Data and Information Flows: Assessing Threads and Opportunities to Ensure Privacy and Investment Returns

Alessio Faccia

Original citation & hyperlink:
Faccia, A., 2019, August. Data and Information Flows: Assessing Threads and Opportunities to Ensure Privacy and Investment Returns. In *Proceedings of the 2019 3rd International Conference on Cloud and Big Data Computing* (pp. 54-59).
https://doi.org/10.1145/3358505.3358508

DOI 10.1145/3358505.3358508
ISBN 978-145037165-0

Publisher: Association for Computing Machinery (ACM)

Data and information flows. Assessing threads and opportunities to ensure privacy and investment returns.

Dr. Alessio Faccia
American University In The Emirates
Academic City, Dubai,
P.O. Box 50300, UAE
+971 504256750
alessio.faccia@gmail.com

ABSTRACT

Timely access to data and information has always been the key to the success of every business. A merchant who can find goods, before anyone else, at a lower price, may succeed in obtaining a higher profit margin for the resale, the researcher who first manages to discover a chemical formula or to invent a new design, can sale a new product. Therefore, science must always face a terrible trade-off between the private need to make profits and the social opportunity of scientific development. Information and knowledge undoubtedly represent a valuable competitive advantage for those who hold them [1-2]. The data holders, in order to continue to benefit from this advantage, try in every way to protect their secrecy. Innovation is a creative process, generated by the human mind, driven by completely different potential factors. Some economic theories claim that innovation is driven by competition. Conversely, other more recent theories highlight the advantages produced by information sharing [3-6]. In this research paper, the author provides a systematic overview of the main factors that (a) influence decisions, (b) limit the development and dissemination of ideas [7], (c) guarantee economic growth. The analysis provided represents an effective tool for companies and policy makers to understand the standard patterns of the data and information flows and the key factors that drive the innovation, the development, and the growth of the world economy.

CCS Concepts

- **Security and privacy → Human and societal aspects of security and privacy.**
- **Social and professional topics → Computing / technology policy → Intellectual property.**
- **Information systems → Data Management Systems → Information integration.**
- **Information systems → World Wide Web → Web searching and information discovery.**

Keywords

Sharing information; big data; data management; data security; investment return.

1. INTRODUCTION

This research starts from the classification of the following four clusters: (a) the main programming languages; (b) the main big data analysis software; (c) the main users of big data; (d) the major big data providers. The link of these clusters allows the flow of information and data to be analysed sequentially and systematically, also considering the related historical periods of appearance (or development)[8] of the technologies or tools used in statistical and econometric studies.

The data confidentiality (privacy) [9] is a fundamental challenge in the analysis, both to protect personal data in general and to protect sensitive data from theft for the purpose of damaging a company's production system or copying information for someone else's interest. Many laws [10] in all the countries have been introduced on the topic. The correct and safe storage of data for their protection allows then to be able to quickly recover damaged data as soon as possible (backup and restore, automatic and manual backup) so as not to damage the user. The correct and safe storage of data for their protection allows then to be able to quickly recover damaged data as soon as possible (backup and restore, automatic and manual backup) so as not to damage the user. Data protection involves people inside and outside the company.

The lack of security systems [11-15], at various levels, can lead to economic and financial damage. For this reason, various IT security systems are applied, such as:

- data storage on different physical supports;
- data storage in protected and decentralized environments (not in a single point. Backups must be performed in any case);
- use of software that guarantees high data protection in order to allow sharing with different levels of access between different people. The data must be stored on a company server.
- data security is guaranteed by the authorizations granted at various levels by the system administrator who controls who accesses the data and which in turn is controlled by company security personnel;
- it must also be established who can have access to the premises where data is stored;
- a risk analysis must also be carried out and consequently the appropriate security measures must be taken (many software allow to do that);
- periodically update the personnel on the risks that occur for the disclosure of the company's confidential data, even if unintentionally.

