Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 196 (2022) 982–989

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2021

Risks of Data Science Projects - A Delphi Study

Cristina Varela*, Luísa Domingues

Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal

Abstract

Risk is one of the most crucial components of a project. Its proper evaluation and treatment increase the chances of a project's success. This article presents the risks in Data Science projects, assessed through a study conducted with the Delphi technique, to answer the question, "What are the risks of Data Science projects". The study allowed the identification of specific risks related to data science projects, however it was possible to verify that over a half of the most mentioned risks are similar to other types of IT projects. This paper describes the research from expert selection, risk identification and analysis, and the first conclusions.

Peer-review under responsibility of the scientific committee of the CENTERIS –International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2021

Keywords: Data Science, project sucess, project risk management, risk assessment, Delphi study

1. Introdução

It is already a consensus that the data are natural resources of the new industrial revolution [1]. The need for using data to transform a business [2] has never been more urgent and thousands of companies have become overwhelmed with data, due to the amount of it and its potential to influence business growth. Data Science (DS) has become an indispensable ally for the development of society. The collection, analysis, and storage of data have allowed companies to understand and improve their business.

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
E-mail address: acava1@iscte-iul.pt

Peer-review under responsibility of the scientific committee of the CENTERIS –International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2021
10.1016/j.procs.2021.12.100

Though all this profit, 85% of Data Science projects fail [3]. Literature of this industry presents some common-sense type reasons: inadequate data [4] technology for its own sake instead of drivers of profits and key performance indicators (KPI) [5] poor communication, insufficient executive support [5] over-complicating, and limited focus on the real problem [6]. In effect, it is estimated that around 90% of available digital data is not being adequately used [1]. And the question every project manager wants the answer is: "how can my project be in the 15% of success rate?". Since there is a large number of events that can lead projects to fail, how can a company prevents and deals correctly with all of them?

The first thought or step in risk management is the identification of the risks to be controlled to assure the project's success. The identification of the risks factors also plays an important role in this process [7]. Throughout the years, several lists of risk factors have been published in the literature, but there is a lack of literature information about Data Science projects risks. This paper aims to report the result of a study conducted with Delphi procedures, to identify the risks of Data science projects and their drivers, in Portugal.

A Delphi approach seemed to be suitable for this investigation since the information would be consensually sourced by the Data Science specialists, who have been facing those risks in past and present experiences. This Delphi study aimed to answer the questions:

- What are the risks of Data Science projects in Portugal?
- How often do they appear and how it impacts the project?
- How normally the risks are assessed and managed?
- Which risks and risks factors are more common for each area of Data Science?

The remainder of this paper is organized as follows. First, it is described the background and motivation for the research. Second, it is presented the research methodology used to conduct the study. Third, it is presented the considerable findings of the study by reporting the list of Portuguese Data Science projects risks and discussing the results. Finally, it is presented the conclusions that can be taken from this research, and discussed the limitations of the study and the challenges faced, considering the potential for future work.

2. Background

The growth in the use of data science is driven by the emergence of Big Data and social media, the speedup in computing power, the massive reduction in the cost of computer memory, and the development of more powerful methods for data analysis and modeling, such as deep learning [8]. Manheim and Kaplan support the idea of technological improvement but also stated that the advances in the data science field brings not just a new age in computing, but also present new dangers to social values and constitutional rights [9]. Those risks are well known as a threat to privacy, social media algorithms, and the Internet of Things. Towards technological achievements society started to see a promised future in human well-being. But those improvements did not come without side effects. The internet gave rise to social media, whose devaluation of privacy has been profound [9]. Factors such as time, money, and unrealistic expectations, among many others, are able to sabotage a promising project if it is not properly managed [10]. In order to get better use of available data, taking into consideration the complexity involved in working with an extensive quantity of data, as well as the difficulties that organizations may face in managing and control them, managers should not underestimate the positive and negative effects that need to be taken into account in the exploitation of data [11]. So, what is Data Science?

Data Science encompasses a set of principles, problem definitions, algorithms, and processes for extracting non-obvious and useful patterns from large data sets [8]. It can also be described as the "art and science of acquiring knowledge through data" [12]. Data science is all about how the data is taken, processed, and used to acquire insights and knowledge [2]. This further helps the organization to make and improve decisions, predict the future, understand the past and/or present and create new business opportunities. Understanding Data Science begins with three basic areas: hacking skills, math and statistics knowledge, and substantive expertise [12].

