

## Big Data Analytics in Islamic Banking

Jamil Sultan <sup>a</sup>, Clemens Bechter <sup>b</sup>

<sup>a</sup> *Dual Studies Program, Al-Quds University, Jerusalem, Palestine.*

<sup>b</sup> *Thammasat Business School, Bangkok, Thailand.*

---

### **Abstract**

Islamic financial institutions must comply with the Shari'ah law. This makes them distinctly different from conventional Western banks. Big Data Analytics (BDA) is a technology and as such regarded as culture-free. There are a few pilot BDA projects in conventional banks but none in Islamic banking. Our paper fills a research gap by analyzing possible applications and implications of BDA for Islamic banks and its regulatory institutions. Our field research showed that the benefits of BDA are already here or will flourish within the coming five years. Implications of our research are that Islamic Banks must start drafting a roadmap for their organization. Risk management is one area where BDA can be of great benefit.

**Keywords:** Big Data Analytics, Islamic banking, Hadoop, MapReduce, machine learning, risk management.

---

## **I. INTRODUCTION**

In contrast to the common meaning of the word 'big', Big Data does not necessarily mean only huge amounts of data. There are additional criteria. Traditionally, data has been kept in databases to serve applications. When the volume becomes too large, the database becomes unmanageable and the response time of applications becomes very slow. Google was faced with this exact problem, therefore the MapReduce algorithm was introduced. The basic idea behind this was to dissect a task into small parts and let several different computers on a network handle a small part of the overall job each. At the end, the results of each small job would be combined and the resulting dataset is compiled. Based on Google's MapReduce, Apache Hadoop today is the most widely implemented platform for Big Data. Apache Hadoop consists of a data storage layer called Hadoop Distributed File System (HDFS) and a data processing layer called Hadoop MapReduce Framework. The system is different from peer-to-peer systems such as blockchains which are also distributed (ledger) systems. Big Data happens within an organization – not publicly. There are private blockchains of course, but they lack the analytical component. Hadoop can handle unlimited numbers of parallel analytical tasks and is, therefore, the leading infrastructure. The emphasis of Big Data is on the fact that real-life data stems from different sources which include structured (e.g. numbers) as well as unstructured data (e.g. pictures). All data, no matter whether structured, semi-structured or unstructured, can be analyzed by Big Data Analytics (BDA) in order to give solid recommendations that may even go beyond human analytical capabilities.

The rest of the paper is organized as follows:

Section II: Literature review on the topics of BDA and an analyzation of the Islamic banking laws.

Section III: Field research results and their discussion.

Section IV: Concluding remarks.

## **II. LITERATURE REVIEW**

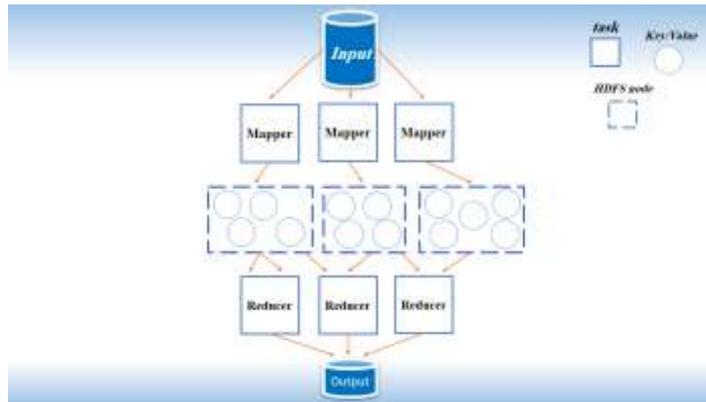
### **2.1 BDA**

There are three main types of analytics: descriptive (simply describing what has happened), predictive (forecasting events), and prescriptive (how should an organization such as a bank prepare for the future). An example of analytics is the following: a client is regularly accessing the bank's website to check exchange-rates (descriptive), something s/he has never done in the past (descriptive). This could mean that s/he intends to buy or sell foreign exchange (predictive). A targeted email or call should be the resulting action (prescriptive) to offer foreign exchange services. Machine-learning is one technique to extract meaningful information out of data and making predictions based on past occurrences. Deep neural networks are a set of machine learning algorithms that are designed to recognize patterns <sup>[1]</sup>. In order for those patterns to be recognized, they must have numerical format, which means that all non-numeric data must be converted into vectors. A good example would be the diagnosis of a patient's disease and prescribing the appropriate therapy. Data would come from different sources such as X-ray machine, medical records, fitness tracking devices, oral communication with the patient, and even social networking activities. The X-ray images and facial expressions of the patient would need to be converted into a numerical format.

