major IT players have provided as a solution to implementing big data and attempt to understand whether it is possible to evaluate its cost effectiveness. The research than has a look at current methods of evaluating economic feasibility, compare different approaches and access whether there is any applicability whatsoever to big data solutions. Furthermore, unlike the fortunate development of Facebook which became financially viable as a consequence of big data as opposed to their actual platform original purpose. This research seeks to understand how big data platforms can be built economically sustainable from the ground up.

### LITERATURE REVIEW

The literature review of this paper looks at two aspects, firstly, the concept of Big Data, secondly economic feasibility. Systematic approach was applied by searching for Big data literature review in major academic outlets such as science direct, Emerald, and further search was applied in Google and Google Scholar. Around 25 relevant articles were identified which were categorised as follows: “Survey, Big data analytics(BDA), technology, challenges, sector specific case studies and other”. After reading through the articles further searches were made to look into “big data feasibility study”. This search was modified a few times to looks for economic

and financial feasibility studies. Unfortunately, little relevant or none was found in this area.

### *Big data*

Big data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process data within a tolerable elapsed time [2]. It is characterised by its three Vs:

- Volume: big data doesn't sample; it just observes and tracks what happens
- Velocity: big data is often available in real-time
- Variety: big data draws from text, images, audio, video; plus it completes missing pieces through data fusion

### *Determine the value of big data*

Raguseo [3] evaluated through empirical investigation what the benefits and risks are of big data adoption. Based on the literature, key benefit and risk factors were identified and used to form questions in a survey. Based on the findings of the survey, it is known which types of big data methods and technology companies tend to use, however, there is no indication of what value it actually brings other than it being a benefit or what that means in monetary terms.

Gunther et al. [4] accessed the literature on big data with the purpose of trying to understand how value can be realised. They defined three levels where literature debate is taking place: "Work practice level, Organisational level, Supra-Organisational level". At each of those there is different ways of taking advantage of big data and producing value. Furthermore, they identify two key features, Portability and interconnectivity as means of gaining value. In addition they suggest that linking between the different levels can generate further value from big data. However, they fundamentally lack a means of quantifying value, instead identify possible areas where value can be generated.

Portela et al. [5] suggested a Big data assessment framework (BIGDAF) based on the complexity of the Big data solution. The complexity is based on the dimensions of the data, Volume, Velocity, and Variety. Depending on the type of data that the organisation is dealing with a complexity level

will be defined between one and five. Each of the dimensions is multiplied by the complexity level category before adding all together. This will provide the final complexity level score. Based on this score the author suggests different types of big data solutions. While the framework certainly allows for a good assessment of the complexity of a Big data solution it has no capability of determining the cost or what the cost would be in relation to the potential benefit.

### *Big data feasibility studies*

feasibility is the study which takes place prior to the project to determine whether the study is viable. The study can take in consideration multiple aspects should be considered such as technology, legal, finance, etc. Below a few papers were summarised which performed a feasibility study on a big data project but failed to address the financial aspects of the project.

A feasibility report in the government sector of the EU [6] has demonstrated that a big data solution is required as the current sources of information obtained from stats are insufficient and provide an incomplete picture. Therefore, the report identified which of the stats provide beneficial information in addition to a set of new alternative information sources. However, this report does not go into the detail of the cost of integration those information sources. With regards to the obtained benefit, no more information is provided other than the EU wanting to get a full picture of the situation.

Li et al. [7] produced a case study on a feasibility study of big data in a library. They investigated what the potential opportunities and obstacles of big data are in libraries. They conducted a survey with librarians to understand the feasibility of these opportunities. The findings indicated that big data would enrich the library database, enhance the skills of librarians, promote interlibrary loan service and provide personalized knowledge service. Although it was concluded that big data would be feasible in libraries, the actual cost it is not established, and the somewhat minimal gains remain unjustified.

### *Economic feasibility in Big data*

Due [8] stated the key method of determining economic feasibility is cost benefit analysis method. Through the application of this method it is possible to take the potential

benefits, quantify the output and correlate it against the actual cost of the Big data solution to determine its economic feasibility for implementation and long-term ownership.

Throughout the literature several papers were found to describe the opportunities, benefits, challenges, and obstacles [9–11]. Furthermore, several papers and reports were found to have conducted feasibility studies, but none seem to have gone that extra mile and quantify the benefits and correlate those to the cost through a cost benefit analysis as part of their feasibility study. Therefore, there seems to be a gap in the literature with regards to whether Big data solution are financially viable in relation to the potential gain.

#### METHODS ANALYSIS

In this section we review the findings of the literature view from the lens of evaluating the cost of a big data project life cycle with regards to two factors, the life cycle cost opposed to the gained benefit in monetary terms.

Throughout the literature several papers were found to summarise the benefits, a framework was identified to access the complexity of the big data solution (BigDAF) [5] and to determine the financial viability cost benefit analysis was identified.

#### Benefits

The benefits identified in the literature are categorised into four sections [3], [12], [13]. First, transactional benefits include: “Saving on supply chain management, Reducing operating costs, Reducing communication costs, Avoiding the need to increase the workforce, Increasing return on financial assets, Enhancing employee productivity”. Secondly, Strategic benefit: “Creating a competitive advantage, Aligning IT with a business strategy, Establishing useful links with other organizations, Enabling a quicker response to change, Improving customer relations, Providing better products or services”. Third, Transformational benefit: “Achieving an improved skill level for the employees, Developing new business opportunities, Expanding capabilities, Improving business models”. Finally, Informational benefit: “Enabling faster access to data, Enabling easier access to data, Improving management data, Improving data accuracy, Providing data in more useable formats”.