2.1. Risks of Data Science Projects

Risk is described as an uncertain event or condition that, if it occurs, has a positive or negative effect on a project objective [13]. Data Science involves many different areas, from Big Data, Data Mining, Artificial Intelligence, Deep

Learning, Machine Learning, Data Analytics, Business Intelligence, and everything that involves extracting knowledge from data [2]. Data sets with many possible input variables bring essentially high risk [14]. As with any business initiative, a project of Big Data, Data Mining, or Data Analytics involves a lot of risks. At some point when the volume, variety, and velocity of the data are increased, the current techniques and technologies may not be able to handle storage and processing of the data [15] and it also brings issues regarding the cost of it and the security providence [11]. Risks as data privacy and confidentiality, poor analyses, lack of skilled expertise to deal with business and technologies [1], [16], have also been identified.

Hasan et al [17], stated that some organizations tend to adopt Business Intelligence (BI) with ambiguous objectives, allied with poor business and data management, limited funding and expertise, leading to a great failure. Artificial Intelligence (AI) has also encountered numerous advantages to the technological processes, but its risks can lead to major failures that can bring significant challenges for organizations. From reputational damage and revenue losses to the regulatory backlash, criminal investigation, and diminished public trust [18]. Risks factors as lack of labeled data, technology troubles, security snags, models misbehaving, human-machine interactions issues, lack of transparency and interpretability and so many others have been identified by many studies [19], [18], [20].

Considering all these risks, the following sections show the research methodology that has been used to investigate the risks of Data Science projects, their factors, frequency, and impacts in the Portuguese industry.

3. Research Method

To reach the information needed regarding Portuguese DS projects, it was conducted a Delphi study. The purpose of the Delphi study is to develop a reliable list of risks of Data Science projects in Portugal and its factors, and determine which of those risks are more important and impactful. This study also aims to get information about the response strategies and tools used to manage those threats and challenges with an acceptable level of consensus.

3.1. The panelist composition

The input source chosen for the study is Portuguese professionals with experience in Data Science projects. Since Data Science has diverse fields of practice and approach, a representative sample of different backgrounds was selected. According to Osborne et al. [21], commonly, the minimum number for a Delphi panel is 10 with a reduction in error and improved reliability with increased group size. [22], [23], [24], [25] revealed that studies tend to utilize between 10 and 30 specialist participants. The definition of the panelists/experts for the study was inspired by [25] steps proposed structure, as shown in table 1.

Table 1. Procedure for selecting panelists – Source: [25]

Steps	Actions
Step 1: Prepare KRNW	Identify relevant skills: practitioners
Step 2: Populate KRNW with names	Write in names of individuals in relevant skills
Step 3: Rank expert	Categorize experts according to appropriate list -> Rank experts within each list based on their qualification
Step 4: Invite experts	Invite experts in the order of their ranking to participate in the study -> target size 10 - 20

* KRNW - Knowledge Resource Nomination Worksheet

The specialist acknowledgment was provided, by LinkedIn searches. To achieve diversity in respondents, a list of specialists with different backgrounds in Data Science fields were chosen as the target populations for selecting the panelists. To find the wanted professionals it was applied the following search keywords were: “Data scientist” - “Data science project manager” – “Data science specialist/ expert “- “Big Data specialist/expert” – “Machine learning specialist” – “Data science Team manager” – “Data analyst specialist/ expert” – “Data mining specialist/ expert” – “Artificial Intelligence specialist/ expert” – “Data engineer”.

For each Data Science field, it was listed the largest number as possible into the appropriate categories. In the ranking step, it was compared the qualifications of those on the large list of specialists and rank them in priority for the invitation to the study. First, it was created the three sub-lists: The specialists, the seniors (people with a career in Data Science rounding 4 years and above, but not identified as a specialist), and the consultants (people with some years of experience in Data Science). The main focus of this study was the participation of the specialists and seniors based on their qualifications and years of experience. Based on the rankings, was sent 48 invitations for the study. Each person on the list was contacted and explained the subject of the study and the procedures required for it, including the commitment required, the schedule, and privacy terms. From the list, 16 accepted to take part in the study, 4 declined for personal/professional reasons and 28 did not respond to the invitation, which was reinforced twice. With 16 confirmed participants, it was created the panel for the first round.