The three Vs - Volume, Variety and Velocity - are characteristics of Big Data. Volume refers to the large amount of available data, variety refers to the structured and unstructured nature of these data and velocity

to the speed of the processing the data (from batch to real time). Besides the three Vs, there are additional characteristics such as Variability which refers to possible inconsistencies within the different sets of data. Veracity is the trustworthiness of the data sources and Complexity refers to the complex processing of the data.

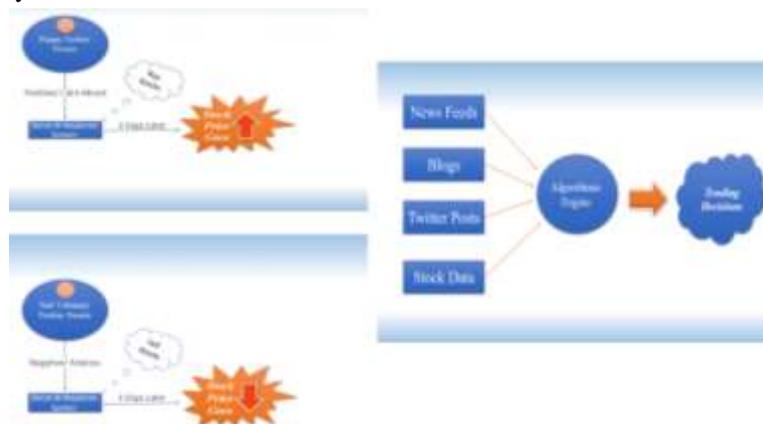
Most popular Big Data tool is opensource Hadoop. See Figure 1.



**Figure 1. Hadoop architecture**

Hadoop's Mapper generates an intermediate key-value pair. HDFS is the primary data storage system. It employs a NameNode and DataNode architecture to implement a distributed file system. The Reducer shuffles and sorts the results. Shuffling bundles similar types of works into single units and sorts the data into a logical order.

Existing commercial implementations of BDA focuses on product innovation, strategy development, market segmentation and growth, compliance with regulations, and intelligent solutions such as faster and better decision making resulting in a closer market orientation [2]. BDA is slowly making inroads into the banking sector. Figure 2 shows a simple example of a stock exchange decision based on an algorithmic engine which analyzes data from various sources.



**Figure 2. Stock trading decision supported by BDA**

One source of making a stock investment decision is Twitter. Tweets can influence the share price. The algorithmic engine of BDA will not just base a decision on the analysis of tweets. It will also look at news feeds, blogs, and of course financial data such as ROI, ROA, free cash flow, debt ratios etc. By doing so, BDA will need to combine numerical stock data with unstructured texts.

Looking at banking services as a whole, input data range from call logs and credit history to social media, see Figure 3. Predictive analytics should forecast the risks associated with a client. Same applies to capital markets. Lots of structured, semi-structured and unstructured data are needed for a reliable forecast.

