Big Data: New Development Opportunities in Islamic Studies

Author: Sally Kawtharany (Lebanon)[1]
Reccive: 04/01/2017 **Accept**: 10/08/2017

Abstract

A Quranic researcher's personality should be identified by the scientific characters because his/her efforts should be based on evidence and truth to acquire knowledge, And without any doubt, the basic role here is to transform the Quranic givens to realistic values that shapes the society into a better place, by trying to provide solutions for the problems and crises that our society is passing through, and develop our society's issues based on the Quranic visions. Moreover, based on our belief that it is necessary to start the development from the needs of the society, and in order for the Quranic researcher to complete this role in the perfect and most delightful way, this paper represents a new flame that will surely help in strengthening the researcher's ability, and in providing new technological techniques and tools of “Data Science” and specifically “Big Data”, in order to provide Quranic students and institutions a huge power that will lead to analyze the new era's situations and needs. Data Science is the process of extracting data from different sources to detect the directions and visions that represents it. This whole process leads us to produce and come up with the right techniques and decisions in any issue that is tackled. However, when such science is putting in the hands of a Quranic sciences researcher, we will be investing a productive mentality, and getting a step closer to treat the polarity that the Muslim lives trying to connect between theories and the practical reality.

1. Trainee in Data Science and Machine Learning Fields, Lebanon, *sallykwtha@gmail.com*

This paper studies the greatest profit that can be produced by the Data Science in understanding and analyzing the Islamic reality. This paper can also be a dependable research that a researcher can rely on through his/her Quranic studies so he/she can start with eliminating the society's worries through knowing the major issues to be handled, detecting the most popular Quranic-understanding misconceptions, analyzing the cultural and social backgrounds that affects the Islamic area, and studying the correlation for different phenomenon. We emphasize in this paper the new role that the Quranic mentalities of students and institutions should play on account the digital civilization that we are facing, through getting the help from the Quranic culture and education first, and the technological revolution second, specifically the Data Science. Since this position is the first of its kind in the Quranic sciences field, we hope that it resembles a start in the path of development and improvement so this goal can be achieved in the few coming years.

Key Words: Big Data, Development Ppportunities, Islamic Studies

Introduction

Nowadays, we are living a digital civilization. We can name it as "undisguised" one, where everything is being in public, news, information and opinions, even the daily life personal events, this seems advanced, but actually we are often using it inappropriate, which leads to a negative mind programming, brainwashing, and deterioration of human values.

Passing through old era's representatives, we observed that every powerless civilization would be destroyed, so that a new civilization would be born. However, this indicates that in our days, human sciences productivity will more expand intercontinental especially the religious science, and we will not be surprised by its wide spread, where the digital technologies and techniques would play a vital role, specifically Data Science.

Thus, in front of this transitive period, the Quranic mentalities of researchers and institutions should invest in Data Science, so that we can develop the creativity of productive minds.

Quran and Reality Needs

The ideal human that Quran describes, is a highly qualified scientist in every life details, where he consists in indications and facts. And that's how the Quranic researcher is.

The most important function for the Quranic researcher is collecting data givens about all reality needs, and then tries to find its solutions through Quran, which was obvious in what prophet Mohammad (S) have done. The Islamic thinker, Mohammad Abu Alkassem Haj Hamad, affirms for Muslims, the importance of facing the social crisis by using the Quranic principles.

He mentioned in his book (The knowledgeable Approach of Quran): "The prophet's mission was to show, clarify and teach people how to reveal Quran principles on real life"[1] and how the Quran conducts the development of knowledge approach dimension, in all fields.

However, our Islamic intellectual realities, in general, haven't achieved this objective yet. Where lots of efforts are being lost in meaningless debates and others are exhausted in profitless subjects and discussions.

In these days, the rift between reality and the theoretical concepts, that the Muslim is living through, is because of the distance between the Quran scientists and the needs of Muslim individuals.

1. Abu Alkassem, 2003, p. 13.

This gap, in addition to raising motivation and cooperation between individuals and religious institutions, requires to start from the scientific information and datum.