Currently, there are many software that allow the shared management of data on servers and that guarantee a high level of confidentiality (SQL Server; Oracle etc). The choice depends on the cost of the licenses, on the maximum data storage space, on portability (different servers, clouds), and, above all, by various IT systems used by the company and which must be able to interface with each other. If the company works with Windows operating systems it is clear that it will be preferred in the choice of SQL SERVER, even if it is possible to access ORACLE environments or other operating systems as well. The software must also guarantee a certain flexibility in the search for information and accuracy of the data. The accuracy of the data also entails the

possibility of making forecasts and investments for the company without taking risks, or at least reducing them, of errors in choices that could damage the company. It is therefore necessary to identify technical and organizational solutions that allow reducing any high risks. The timing of implementation, the expected risk after the intervention and any costs must be defined. This analysis allows to make a decision in line with company policies and needs. Periodic safety level checks must be performed and every time the requirements are changed the risk analysis must be performed again. This makes it possible to improve the results of the evaluations obtained, to reduce and optimize management costs, to eliminate unnecessary and operationally damaging security measures (for example those that overlap or duplicate).

There is therefore an evident trade-off between opposing needs:

- companies to keep information secret (moreover costly obtained thanks to the use of private economic resources) to make an economic profit;

- the community to have access to information in order to increase and feed scientific research to produce innovations that we could all potentially benefit from.

In this research paper the author analyse therefore not only the technical aspects that regulate the management of data and the related information obtained, but also the social implications that are regulated by the laws on security and privacy. The wealth and prosperity of the community today, moreover, depends predominantly on the correct balancing of private and public interests. If a balance is not reached in relation to the protection of all the interests at stake, the entire economy of a country or the survival of businesses is endangered.

1.1 Programming Languages

The programming languages are linked to the different environments and operating systems in which they have been implemented. The company choice of the operating system to adopt therefore also affects the programming language to be preferred. Among all programming languages, the author provides below a useful list (although not exhaustive) of those mainly used for relational databases and for performing predictive analysis on big data.

Apache Groovy (2004)	Object oriented. Static and dynamic. Java Platform.	
C (1972)	General-purpose, supporting structured programming, static.	
C# (C Sharp) (2000)	general-purpose, modern and object-oriented programming language (Microsoft).	
C++ (1983)	general-purpose, imperative, object-oriented and generic programming features.	
CoffeeScript (2011)	programming language that transcompiles to JavaScript.	
Go (Golang) (2009) Google	statically typed, compiled programming language designed at Google.	
Haskell (1990)	statically typed, purely functional programming language.	
Java (1995)	general-purpose computer-programming language that is concurrent, class-based, object-oriented	
JavaScript (JS) (1996)	high-level, interpreted programming language, dynamic, weakly typed, prototype-based and multi-paradigm	
Lua (1993)	lightweight, multi-paradigm programming language designed primarily for embedded use in applications.	
MATLAB (1978)	multi-paradigm numerical computing environment developed by MathWorks.	
Objective-C (1982)	general-purpose, object-oriented programming language, supported by Apple for the macOS and iOS operating systems.	
Perl (1987)	family of two high-level, general-purpose, interpreted, dynamic programming languages, Perl 5 and Perl 6	
PHP (1994)	general-purpose programming language originally designed for web development.	
Python (1991)	interpreted, high-level, general-purpose programming language.	
R (1996)	programming language and free software environment for statistical computing and graphics	
Ruby (1993)	dynamic, interpreted, reflective, object-oriented, general-purpose programming language.	
Scala (2003)	general-purpose, providing support for functional programming, strong static type system.	
Solidity	object-oriented for writing smart contracts. It performs smart contracts on various blockchain platforms	
SQL (1978) IBM	domain-specific language used in programming and designed for managing data held in a relational database management system.	