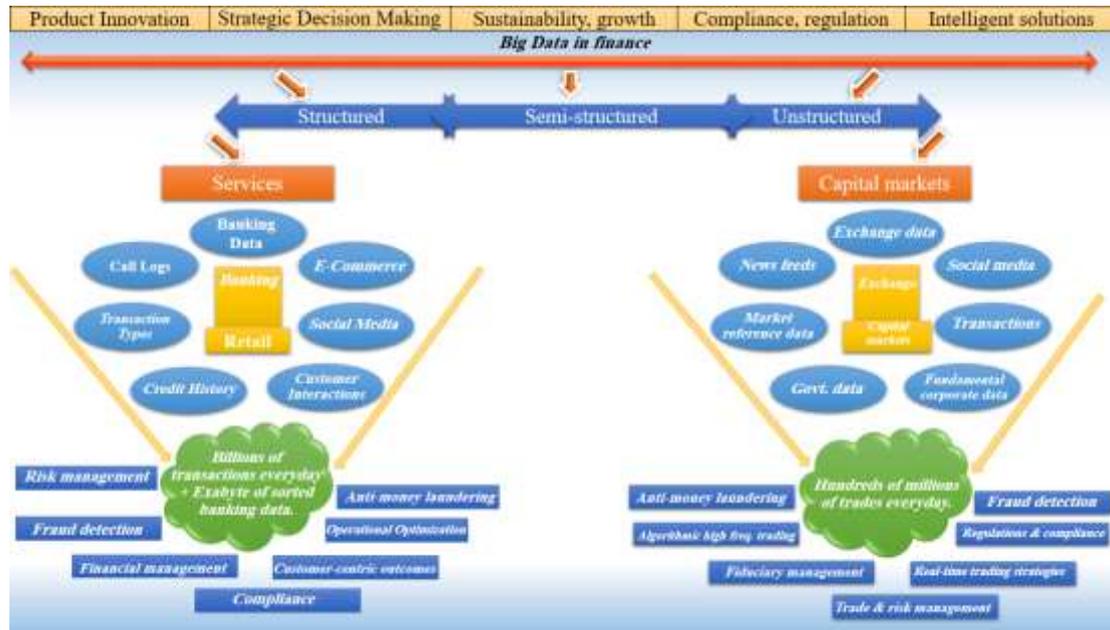
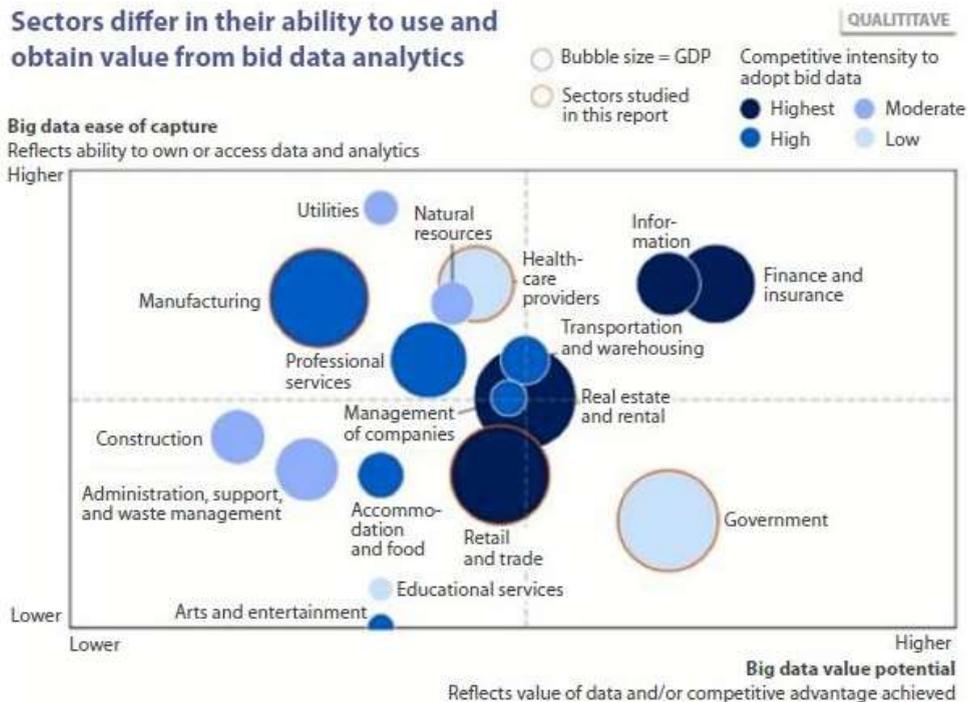


Figure 3. BDA in Finance

Figure 3 illustrates the impact of the three Vs. BDA has to deal with data volumes from very different sources in different formats. Results should be presented in real time to make instant decisions. Lehrer et al. <sup>[3]</sup> see potential for BDA in personalization of services and creation of innovations. Martens et al. <sup>[4]</sup> stressed the importance of detailed data. Their research showed that BDA does not automatically achieve better results. Only when the data goes to a very deep level, then BDA can show its full predictive power. Not all sectors will benefit from the use of BDA in the same way. Looking at Figure 4, the finance sector may benefit the most.



**Figure 4. Beneficiaries of BDA (source: US Bureau of Economic Analysis; McKinsey Global Institute Analysis)**

The value of BDA for the manufacturing sector seems well below the finance sector. Pérez-Martín et al. <sup>[5]</sup> outlined how BDA can for example measure credit banking risks for mortgages.