Therefore, displaying positive role models would be boosting motivation in providing a guide to achieve that.

One of these models is the theologian Ayatollah Mahmoud Taleghani (1911-1979) who believed that the movement in the midst of reality concerns and events of the world is the best way to understand the Quranic purposes.

Thus he said: "some of the Quranic cognitions that we often don't pay attention for, is that some of Quranic issues and facts are revealed to man through movement, and not through the view of this interpretation and that, and the view of this opinion and that, or through the combination of them.

I present to you an example: I have written the interpretation of "Al-Nazia'at" as seen by people in the interpretation of "Qabas of the Quran", but when I went out from prison and saw this movement and revolution, new issues were emerged in front of me from Al-Surah and this is the richness of Quran."[1]

We conclude that the fundamental key is relying on established facts, including the significance of starting from reality needs, understanding, and studying it, so that the Quranic researcher can perform his role and responsibility by processing creative minds in the discovery of daily life complication. And then works to find resolutions from the Quranic perspective.

1. Khoshmanesh, 2010, p. 47.

Therefore, this paper presents the necessity of providing new digital tools and techniques in Data Science, which gives researchers and religious educational institutions a tremendous power in diagnosing and analyzing the circumstances and needs of this age.

And even if we couldn't immediately involve Data Science in the religious curricular, since it needs to prepare enough time to provide the necessary resources and capabilities, but the educational institutions of Quranic sciences can play today this role and develop the research department, because this contributes to improve their curriculums and carry out the priorities for researchable topics that their students and researchers work on.

Data Science: An Overview

By 2020, the number of connected devices in the internet is expected to reach 60 billion, the equivalent of 6.58 per person[1] (this forecast is pretty conservative).

This development is known as the internet of things (IoT), or what is now called internet of everything (IoE), in which it is changing radically our lives. In view of this increasing interaction between individuals and devices components, there is a need for serious research trends that can help specialists in every field, business sector or subject of study to work on digital methods and tools to extract all data available for analysis; Whether it is the objective of a specific problem analysis to formulate appropriate solutions, or develop predictions in order to make strategies.

1. Cisco's forecast, Entitled, "Number of connected objects expected to reach 50bn by 2020".

Each interaction between the individual and the internet-related devices is an additional opportunity to measure and understand the individual's behavior.

"Big Data's ascendancy represents three shifts in the way we analyze information that transform how we understand and organize society".[1]

It's called Big Data relative to the huge size, the universality of types, in addition to the constant change of the data. Indeed we can imagine billions of tremendous available content through the Internet.

Such as Open Data sources, for those data that are amenable to scraping, like social network data, in which different programming tools can extract. But there are private data that are concerned to specific sides, like owners of websites, phone applications, web hosting, etc. or any entity that have developed sensors technology[2] to collect data in all its forms; audio, visual, written or otherwise.

And for the importance of differentiating between Big Data and traditional database terms, we identify the content of database as records or inputs, while Big Data may include e-mails, audio files, videos, online clicks, tweets, transactions, sensor readings, etc.

The database helps us to understand and analyze a past situation, while Big Data contributes to predict the future. Actually, we can say that Big Data is everywhere. But there is no sense in compiling data without turning it into meaningful indicators, because data is a fancy word without scientific analysis.

---

1. Mayer-Schönberger and Cukier, 2013, p. 12.
2. A sensor is a device that detects and responds to some type of input from the physical environment. The specific input could be light, heat, motion, moisture, pressure, or any one of a great number of other environmental phenomena.

Therefore, we present in this paper the concept of Data Science, because it is the one that reveals the significant observations expressed by the data sets from different sources. This will lead us to make appropriate decisions and strategies in any subject given to be processed.

At the present time, social networks play a key role in the enhancement of Big data sets.

Twitter, for example, is a very useful source of generating data, and can be used in the Sentiment Analysis, such as the feedback of any product or service, elections, marketing etc.