Swift (2014) Apple	general-purpose, multi-paradigm, compiled programming language developed by Apple for iOS.	
TypeScript (2012) Microsoft	open-source, static, language developed and maintained by Microsoft. Strict syntactical superset of JavaScript.	
VB.NET (2001) Microsoft	multi-paradigm, object-oriented programming language launched by Microsoft.	
VBA Microsoft	implementation of Microsoft's event-driven programming language Visual Basic 6	

1.2 Big data analysis software

The analytical data, from descriptive to predictive, are fundamental for maintaining customers and for company growth. Information management tools allow companies to leverage the most valuable business resource: data. Data and information allow managers to get insights on customers, to protect the organization and to promote new revenue opportunities. It is only in 1999 that the word "Big Data" appeared. Even in this period the problem of having such a vast amount of stored data is very much felt, but with the impossibility of having enough elaborate software to be able to manage and analyse these quantities of data.

In order to highlight the advantages of the use of big data analysis software it appears first of all necessary to specify that there are 4 large types of data Analysis: (a) Descriptive Analytics, which is made up of all the tools that allow us to represent and describe in a graphical way the reality of certain situations or processes. In the case of companies, for example, it is related for example to the representation of business processes. The Descriptive Analytics allows the graphical display of performance levels; (b) Predictive Analytics, based on solutions that allow data analysis to be carried out in order to design future development scenarios. Predictive Analytics are based on mathematical models and techniques such as Predictive Models, Forecasting and others; (c) Prescriptive Analytics are related to tools that combine data analysis with the ability to take and manage decision-making processes. Prescriptive Analytics are tools that provide strategic indications or operational solutions based on both Descriptive Analysis and Predictive Analysis; and (d) Automated Analytics, that allow users to enter the automation field with Analytics solutions. Against the results of descriptive and predictive analyses, Automated Analytics allow to activate actions defined on the basis of rules. Rules that can in turn be the result of an analysis process, such as the study of the behaviours of a specific machine in relation to certain conditions being analysed.

Among all big data analysis software, the author provides below a useful list (although not exhaustive) of those mainly used for data storage, data analysis and for performing predictive analysis on big data. All the following software are based on some of the programming languages listed in the previous paragraph. Amazon Web Service (SQL), Google BigQuery (Golang, C#, Java, PHP, Python, Ruby), Arcadia Data (SQL), Microsoft Azure (SQL), GoodData (Golang), Actian Analytics Platform (SQL), Google Bigdata (Golang, SQL), Wavefront (SQL), IBM Big Data (SQL), Datameer (SQL), Attivio Active Intelligence Engine (SQL), DataTorrent (SQL), Cloudera Enterprise Bigdata (SQL), Opera Solutions Signal Hubs (SQL), FICO Big Data Analyzer (SQL), Palantir Bigdata (SQL), Oracle Bigdata Analytics (SQL), Qubole (SQL), Syncsort (SQL), Eviews (SQL), MapR Converged Data Platform (SQL), Hortonworks Data Platform (SQL), Amdocs Insight (SQL), Splunk Bugdata Analytics (SQL), Celebrus Technologies (SQL), VMware (SQL), HPCC Systems Big data (SQL), Pentaho Big Data Analytics (SQL), Pachyderm, BlueTalon (SQL), Flyxt, MongoDB (SQL), BigObject (SQL), Opera Solutions Signal Hub (SQL), Rubikloud (SQL), SAP Bigdata Analytics (SQL), Next Pathway (SQL), CSC Big Data Platform (SQL), Kognitio Analytical Platform (SQL), 1010data (SQL), GE Industrial Internet (SQL), DataStax Bigdata (SQL), SGI Bigdata (SQL), Teradata Bigdata Analytics (SQL), Intel Bigdata (SQL), Guavus, HP Bigdata (SQL), Dell Bigdata Analytics (SQL), Pivotal Bigdata (SQL), Cisco Bigdata (SQL), Mu Sigma Bigdata (SQL), MicroStrategy Bigdata (SQL), Opera Solutions Bigdata (SQL). It is undoubtedly evident that SQL [16-18] (see Figure no.1) is the most widespread programming language, used and preferred for interaction with database management systems, in particular with relational ones. Relational databases are based on unique data relationships.