## 2.2. Islamic Banking

Literature on the possible impact of BDA on the Islamic financial sector is sparse. Islamic finance differs in various aspects from conventional Western finance:

- “Islamic finance refers to the provision of financial services in accordance with Islamic jurisprudence (Shari’ah)” <sup>[6, p.9]</sup>. Islamic finance is a broad sector which encompasses any financial activity that respects the underlying criteria defined by the Shari’ah law.
- Overall market shares are: 80% is accounted for by Islamic banks, 15% Sukuk (Islamic bonds), 4% Islamic mutual funds and 1% Takaful (Islamic insurance) <sup>[7]</sup>.

Main regulations that must be observed are:

- Prohibition of “Riba” which can generally be translated by interest <sup>[6]</sup> or, to be more precise, “the term riba, which means ‘hump’ or ‘elevation’ in Arabic, is an increase in wealth that is not related to engaging in a productive activity” <sup>[6, p.4]</sup>. Riba prohibition has been explained by “Islam considers money a mere means of exchange, without intrinsic value. Since its circulation does not reflect a real economic activity, therefore, it would be immoral for it to yield any premium whatsoever” <sup>[6, p.4]</sup>.

- Prohibition of products with excessive uncertainty “as manifested by contract ambiguity or elusiveness of payoff”<sup>[6, p.4]</sup>, gambling (Maysir) or short sales not abiding to specific rules<sup>[6]</sup>.
- Prohibition of harmful, immoral or illicit activities e.g. “Islam does not endorse every human wish, and it prohibits on moral grounds activities related to tobacco and other drugs, alcohol, pork products, gambling involving money and non-money assets (Maysir), speculation, pornography, and armaments and destructive weapons”<sup>[6, p.3]</sup>.

However, as many large firms receive a modest proportion of income from such prohibited activities (for instance, hotel chains with alcohol sales), modern Shari’ah scholars tend to allow investment in companies with tolerable proportions of revenues from prohibited activities under the condition of ‘Haram’ purification, which comes down to calculating the portion of these illicit revenues and compensating for it by donating them to charity<sup>[7]</sup>.

Only Iran and Sudan rely exclusively on Islamic finance for their banking sector, while in other countries Islamic banks compete with conventional banks. Thus, even in very conservative and religious countries such as Saudi Arabia, only 35% of the banking sector assets are Shari’ah compliant. This proportion stands at 20% in countries such as UAE, Qatar, Palestine and Malaysia<sup>[8]</sup>.

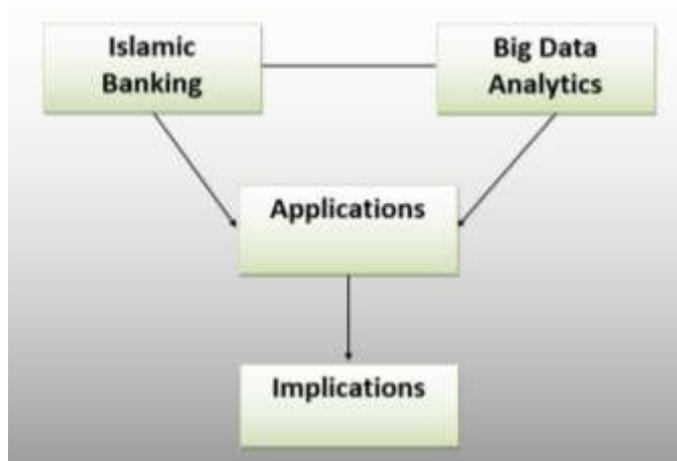
In essence, Islamic banks are neither permitted to charge interest of any kind nor to shift the entire risk to their customers. On contrary, they are expected to charge moderate administrative cost and to share risks as much as possible<sup>[9]</sup>. Banking scholars have tried to compare the actual performance of Islamic banks with conventional ones, but no clear consensus seems to be emerging out of these studies<sup>[10, 11]</sup>.