Sentiment Analysis is one of the aspects addressed by the Data Science that depends on Big Data. It is the process of extracting and classifying data according to the opinions, feelings and attitudes such as positive, negative or neutral, in order to employ it in rehabilitating specific issues, and maintaining decision-making processes on the basis of scientific givens alternative to guesswork and hypotheses.

Consequently, contemporary companies rely their investments on Big Data. For example, Amazon uses Big Data to deliver better service for its customers. If a customer purchases an item, he will automatically receive a suggestion items list that might concern him. This technique is based on the behavior analysis of all customers.

In this case, Data Science intersects with Machine Learning[1], which is also a fertile research area for the development of online religious learning, but what concerns us here is one of the distinctive features of Big Data.

1. Machine learning is a method of data analysis that automates analytical model building.

It should be noted that data Science is exploding in the business sector. Even though, Data driven decision making (DDDM) has been around for a long time, however, the sources of the assessment information were different in the past. But the old tools do not seem to be enough anymore, we must balance them with the new data technologies.

Human capacity simply cannot handle the amount of data with which we are being confronted. Thus, businesses have turned to a variety of technology solutions to help them deal with this increasing data load. But because Data Science is a big field, this paper is only concerned on specific purpose, which is how Big Data improve the Quranic research, such as detecting the most popular Quranic misconceptions, analyzing the cultural and social backgrounds that affects the Islamic area, and studying the correlation for different phenomenon.

Quran and Data Science

Because that we are Quran science students, we are interested, in this paper, to combine Quranic vision with our topic, since the Data Science is included by the tenor of Quran.

There is perceived lack of Quranic interpretation interest related to research in Data Science, but there are few researches that started appearing, such as “ICT and Islam” book which was written by Dr. Mohammed Fouzan Nour Al-din, it includes Quranic vision toward privacy, data security, and technology age.

As well, he has involved in many papers that contribute to generate clearly the relation between Data Science and Quran. In one of these papers, he confirms that the process

of Big Data within its concepts has marked by the Quran knowledge approach, from more than 1400 years ago.

"If we take "Big Data" in the context of the Qur'an, it would include data about everything that exists, encompassing the whole universe and everything that exists inside of it. The size of the "Big Data" that exists in the form of digital data in our modern world would not even add up to a fraction of the "Big Data" that the Qur'an talks about."[1]

At this point, we are looking forward to increase this kind of researches, by interpreting the fundamental Quranic principles related to Data Science, as long as this Science is the primary core in society development, where the governments, companies and individuals depend on it. Therefore, we are proposing a Quranic verse interpretation, as a sample, which lead us to understand the Big Data from the divine scale:

«وَعِندَهُۥ مَفَاتِحُ ٱلْغَيْبِ لَا يَعْلَمُهَآ إِلَّا هُوَ ۚ وَيَعْلَمُ مَا فِى ٱلْبَرِّ وَٱلْبَحْرِ ۚ وَمَا تَسْقُطُ مِن وَرَقَةٍ إِلَّا يَعْلَمُهَا وَلَا حَبَّةٍ فِى ظُلُمَٰتِ ٱلْأَرْضِ وَلَا رَطْبٍ وَلَا يَابِسٍ إِلَّا فِى كِتَٰبٍ مُّبِينٍ»:

And with Him are the keys of the unseen; none knows them except Him. And He knows what is on the land and in the sea. Not a leaf falls but that He knows it. And no grain is there within the darknesses of the earth and no moist or dry [thing] but that it is [written] in a clear record.[2]

This verse indicates how the knowledgeable authority controls our universe, and also shows us the ability to recover any data sets for any specific thing, as seen on the representation of leaf falling.

1. Qadri and Others, 2015.
2. Surah Al-An'am, 59.

In addition to this general view, we are contrived through it, the importance to acknowledge every single slight data set without any exception, because it would be related with other's effects, they are all into one united complication system. And also, the significant role of data sets classification, as seen by the verse's example (moist or dry).

This short approach has to say that researchers would be more concerned in this additional Quranic features during their research path.

They could base on the Quranic view to proceed within conditional treatment concerning any kind of data. So that they apply unified Quranic cultural context of research codes, either during the social givens needs diagnosis primary stage, or eventually, the interpretation of Quranic verses searching for specific issues.