#### BigDAF

The BigDAF [5] allows to assess the type of solution required. Depending on the metrics of the dimensions and the complexity level as demonstrated in table I the type of solution is calculated as defined in table II. The final score between 100 and 200 is a traditional BI issue, 200 -300 is a BI issue near to Big Data Challenge, 300 – 400 is a big data issue, and 400 to 500 is a complex big data issue.

BIG DATA COMPLEXITY FRAMEWORK

| CL/<br>Dimension | Volume         | Velocity       | Variety   |
|------------------|----------------|----------------|---|
| CL1              | <1000GB        | Batch          | Structured Data                                   |
| CL2              | 5TB - 50TB     | Intra-day      | Docs: XML; TXT, JSON                              |
| CL3              | 50TB - 500TB   | Hourly-refresh | Web-log; sensors and device events                |
| CL4              | 500TB - 2000TB | Real-time      | Image; social graph feeds; geospatial information |
| CL5              | >2PB           | Streaming      | Video; Voice                                      |

BIG DATA COMPLEXITY LEVEL (CALCULATION)

| CL/<br>Dimension | Volume  | Velocity | Variety |
|------------------|---|----------|---------|
| CL1              | $CL = (WVolume * CLx) + (WVelocity * CLx) + (WVariety * CLx)$ |          |         |
| CL2              |   |          |         |
| CL3              |   |          |         |
| CL4              |   |          |         |
| CL5              |   |          |         |

#### Cost benefit analysis

To determine whether the big data solution is economically viable a cost benefit analysis should be performed [8]. Although the benefits of a big data solution were previously identified, they are limited in applicability. Overall, benefits should increase profits or decrease cost, improve performance, which can be tangible or intangible [14].

However, so far, no indication has been provided on how to determine the cost of the big data solution which will be

incurred throughout its life cycle. For the cost we would consider factors such as development, maintenance, fixed, and variable costs [14], [15].

Once the benefits and cost have been established different methods can be applied to evaluate the correlation between benefits and cost to determine economic feasibility. Methods are subjective estimation, decision tree, or working back from the solution [14].

#### CONCLUSION

Big data solutions are taking every industry by storm providing great benefits and revolution. However, it seems that these solutions are growing organically rather than being planned for in advance ensuring that there will be a return on investment. Since there seems to be a lack of research on the matter, this paper has suggested a method based on the identified benefits of big data solutions, the BigDAF and the cost benefit analysis method, such that big data systems can be planned for in advance ensuring economic feasibility.

#### FUTURE RESEARCH

As this is a very brief research paper to establish a research direction, future research can look at reviewing the literature and ensuring that there is certainly no current research which looks at the economic feasibility of big data solutions. Furthermore, the suggested approach for accessing big data solution could be tested in empirical research.

#### RESEARCH LIMITATIONS

Some of the limitations identified in this paper are the dimensions of big data dimensions should be extended. Actual big data technologies should be reviewed and considered during the cost benefit analysis to understand the cost implications in correlation to the benefits. And a case study is required to evaluate a system its true benefits of big data in comparison to the cost

#### REFERENCES

[1] G. S. Sadasivam, M. Subrahmanyam, D. Himachalam, B. P. Pinnamaneni, and S. M. Lakshme, "Corporate governance fraud detection from annual reports using

big data analytics," *Int. J. of Big Data Intelligence*, vol. 3, 2016.

- [2] T. Ahlawat and D. R. K. Rambola, "Literature Review On Big Data," May 2016.
- [3] E. Raguseo, "Big data technologies: An empirical investigation on their adoption, benefits and risks for companies," *International Journal of Information Management*, vol. 38, no. 1, pp. 187–195, 2018.
- [4] W. A. Günther, M. H. R. Mehrizi, M. Huysman, and F. Feldberg, "Debating big data: A literature review on realizing value from big data," *The Journal of Strategic Information Systems*, vol. 26, no. 3, pp. 191–209, 2017.
- [5] F. Portela, L. Lima, and M. F. Santos, "Why Big Data? Towards a Project Assessment Framework," *Procedia Computer Science*, vol. 98, no. Supplement C, pp. 604–609, 2016.
- [6] "Feasibility study on data collection and analysis in the cultural and creative sectors in the EU," 2015.
- [7] J. Li, M. Lu, G. Dou, and S. Wang, "Big data application framework and its feasibility analysis in library," *Information Discovery and Delivery*, vol. 45, no. 4, pp. 161–168, 2017.
- [8] R. T. Dué, "Determining Economic Feasibility: Four Cost / Benefit Analysis Methods," *Journal of Information Systems Management*, vol. 6, no. 4, pp. 14–19, 1989.
- [9] S. Akter and S. F. Wamba, "Big data analytics in E-commerce: a systematic review and agenda for future research," *Electronic Markets*, vol. 26, no. 2, pp. 173–194, May 2016.
- [10] A. Alharthi, V. Krotov, and M. Bowman, "Addressing barriers to big data," *Business Horizons*, vol. 60, no. 3, pp. 285–292, 2017.
- [11] K. G. S. Somani, "Big Data Challenges: A Survey," *International Journal of Computer Systems*, vol. 01, 2014.
- [12] S. Gregor, M. Martin, W. Fernandez, S. Stern, and M. Vitale, "The transformational dimension in the realization of business value from information technology," *The Journal of Strategic Information Systems*, vol. 15, no. 3, pp. 249–270, 2006.



- 
- [13] S. J. Ren, S. F. Wamba, S. Akter, R. Dubey, and S. J. Childe, "Modelling quality dynamics, business value and firm performance in a big data analytics environment," *International Journal of Production Research*, vol. 55, no. 17, pp. 5011–5026, 2017.
- [14] "Feasibility and Cost-Benefit Analysis." .
- [15] P. Katimuneetorn, "Feasibility Study for Information System Projects." .