In an IMF staff discussion note, Prasad et al.<sup>[6]</sup> suggested that the Islamic finance, because of its risk sharing feature, entices a better risk management from both the financial institutions and their customers and thus Islamic finance is less prone to crisis. The note added to explain this hypothetical higher resilience as follows “(i) Islamic finance involves prohibitions against speculation; (ii) financing is asset-based and thus fully collateralized; and (iii) it is founded on strong ethical precepts”<sup>[6, p.8]</sup>, all of which could make the Islamic finance more stable and resilient. However, the last two financial crises have shown that asset-based lending (e.g. mortgages) can trigger a financial crisis. Cash-flow based lending might be a better way.

From a regulatory perspective, the Basel Committee on Banking Supervision (BCBS) seems to be accepted globally as an independent watchdog. In addition to Islamic banking, there are the Islamic Financial Services Board (IFSB) and the Accounting and Auditing Organization for Islamic Financial Institutions (AAOIFI)<sup>[12]</sup>. Regulatory bodies rely commonly on CAMELS (Capital-Asset-Management-Earnings-Liquidity-Sensitivity) rating systems to assess the health of a bank<sup>[13]</sup>.

Only a few authors investigated the possible impact of BDA on Islamic banking. Ridhwan et al.<sup>[14]</sup> recommend BDA for risk management. Not only assessing the financial risk of a client, but also the risk of non-compliance with Shari’ah law. Johnson<sup>[15]</sup> argues that BDA may lead to more financial market stability in the context of Islamic banking. Ogden<sup>[16]</sup> sees the potential of BDA in identifying new business opportunities, new market segments and more targeted marketing activities. Javed et al.<sup>[17]</sup> developed an ontology-based framework to increase the effectiveness of approach to enhance the effectiveness of BDA, especially the data retrieval. Haseeb<sup>[18]</sup> sees the risk that Islamic banks will be falling behind if they are no not react soon and take serious actions concerning BDA. However, even financially advanced economies, such as the UK, do not make widespread use of BDA. Empirical

research by Grishikashvili and Bechter <sup>[19]</sup> have shown that large financial groups are just at initial launching stage, and financial advisors are most likely not making full use of the BDA systems provided.



**Figure 5. Research Framework**

Figure 5 shows our research framework. Possible applications of BDA have to take into account the regulations that are specific to Islamic banking, e.g. the Shari’ah law. After having identified potential applications, the implications for clients and society at large was assessed. We also looked into the challenges that come with BDA.

The objectives of our research are:

- Assess perceptions of Islamic bank’s employees on BDA.
- Evaluate the opportunities of BDA for Islamic banking.
- Outline how sentiment analysis can be used for systemic risk management.

The limitation of our research is in the small size of the sample. One may characterize our findings as preliminary.

### **III. FIELD RESEARCH RESULTS AND DISCUSSION**

In autumn 2018, we asked the staff of the Islamic banks to fill in an online questionnaire. The respondents were familiar with BDA from a theoretical perspective. We received and analyzed 17 responses. The number seems small, but most employees are still unfamiliar with BDA, so it would not make much sense asking them.

The first set of questions inquired about the value of BDA and when it might flourish.

Three choices were given: First, it would create value now, second, it would do so in 5 years, third, it may never do so. or never. Based on our literature review, we identified the following areas where BDA can improve customer segmentation, loyalty and retention. It can also improve service/product design, better risk/financial management, identify new business models, time saving, faster decision making, and better decision making. Table 1 shows the results.

**Table 1. Top Five Benefits**

Benefit:	now	In 5 years
faster decisions	76.50%	23.50%
better decisions	64.70%	35.30%
market segmentation	70.60%	23.50%
risk management	58.80%	41.20%
design new products	58.80%	41.20%

The findings are in line with current research and literature, apart from risk management. This is very banking specific.

The second set of questions dealt with the likely challenges associated with the introduction of BDA. Challenges were also based on the literature review. They include: timeliness, overwhelming data volume, managing unstructured data, data quality, availability of data, access rights to data, data ownership issues, cost of data, lack of facilities, IT infrastructure, lack of pre-processing facilities, lack of technology, shortage of talent/skills, privacy concerns and regulatory risks, security, and corporate culture. All items were ranked from 1 (major challenge) to 5 (no challenge at all).