Text Data Mining and Natural Language Processing (NLP)

According to the annual infographic report released by Domo[1], titled by "data never sleeps", which started since 2013, where its fifth version in this year mentioned that people send more than 15.2 million texts per minute everyday around the world, this was duplicated 334 % time from the previous report.

Against this new reality, the importance of Text Mining increased, where Data Scientist's tasks include: clustering, classification, entity extraction, Sentiment Analysis, and modeling correlations, etc.

Hence the Quranic student or researcher can rely on Big Data analytics, which is considered as the richest source that helps in understanding more the world, instead of using the traditional research techniques (interviews,

1. Domo, is a computer software company based in United States. It specializes in business intelligence tools and data visualization.

questionnaire forms, traditional references, etc) "Text Mining brings together a broad range of contemporary qualitative and quantitative methods to provide strategic and practical guidance on analyzing large text collections."[1]

Text Mining analysis allows the researcher to obtain accurate study, by realizing huge data sets. But that requires many additional technical skills, as well as specialized collaboration.

Furthermore, using Text Mining prompts us to ask new kind of examination, as how can we present Big Data in meaningful way, amenable to understand its indicators, and what it can afford to measure within any determined issue.

Moreover, Natural language processing (NLP) is a component of Text Mining that performs a special kind of linguistics analysis, such as summarizing blocks of text, generating keyword tags, reducing words to their root, identifying the type of entity extracted (it being a person, place, or organization, etc.), or Sentiment Analysis, and many more.

NLP allows for researchers to study the context of texts meaning, by using for example hashtags to extract related results, and this also what Quranic researcher can work on.

Thus, there are a lot of analytics methods for different uses considering researcher's purpose and the algorithm[2] determined. For that, it's significant that researcher should has a clear vision toward his topic, in order to use these methods perfectly.

---

1. Ignatow and Mihalcea, 2016.
2. Algorithm is a process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.

Practical Applications

We mentioned briefly in the previous section the most important concepts related to Data Science and Big Data, before presenting forthwith some applications samples that are helpful for Quranic researcher, by highlighting two open data sources: Digital newspapers and social data. Because, both contribute in manipulating the public trends, and engineering our realistic thoughts.

Regarding the sentiment analysis methods, some researchers apply it on New York Times' news headlines between the years 1990-2014 within their paper entitled "Are Muslims Collectively Responsible?",[1] to determine the rate of bias toward Islam and Muslims. They realized that 57 % of news headlines, which include Islam/Muslims keywords, are classified as negative, while 8% of them as positive.

Furthermore, their results indicate the lack of positive words through analyzing the popular context related, whilst conceptions as: "Rebels" and "Militant" are the top words that followed by Islam/Muslims keywords.

This study is an analytical sample in facing the media reality around the world. However, it's huge source to analyze the trends and their temporal related, as well their correlations in specific context, and predict also the future of streaming.

Further, the combination between multi sources of data is richness for research, because each time that news publishes, they will be expanded with tremendous speed in the social network in the form of feedback loop.

Another article entitled "Muslims in social media discourse: Combining topic modeling and critical discourse analysis" which relies on Social Data analysis,

1. Arshad, Setlur and Siddiqui, 2015.

depends on a similar research approach, by using Islam and Muslims keywords.

One of the top Islamic topics that this research mentioned is "how women are portrayed in the Quran, and whether gender oppression is an inherent trait of Islam. This also shows that the representations of Muslims and Islam are gendered".[1]

The researchers reached for a conclusion that the topics related to Islam and Muslims are increasing with time, but Muslims are taking defensive positions against discrimination and accusations, instead of dealing with deep intellectual topics related to the Muslim's needs, as in previous years, in which they dealt with geographical context in Sweden during the analysis of social data. It is appropriate to say here, that we can set the research limits, as place, time and language, and many more attributes.

Adding to analyze Social Data from a universal scale, we can also indentify specific known people who impact on the public streaming, and then investigate their followers' attitudes, by analyzing their common characteristics and trying to understand their personal circumstances surrounding any external influence.