The study was planned to be conducted in three stages to get a consensus. Brooks in [26] identified consensus as ‘a gathering of individual evaluations around a median response, with minimal divergence’. A general guideline is to conduct 3 rounds, before panel fatigue becomes an issue [25] and so that a strong consensus can be obtained, when Kendall's coefficient of concordance reach W>= 0.7 [22].

3.2. The Data collection

In the first round of this process, the 16 specialists/experts received a message on LinkedIn with the link to the questionnaire. In the introduction of the questionnaire, there was a text that oriented them to the purpose of the round and the completion of the questionnaire. The questionnaire was structured with a list of 41 risks of Data Science projects (grouped by category), identified during the literature review. The specialist role was to validate the risks that he/she has already identified in a Data Science project or that is considered relevant. It was also requested that panellists could present other risks by categorising them. The main goal of the first round was to undertake a brainstorming round to identify risk. The list was categorized following PMI [13] Risk Breakdown Structure. The risks categories selected were: Functional scope - Project management - Operation management – Technological – Quality - Resource - Surrounding environments – Legal – Communication – Organizational – others.

The second round involved classifying all the risks identified in the first round by probability, impact, outcome, and level of impact. The questionnaire was structured in an excel workbook comprising: an introduction sheet, a business area identification sheet, and a classification sheet, where the risks were listed.

For this round, it was sent more than 14 invitations (4 of them were the reinforcement of the first-round invitation), 6 specialists accepted the invitation, one declined and the 7 did not respond. From the first group (16 specialists), just 7 of them participated in the second round, and from the new set, just 4 participated. Therefore the second round had 11 participations. Based on 60 risks (41 risks from the first round plus the 19 suggested by the experts), organized by category and the areas in which the participant has professional experiences, the goal was to classify those risks by probability, impact, outcome, and level of impact. The probability and impact level options were: Very Low (0% – 10%) – Low (11% - 30%) – Medium (31% - 50%) – High (51% - 70%) - Very High (71% - 99%). The impact type options were: Integrity / Credibility - Availability (service, infrastructure, human...) – Security (public, infrastructure, business...) - Quality of Service – Cost - Productivity / Motivation - Legal process. Finally, the outcome options were: Opportunity or Threat.

3.3. Findings

In the first round, within the 41 listed risks and 16 participations, some risks got a remarkable identification rate compared to the others. About 69% of the experts identified risks such as “starting the project with wrong questions”, “project scope poorly outlined with the client”, and “lack of concern about data quality”. 62% identified risks such as “misinterpretation of requirements”, “lack of information in the data”, “poor analysis/validation of model inputs/outputs”, “lack of talent (lack of knowledge and training in data intelligence)”, “violation of data privacy”, “communication problems”, and “lack of transparency”. Such risks are mostly of a functional and project management nature. Other risks, related to “poor time estimation”, “lack of specialized staff in the team”, “use of wrong methods in data collection and processing”, “model instability” and “loss of data confidentiality” were also highly identified, reaching an identification rate of 50% among the experts. While behaviors related to 2negligence towards data security" were often identified (44% identification rate), risks related to "lack of security awareness", "cyber-attacks" and "theft or loss of information" were rarely pointed out, reaching only a 10% identification rate. Implying that companies and employees are aware of security measures but they are not always applied, and this neglect of security may explain the high identification rate of data privacy breaches and loss of confidentiality of data (almost 70%). Risks related to political, economic and social issues were rarely pointed out (about 18% of the experts identified them).

However, the dynamics defined for this first round also allowed the experts to suggest factors and risks considered relevant and that were not presented in the list initially presented, thus conditioning the brainstorming. Within the structured categories, there were suggestions in almost all them. They are:

- Functional scope risks ("poor management of customer expectations" and "unclear or lack of documentation (technical and functional)"),
- Project management risks ("lack of clarity about which information sources to use", "poor management of the team skills and high team volatility"),
- Operational management risks ("concentrating on model performance rather than usability" and "lack of knowledge of data segmentation practices")
- Quality risks ("limited data processor capacity (memory, processor)", "lack of a quality management system on input data (lack of validation on input formatting)" and "poor quality testing by of client")
- Communication risks ("customer unavailability to meet", "lack of alignment between project manager, team and client" and "lack of model explanation for a non-technical language")
- Technological risks ("lack of clarity about the expected growth/Upgrade of the source systems", "poor software choice/selection", "systems incompatibility (migration problems)" and "introduction with a new technology")
- Organizational risks ("lack of information system structure support").