**Table 2. Challenges**

	N	Mean	Std. Deviation
facil.	17	1.76	1.091
volume	17	1.82	1.131
pre_proc	17	1.82	0.728
techno	17	1.94	0.827
quality	17	2	1.173
owner	17	2	0.866
available	17	2.06	1.197
security	17	2.12	0.857
HR	17	2.12	0.781
timeliness	17	2.18	0.883
unstructured	17	2.18	1.015
rights	17	2.24	0.831
cost	17	2.29	0.92
privacy	17	2.29	0.849

As top challenges were perceived, the lack of facilities, volume and pre-processing of data, (see Table 2) while the privacy issues and the costs involved seem to be less challenging. A factor analysis resulted in four components. The variance explained by each component varies, see Table 3. Factor 1 can explain the most with 27%.

**Table 3. Variance explained**

Component	Initial Eigenvalues		Rotation Sums of Squared Loadings		
	Total	% of Variance	Total	% of Variance	Cumulative %
1	6.743	48.166	3.778	26.984	26.984
2	1.889	13.494	3.313	23.667	50.651
3	1.447	10.335	2.264	16.171	66.823
4	1.385	9.895	2.109	15.066	81.889

Table 4 shows the factor loadings. We regarded values above 0.7 as high. When looking at the high loading variables of factor 1. Timeliness relate to velocity, availability to variety and the third V is volume, quality refers to veracity. The second factor refers to often overlooked challenges such as human resources and legal (ownership) rights on the data. Factor three represents challenges in terms of lack of facilities. The least important factor covers security and privacy.

**Table 4. Rotated component analysis**

	Component/Factor/Dimension			
	Classic Vs	HR & Legal	Facilities	Privacy & Security
timeliness	0.866	0.07	0.165	0.02
available	0.85	0.318	0.106	0.087
volume	0.817	0.18	0.153	0.282
quality	0.741	0.421	-0.003	0.216
unstructured	0.672	0.482	0.383	0.17
HR	0.149	0.88	0.167	0.134
rights	0.15	0.799	0.239	-0.064
cost	0.24	0.781	0.219	0.206
owner	0.283	0.712	-0.026	-0.052
facil.	-0.023	0.189	0.931	-0.051
pre_proc	0.492	0.104	0.763	0.036
techno	0.338	0.407	0.666	0.4
security	0.054	-0.031	0.146	0.955
privacy	0.309	0.148	-0.094	0.896

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Most literature highlights technical challenges. Our findings suggest that personnel and legal issues are also very important as well as privacy and security.

#### IV. CONCLUSION

Our field research has shown that BDA offers tangible benefits which already exist or will show up within the next five years. Non-technical challenges must be addressed. Similar to the introduction phase

of an ERP system, there must be a clear roadmap for the implementation and the involvement of all level, not just the management.

BDA can help banks to identify sanctioned, blacklisted individuals, and companies who often have various identities which are difficult to trace. This can be done in almost real time and at minimal cost. Fraudulent transactions can be identified by highlighting unusual transaction patterns.

In addition, other risks can be identified early since multiple input parameters can be taken into the overall risk assessment. Risk can be a non-performing loan as well as total bank failure. Faster and better investment decisions (e.g. stock trading) can be done. Historical data and latest news can be combined, and recommendations given in real-time due to the parallel execution of different tasks.

Individual customer behavior can be analyzed and predicted whether s/he is satisfied or is considering switching to another bank. Such sentiment analysis can be based on phone conversations, emails and other ways of communication with the client. This does not only apply to individuals, but also to banks. Analyses of postings on social media can indicate people's perception of the bank, or even the whole banking sector. The perceptions of clients on new banking services can be assessed by analyzing social media in combination with own desk and field research.

Since banks know the purpose and the locations where credit cards are used, they could cooperate with retailers to suggest personalized offerings.

Another area is systemic risk assessment. Large banks form a systemic risk for the whole banking sector. The traditional approach of calculating ratios (e.g. capital adequacy, Basel III) may not prevent disaster because they are based on probability distributions and many transactions happen off the books. Even ratings by rating agencies can give a false sense of security as the case of Lehmann Brothers has shown. The authors recommend a sentiment analysis to complement traditional ratios and ratings. This can be done by analyzing tweets daily i.e. the sentiment towards a bank can be measured day by day, and therefore, the sentiment's ups and down will be gathered over time. As on the web in general, there are tweets from authoritative and independent sources (e.g. Reuters) but also fake news. Looking at trustworthy tweets a sentiment analysis can be conducted. Sentiments can be bullish, bearish, optimistic, pessimistic etc. The allocation to a sentiment can be dictionary based i.e. a certain word stands for a certain sentiment. Emoticons, also, express sentiments. An overall positive or negative sentiment will be part of the assessment of risks.