Fig 1 – Most popular database programming languages

In practice, all the data to be processed are stored in fixed structures, called tables, where they are placed one after the other to form individual tuples or records (ie the rows of each table). Each data contained in a table is transformed into real information only if associated with its corresponding attribute, which represents the header of the column in which the data is stored. This structure requires the fragmentation of information between different tables, even when the data describes the same object. To identify the information belonging to the same object, a series of logical operations are used, such as the JOIN, which base its calculations on foreign keys or foreign keys that link one table to another. For this reason, using too large SQL tables is strongly discouraged: they are difficult to manage even with very powerful computational systems. This peculiarity requires SQL systems to carry out continuous queries, as the data must be collected and combined together from different tables, even for simple operations, such as insertion, deletion and updating. This strong fragmentation also requires strict control over the relationships and validity of data residing in the different tables in order to preserve the integrity of the database, to the detriment of flexibility. It is also for this reason that relational databases are based on tabular schemes (defined entity-relationship schemes) predefined upstream and that can hardly be readjusted to new situations, unless there is a risk of data corruption.

1.3 Big data main users

The main users of big data are undoubtedly:

- public institutions for financial and economic data to be able to perform econometric simulations of economic models and adopt the best economic and monetary policies;
- private companies, in order to analyze markets, consumption habits, commercial trends, consumer preferences, in order to best address their marketing and production strategies.

1.4 Financial big data vendors

The collection and analysis of big data allows companies to develop competitive and innovative business strategies and models that can constantly improve. In fact, studying this data allows companies to make targeted business decisions in order to increase the number of potential customers and therefore sales, to make customer loyalty strategies more efficient and to devise cross-selling and up-selling strategies studied in based on user needs. It is the technological skills of Analytics that transform the raw data received from the system into valid and interesting information for corporate decision makers: in a market in continuous development and change such as online, it is fundamental for companies of any type and size to take timely decisions. All over the world the Big Data market is now a consolidated business. The big data business exists and is evident. The data collected helps companies to understand how users perceive a brand and how they react to markets: by analyzing the data, the company can gather a large amount of information regarding the effectiveness of the communication strategies adopted and therefore inherent to the trend company itself. Big data make a company able to segment its user base in order to significantly customize the marketing and sales strategies implemented. Knowing and differentiating the purchasing behavior of users are essential actions for a company with a good online market. The data collected identify the key factors that move people to buy a specific product, allowing companies to leverage on the most profitable market choices. The study of Big data allows the effectiveness of individual media to be examined in order to identify which platforms are the most appropriate for communicating a given content. The information gathered allows users to design predictive simulations to understand how the market in which the reference company operates will evolve. Data analysis can enable new business models based on innovative purchasing trends recorded in real time. The speed and the exponential dynamics of the technological impact represent a demanding challenge for public institutions as well. The present and the future of digital transformation require adequate public systems capable of adapting and accompanying social and economic changes. Big data and algorithms are increasingly used also within public administrations. The risks are different, however: in addition to the error of the algorithm, the protection of privacy, the security of data, there is also that of discrimination and an "excess of surveillance". The usefulness of big data for international financial institutions is also undoubted in order to obtain more and more detailed economic forecasts and provide valid analytical tools to be able to present more incisive and effective economic and monetary policies.