As in the "Exmuslim" trend (exmuslim hashtag) which has grown widely since last year, that expresses atheists or agnostics concerns. In which thousands of posts have been published from the Islamic world, although the data showed that the campaign was from abroad. It should be noted that, the bigger percent of the publications through this hashtag were retweeted, and not for the original ones, which means that the republication is the bulk of this campaign.

1. Törnberg, 2016.

As we can analyze the Social Data concerned by this subject, and examine the opinions received, we can also effectively identify the most influentials who where in that campaign, and do an analytical study of their followers.

We designated this, to point out that the Islamic researchers must be aware of the theories of Social Network Analysis (SNA) to discover the impacts and how the interaction system within social networks works.

"The science of Social Network Analysis (SNA) boils down to one central concept—our relationships, taken together, define who we are and how we act. Our personality, education, background, race, ethnicity—all interact with our pattern of relationships and leave indelible marks on it. Thus, by observing and studying these patterns we can answer many questions about our sociality".[1]

Based on that, the researcher can indicates the nature of cultural interactions, and approach more to understand the reasons beyond any social phenomenon, before declaring his opinion or providing solutions and appropriate Quranic research.

Conclusion

Today, the religious institutions address several new responsibilities, including keeping up the development of Data Science, because, as we had known, it has powerful link with Quranic principles toward the knowledge approach. Which helps in diagnosing society needs, and enabling the discovery of new information, facts, relationships, indicators, and pointers that could not have been previously realized.

1. Tsvetovat and Kouznetsov, 2011.

Through the paper, we present the most important concepts related to Data Science that serve the Islamic researcher in general, and the religious educational institutions in particular, without explaining the relevant technical methods, for it is not its purpose.

Moreover, we pointed to the importance of Big Data analytics, in its Open Sources such as digital media websites and Social Data, and we emphasized the need to improve the traditional research methods, and integrating Data Science in the Quran science curriculum, so that it provides the Quranic researcher with accurate givens, in form of data visualization[1], that can be understood and analyzed.

We have included samples from recent researches, to confirm the importance of developing this trend, not only in general studies, but also specialized that's done by Quranic specialists, such as detecting the most popular Quranic misconceptions, or studying of the impact of social conditions on how to form special perceptions toward Quranic concepts and verses.

1. Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization software.

References

1. **The Holy Quran**.
2. Abu Alkassem Haj Hamad M., (2003), **Manhajiyat Al-Quraan Al-Ma'rifiya** [The Knowledgeable Approach of Quran], Beirut: Dar Al-Hadi.
3. Khoshmanesh, A., (2010), **Mahmoud Al-taliqani: Rajol Al-ihya'a Al-Qurani** [Mahmud Talqani: The Quranic revival], Beirut: Center of civilization for islamic thought development.
4. Mayer-Schönberger V., Cukier K., (2013), **Big Data: A Revolution That Will Transform How We Live, Work, and Think**, London: John Murray.
5. Qadri, M., Maree, M., Fauzan, Noordin M., (2015, september 9), Data Recording, Data Retrieval and the Big Data Phenomenon: A Quranic Perspective, *International Journal of Information and Communication Technology Research*. *http://esjournals.org/journaloftechnology/archive/vol5no9/vol5no9_3.pdf*
6. Ignatow G., Mihalcea R., (2016), **Text Mining: A Guidebook for the Social Sciences**, United states: SAGE Publications.
7. Arshad O., Setlur V., Siddiqui U., (2015), **Are Muslims Collectively Responsible? A Sentiment Analysis of the New York Times**, *http://416labs.com/nytandislam*
8. Törnberg A., Törnberg P., (2016), **Muslims in Social Media Discourse: Combining Topic Modeling and Critical Discourse Analysis**. *http://www.sciencedirect.com/science/article/pii/S2211695816300290*
9. Maksim Tsvetovat M., Kouznetsov A., (2011), **Social Network Analysis for Startups**, United States: O'Reilly Media, Inc.