The main goal of the first round was to develop a list of factors that could drive a range of diverse and different risks. Undertaken this list and 11 participations, the second round results analysis, revealed that most of the risks were classified as threats. The most identified risk categories were "Project functional scope", "Communication" and "Project Management", and the most identified areas of impact were on "Project scope", "Cost", "Time" and "Service quality". Considering that the fields were different it was expected that some risks would occur more frequently in some fields than in others. Based on this assumption, it was created the top 25 most frequent risks in every single DS field taken in analysis. Since some risks were in a similar context the list was reconfigured to better describe each risk and the scenarios it could occur. It was considered to be frequent, those risks with the probability level classification of "Medium" to "Very high". Table 2, represents the 25 most identified risks, considered frequent in DS projects.

Table 2. Top 25 most identified risks as "frequent" in the second round

Risks	Description
Project scope poorly defined	It refers to scenarios where the project is started with the wrong questions/focus. Often, people focus more on how interesting a project is and not on how much money/time it will save the company. This can sometimes cause the project questions and answers to be wrongly planned
Lack of documentation	It relates to scenarios where there is a scarcity of (technical / functional) documentation and its lack of clarity, when it exists. For it is necessary that there is an explanation of how the models work and how the conclusion was reached
Lack of transparency and understanding of the project (communication problems)	Refers to scenarios in which there are communication problems. For example, lack of alignment of ideas between the project manager, team and client, lack of interest or critical sense on project issues, misinterpretation of requirements or refusal to ask for help in time
Little analysis/validation of model inputs and outputs	Refers to scenarios where there is a lack of verification of the quality and format of data (input and output) by the customer or the team responsible for the project
Negligence in risk planning	Refers to the lack of emphasis given to planning and identifying possible risks to the project
Poorly defined estimates	Refers to underestimated deadlines that are difficult to meet
Data system failures	Refers to scenarios related to migration problems, system incompatibility or data growth not assumed by the target system

Complex data - faulty data	Relates to scenarios where there is little margin for manipulation, complex data label
Lack of awareness about future changes	Refers to the scenarios in which project planning and development is not carried out aiming at the constant changes and needs of the market
Poor skills management in the team	Refers to scenarios in which team members are responsible for tasks outside their area of expertise, or when the inputs of some members are not taken into consideration
Budget underestimation	Refers to the scenarios in which the budget planned for the project, is not able to meet the project's requirements
Inability to strengthen budget	Refers to the impossibility of increasing the project's financial resources
Negligence towards data security	It relates to the lack of training and awareness of employees about data security and digital vulnerability security and digital vulnerability
Human operational errors in project development	Refers to errors made by the project team during the development of the project (e.g: use of improper methods of data processing and analysis)
Conflict of objectives among stakeholders	Refers to the scenario where project stakeholders have diverging opinions and goals
Very complex models	It is related to the scenarios where there is difficulty in the usability of the models and lack of explanation of them in a non-technical language. Or when it becomes difficult for whoever continues with the maintenance to maintain the algorithm
Complex projects	Refers to projects with a high rate of complexity and demand
Theft/leakage of information	Refers to unauthorized access to data
Poor team management	Refers to scenarios that includes, high member volatility and the lack of adequate project specialists (scientists with limited business knowledge)
Poor management of customer expectations	It refers to cases in which the client is not made aware of the limitations of the work to be developed. By carelessness or lack of experience of the manager/leader
Instability/degradation of the model or software (algorithm failure)	Refers to scenarios where models are made available without having a solid stability or scenarios where the models stop working well after some time
Cyber attacks	It relates to unauthorized access under a computer system, with the aim of altering, disabling, destroying or stealing information
Lack of knowledge of data segmentation laws	Refers to scenarios where there is little knowledge about best practices for segmenting the data in use
Unavailability of the customer to meet	It relates to the scenario where face-to-face contact with the client to discuss important points of the project becomes very scarce