The digital revolution continues to renew the landscape of traditional industries. Before spreading Internet access, news and financial information was transmitted through newspapers. In the US, The Wall Street Journal, The New York Times and the Financial Times were among the most popular sources of business news. But the demand for newspapers continues to decline as digital means of accessing information grow. Many companies have now turned away from the press and have only provided digital information. In addition to the vastness of digital news, access to financial information has grown. Bloomberg and Thomson Reuters are leading, supporting the majority market for commercial information. Despite their robust media platforms, both companies are known for technological devices: the Bloomberg Terminal and Thomson Reuters Eikon, respectively. For those who work at large financial institutions, the cost of both programs is probably negligible. However, for higher education institutions, government agencies and small businesses, the costs are amazing. Bloomberg Terminal is the most expensive financial data provider, costing $ 24,000 a year. For customers with two or more subscriptions, Bloomberg costs $ 20,000 a year. In comparison, a fully loaded version of Eikon costs $ 1 800 per month, with a simplified version that costs $ 300 a month. Other big data financial sellers are: Accern, AlphaSense, Cbonds, CQG, Dealogic, Eurekahedge, FactSet, Fidessa, FIS (SunGard), ICE Data Services, Interactive Data Corporation, IRESS, Markit, Mergermarket, Money.Net, Moody's Analytics, Morningstar, Inc., NYSE Technologies, Preqin, Quandl, S&P Capital IQ, SIX Financial Information, TraderMade, Vela Trading Systems.

2. LITERATURE REVIEW

Since the introduction of the "Human and societal aspects of security and privacy" concepts in the 2012 CCS classification (no equivalent concepts existed in the previous version), no recent papers are still displayed in the ACM Digital Library [19].

Since this research is mainly exploratory (as it will be clarified in the paragraph on methodology) and there are no recent studies for the systematic integration of all the topics analyzed in this paper, the author proceeds to briefly analyze the literature reviews related to all the subjects.

2.1 Data science

Data science is the set of methodological principles (based on the scientific method) and multidisciplinary techniques aimed at interpreting and extracting knowledge from data through the relative analysis phase by an expert (Data Scientist). The methods of data science (often associated with the concept of data mining) are based on techniques from various disciplines, mainly from mathematics, statistics, information science and computer science, especially in the following subdomains: databases and data visualization o business intelligence, artificial intelligence or machine learning (also known as *Big data*). The role of data scientist was defined by the Harvard Business Review as the sexiest profession of the twenty-first century [20] and is considered one of the four key roles in the systematic use of Big data in companies.

2.2 Sharing information age

Sharing knowledge and information is an activity through which knowledge (information, knowledge or experience) is exchanged between people, friends, families, communities (such as Wikipedia, or organizations) . Organizations have recognized that knowledge is a valuable intangible asset for creating and sustaining competitive advantages. Knowledge sharing activities are generally supported by knowledge management systems. However, technology is only one of many factors that influence the sharing of knowledge in organizations, such as organizational culture, trust and incentives. Knowledge sharing is a major challenge in the field of knowledge management because some employees tend to resist sharing their knowledge with the rest of the organization. [21-22] Knowledge and information are added values, intangible activities to build and sustain a competitive advantage within organizations. Several factors influence knowledge/information sharing in organizations, such as organizational culture, trust, incentives and technology. Knowledge sharing activities are commonly supported by knowledge management systems, a form of information technology (IT), which facilitates and organizes information within a company or organization. Information technology (IT) deals with tools that help facilitate knowledge sharing and knowledge management. The main role of IT systems is to help people share knowledge through common platforms and computer storage, to help make access easier by encouraging the economic re-use of knowledge. The computer systems are able to provide the coding,

the personalization, the electronic storage of information and can help people to communicate directly with others. With appropriate training and education, IT systems can make it easier for organizations to acquire, store or disseminate knowledge.

2.3 Privacy, copyright, and security regulations

Many problems have arisen in the management of privacy, in today's world that is strongly focused on the collection, processing and dissemination of large amounts of data (the so-called *Big data* phenomenon). At this level of criticality another one is added, independent but in any case superimposed: that of copyright. A large part of the documents and data, in fact, although often also in the public domain and outside the protection of copyright, are collected in large proprietary databases marketed (often at great cost) by publishers. The legislator finds two types of problems in the regulation of these two issues: (a) maintain a degree of generality in the definitions so as to be able to include specific new cases that follow the technological evolution; (b) to ensure and guarantee a fair balance between private and public interests.