## REFERENCES

- [1] LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning, *Nature*, 521(7553), 436-444.
- [2] Dumbill, E. (2012). *Planning for big data*. O'Reilly Media, Inc.
- [3] Lehrer, C., Wieneke, A., Vom Brocke, J., Jung, R. & Seidel, S., (2018). How Big Data Analytics Enables Service Innovation: Materiality, Affordance, and the Individualization of Service. *Journal Of Management Information Systems*, 35(2), 424-460.
- [4] Martens, D., Provost, F., Clark, J., & de Fortuny, E. J. (2016). Mining massive fine-grained behavior data to improve predictive analytics. *Quarterly*, 40(4), 869-888.
- [5] Pérez-Martín, A., Pérez-Torregrosa, A., & Vaca, M. (2018). Big Data techniques to measure credit banking risk in home equity loans. *Journal Of Business Research*, No. 89, 448-454.

- [6] Prasad, A., Tow, C., Pinon, M., Norat, M., Kammer, A., Zeidane, Z., Aljabrin, S., Sul Aiman Alsaeed, K., Arda, A., Awad, R., Barajas, A., Chartouni, C., Chung, K., El Hamiani Khatat, M., El Khoury, C., Galand, J. Harutyunyan, Artak, H, Y, Jansen, B & Kargbo-Sical, D. (2015). Islamic Finance: Opportunities, Challenges and Policy Options. *Staff Discussion Note*, International Monetary Fund.
- [7] Abedifar, P., Ebrahim, S. M., Molyneux, P. & Tarazi, A., (2015). Islamic Banking and Finance: recent Empirical Literature and Directions for future Research. *Journal of Economic Surveys*, 29(4), 637-670.
- [8] Hussain, M., Asghar, S. & Turk, R., (2016). An Overview of Islamic Finance. *Journal of International Commerce, Economics and Policy*, 7(1).
- [9] Sirajo, A., M. Kabir, H., Rosylin Mohd, Y. & Nasri, N., (2017). Islamic Banking Sustainability: A Review of Literature and Directions for Future Research, *Emerging Markets Finance & Trade*.
- [10] Beck, Thorsten, Demirgüç-Kunt, Asli, and Ouarda Merrouche. (2013). Islamic vs. conventional banking: Business model, efficiency and stability, *Journal of Banking and Finance*, No. 37, 433-47.
- [11] Zaheer, S. & Moazzam F., (2014), Liquidity Crisis: Are Islamic Banking Institutions more Resilient?, paper presented at the *IMF Conference on Macroeconomic Challenges Facing Low-Income Countries*, Washington, DC, January 30-31.
- [12] Song, I. & Oosthuizen, C., (2014), Islamic Banking Regulation and Supervision— Survey Results and Challenges, *IMF Working paper* No. 14/220.
- [13] Archer, S., & Karim, R. A. (2012). The structure, regulation and supervision of Islamic banks. *Journal Of Banking Regulation*, 13(3), 228-240.
- [14] Ridhwan, M. S., Khairuddin, N. D., & Suhaimi, M. A. (2014). Big Data Technology in Shari'ah Compliance Risk Management (SCRM): A View of Information Technology in Islamic Finance Risk Management. *Journal Of Islamic Banking & Finance*, 31(3), 84-97.
- [15] Johnson, D. (2015). Advancing Islamic finance. *Asian Investor*, No. 40.
- [16] Ogden, G. (2015). Islamic financial institutions need to jump on the Big Data bandwagon. *Islamic Business & Finance*, No. 91, 48-49.
- [17] Javed, A., Othman, R., & Youji, K. (2016). Conceptualizing the Ontological Framework for Service Dominant Logic for Islamic Banking and Finance. *Journal Of Emerging Issues In Economics, Finance & Banking*, 5(1), 1776-1792.
- [18] Haseeb, M. (2018). Emerging issues in Islamic banking & finance: challenges and solutions. *Academy Of Accounting & Financial Studies Journal*, No. 22, 1-5.
- [19] Grishikashvili, K., & Bechter, C. (2018). Perceptions of independent financial advisors on the usefulness of Big Data in the context of decision making in the UK. *International Journal of Big Data Intelligence*, 5(3), 1-12.