3. METHODOLOGY AND QUESTIONS

The author used an exploratory qualitative research method, based on investigative techniques. The main objective is moreover the search for possible causes in a relatively unknown and not yet thoroughly investigated field. The author highlighted the need to obtain a detailed representation of a dependent variable: the flow of data and information. Therefore, this exploratory research expresses the need to formulate hypotheses on the possible causal factors and possible extraneous factors. The results obtained may represent an excellent basis for the formulation of more specific hypotheses, but they will have to be validated through further investigations, which include these variables in the causal model.

The questions that the author asked himself and to which he tried to find an answer, following the scientific method are the following:: do data protection and competition have a distorting and retarding effect on scientific development? What are the best incentives for research? Is it pure and general research neglected with respect to that applied to favour economic returns (profit)? Can these aspects be measured?

4. FINDINGS AND CONCLUSIONS

In the present research paper, the evolution and diffusion of programming languages and related software most used for big data analysis were analysed. Thanks to this analysis it was possible to verify that the mathematical and technological systems have focused and addressed mainly (almost exclusively) to SQL based IT platforms, as this language was the first and the most reliable developed. It should also be taken into account that Microsoft (fully compatible with SQL) and IBM (developer of the SQL language), also constitute the most commonly used work environments for businesses. However, this monopoly in the development of software that analyses data does not in any way limit the circulation and innovative development of data analysis techniques (mathematics and statistics), since the SQL programming language is highly flexible and customizable based on needs of researchers. The most controversial aspect of information management is instead represented by the conflicting interest shown by private companies and the community. There is indeed a clear trade-off between:

(a) the needs of private companies:

- to collect, purchase, or in any case have, in a timely manner, a large amount of data (often very personal) of consumers;

- to carry out research that leads to innovations whose knowledge must be kept secret and protected in order to allow the company to benefit from the competitive advantages deriving from the investment made with private funds;

(b) the social need of the community to obtain the sharing of information, technological innovations and discoveries in order to accelerate the process of technological and scientific development of the entire community itself.

The balancing of the same conflictual interests in the sphere of regulations on protection of privacy, competition and private economic development is therefore difficult. Several economic revolutions have occurred in modern and contemporary times: commercial, industrial, digital, information. It seems easy to observe that all these revolutions were caused, or at least facilitated, by similar discoveries and technological innovations. According to the author, the last revolution we are facing, that of information management (in particular of big data), can facilitate an incredible and even faster scientific development only with information sharing. However, it should not be overlooked that unregulated meritocracy, currently measured only by the ability to accumulate private wealth, is causing great inequalities that must somehow be resolved to avoid otherwise inevitable social wars. Further studies could focus on the analysis of the current regulations related to (a) intellectual property, (b) fairness of knowledge sharing; (c) competition. The biggest challenge is the need to avoid overlapping R&D that, as has happened so far, it has slowed down technological and social development. It is easy to observe how often many companies offer the same products, with some small incremental innovations. The same patented innovations can only be used by the companies that invented them. We imagine that a company A produces product A, which incorporates an incremental innovation Y (protected by an industrial patent), and company B, which produces the same product, which incorporates an incremental innovation Z (also protected by an industrial patent). In order to take advantage of the innovations, the consumer should buy the same product twice to benefit from innovation Y and Z. Providing a practical example, considering mobile phones, some manufacturers supply models with long-life batteries, but with obsolete software, other models producing models with advanced software, but with a low-capacity battery. A customer who intends to purchase a product with advanced software and high battery performance is currently not satisfied. The monopolistic situation, on the other hand, would have different defects, as probably, without the pressure of competition, the two manufacturers would not have developed those innovative and characteristic features of their product in order to survive in the market, at the same time economic theory shows how product prices practiced on the market by a monopolist would have been higher than the market in free competition. It has also been shown, however, how large companies (which occupy dominant positions in the market) have contributed to the wealth of nations [23]. The contribution of this research aims to better direct towards a solution of the dilemma between favouring a market concentration (which produces benefits in terms of greater resources potentially available for R&D and reduction of the overlap of research efforts, while generating higher prices for the consumers) or free competition (which stimulates innovation to ensure survival in the market, but in a context in which resources are lost and consumer satisfaction is not guaranteed).

5. REFERENCES

[1] Michael, K., & Miller, K. W., 2013. *Big data: New opportunities and new challenges* [guest editors' introduction]. Computer, 46(6), 22-24.

[2] Richards, N. M., & King, J. H., 2014. *Big data ethics*. Wake Forest L. Rev., 49, 393.

[3] Matofska, B., 2014. *What is the sharing economy*. The people who share, 444.

[4] Jorgenson, D. W., & Vu, K., 2005. *Information technology and the world economy*. Scandinavian Journal of Economics, 107(4), 631-650.

[5] Teubner, T., & Flath, C. M., 2019. *Privacy in the sharing economy*. Journal of the Association for Information Systems, 20(3), 213-242.

[6] Wyckoff, J. P., Asao, K., & Buss, D. M., 2019. *Gossip as an intrasexual competition strategy: Predicting information sharing from potential mate versus competitor mating strategies*. Evolution and Human Behavior, 40(1), 96-104.

[7] Allen, T. J., 1984. *Managing the flow of technology: Technology transfer and the dissemination of technological information within the R&D organization*. MIT Press Books, 1.

[8] Keklik, M., 2018. *Schumpeter, Innovation and Growth: Long-cycle dynamics in the post-WWII American manufacturing industries*. Routledge.

[9] Sangeetha, S., & Sadasivam, G. S., 2019. *Privacy of Big Data: A Review*. In Handbook of Big Data and IoT Security (pp. 5-23). Springer, Cham.

[10] Solove, D. J., 2008. *Understanding privacy* (Vol. 173). Cambridge, MA: Harvard university press.

[11] LINDSAY, Jeffrey Dean., 2009. *Security systems for protecting an asset*. U.S. Patent No 7,552,467.

[12] Bleha, S., Slivinsky, C., & Hussien, B., 1990. *Computer-access security systems using keystroke dynamics*. IEEE Transactions on pattern analysis and machine intelligence, 12(12), 1217-1222.

[13] Stallings, W., Brown, L., Bauer, M. D., & Bhattacharjee, A. K., 2012. *Computer security: principles and practice* (pp. 978-0). Upper Saddle River (NJ: Pearson Education.

[14] ROUNDY, Kevin, et al. 2019. *Systems and methods for providing integrated security management*. U.S. Patent Application No 10/242,187.

[15] Thomas, T. D., Quintus-Bosz, H., Menon, M. C., Parker, A. C., Renner, K. H., Aknin, J. P. M., ... & Stadille, J.. 2018. *U.S. Patent No. 9,979,427*. Washington, DC: U.S. Patent and Trademark Office.

[16] Kamara, S., & Moataz, T., 2018. *SQL on structurally-encrypted databases*. In International Conference on the Theory and Application of Cryptology and Information Security (pp. 149-180). Springer, Cham.

[17] Johnson, N., Near, J. P., & Song, D., 2018. *Towards practical differential privacy for SQL queries*. Proceedings of the VLDB Endowment, 11(5), 526-539.

[18] Batra, R., 2018. *A History of SQL and Relational Databases*. In SQL Primer (pp. 183-187). Apress, Berkeley, CA.

[19] https://dl.acm.org/ccs/ccs.cfm?id=10003032&lid=0.10002978.10003029.10003032

[20] Davenport T.H., Patil, D.J., 2012, *Data Scientist: The Sexiest Job of the 21st Century*, in Harvard Business Review.

[21] Bernhard Streitwieser & Bradley Beecher, 2017. *Information Sharing in the Age of Hyper-competition: Opening an International Branch Campus, Change*. The Magazine of Higher Learning, 49:6, 44-50.

[22] Schement, J. R., 2018. *Tendencies and tensions of the information age: Production and distribution of information in the United States*. Routledge

[23] Chandler, A. D., Amatori, F., & Hikino, T. (Eds.). (1999). Big business and the wealth of nations. Cambridge University Press.