Although this study is not yet complete, the result obtained in this second round shows that there are many risk situations in the project scope management, which in general end up influencing the appearance of other risks. 64% of the experts consider "poorly defined project scope" as frequent to very frequent, influencing delays, additional costs and the overall result. The project scope defines all the paths and resources needed to achieve project success, and if it is not critically and clearly defined, the probability of project failure is very high. Some experts have pointed out that this risk is very serious because it tends to trigger many negative events in the course of the developments and have a direct impact on the success of the project and the customer's actual business needs. Even if the project is delivered, it could not be useful to the customer. "Communication problems" was rated as frequent to very frequent 89% of the time and has been pointed out as a risk that impacts cost, time and scope, the fundamental pillars of a project [13]. As for the “lack of documents”, considered frequent 46% of the time, is considered a trigger to delays. Still, in project management, other risks identified as frequently to very frequent, are the ones related to “poor time estimation” (73%) and “poor management of customer expectations” (55%), but not always considered to have a high impact. Some experts consider that it depends a lot on the type of project, the client and the project scope. Other quality and performance risks as “lack data labelling - faulty data”, “focus only on model performance instead of usability” were also pointed as frequent 64% and 73% of the time, respectively. The lack of data quality assessed was pointed to be a frequent problem, since it can change the defined scope, delay the project and contribute to wrong decision making. Risks, such as “data security negligence” were also mentioned 64% of the time, as frequent. Some of the experts reported that in many projects they participated in, the issue of security at all levels is not yet practiced. Small actions that could create openings for information leaks that could compromise the organization or the customer are quite frequent.

Although it is not yet possible to present conclusive statistical data as of this study, it can be seen that there is some consensus about the frequency of the 25 risks presented in table 2. And it can be seen that the type of Data Science project contributed to the severity/impact level not always being the same, but in general, the risk impacts the same business areas (time and cost), regardless of its severity. Only a first analysis of the results of this study has been presented here. The additional results from the third round will be the information needed to conclude this study using Delphi techniques.

4. Conclusion

For this study, the use of the Delphi technique was very valuable. It was possible to gather information, analyze it to be later discussed with experts to create a consensus among all risks identified. It represented a major contribution to find the leading risks that could influence the success or failure of a project. So far, this study has provided an understanding of the most critical risk issues among projects of Data Science in Portugal. In addition to the rounds of the study already completed, a final round will be conducted, with the goal of obtaining a conclusive classification of the 25 most frequent risks identified. In this last round, it is also expected to have more direct contact with experts, who in addition to classifying the risks, may also share experiences about Data Science projects in Portugal. The risks, the biggest challenges, success, and failure rates of the projects, and the factors that led to such closure. For that, the third round will be conducted throughout meeting/interview sessions with the experts. This last round of the study will allow the experts to reflect on the collective opinion regarding what each risk represents in the Data Science field and also raise awareness about these issues. Consequently, we consider the Delphi technique to be a very powerful research tool. We believe that this study was a very valuable starting point for addressing the risks in Data Science projects in Portugal and that there is still much more research and information to be gathered about the success of Data Science projects and how organizations can design projects targeting the risks identified in this study in order to achieve better results.

Acknowledgements

We thank all the experts, who kindly accepted the invitation to be part of this study. A special thank also goes to ISCTE -IUL, for the research funding supported by the FCT UIDB/04466/2020 project.

References

[1] C. Taurion, Big Data, Rio de Janeiro: Brasport Livros e Multimídia Ltda, 2013.

[2] E. Robinson e J. Nolis, Build a Career in Data Science, New York: Manning Publictions Co, 2020.

[3] J. Walker, "Digital Journal - Big data strategies disappoint with 85 percent failure rate," 23 November 2017. [Online]. Available: http://www.digitaljournal.com/tech-and-science/technology/big-data-strategies-disappoint-with-85-percent-failurerate/article/508325#ixzz5JvGk3AV0. [Accessed on December 10th, 2020].

[4] M. Asay, "3 ways to massively fail with machine learning (and one key to success)," 12 July 2017. [Online]. Available: https://www.techrepublic.com/article/3-ways-to-massively-fail-with-machine-learning-and-one-key-to-success/.[Accessed on October 23rd, 2020].

[5] B. Taylor, "Why Most AI Projects Fail," 16 October 2017. [Online]. Available: https://www.datarobot.com/blog/why-most-ai-projects-fail./. [Accessed on October 23rd, 2020].

[6] K. Veeramachaneni, "Why You're Not Getting Value from Your Data Science," 7 December 2016. [Online]. Available: https://hbr.org/2016/12/why-youre-not-getting-value-from-your-data-science. [Accessed on October 24th, 2020].

[7] H. Barki, S. Rivard e J. Talbo, "Toward an Assessment of Software Development Risk," Journal of Management Information System, pp. 203-225, 1993.

[8] J. D. Kelleher e B. Tierney, Data science, Cambridge, Massachusetts: MIT Press, 2018.

[9] K. Manheim e L. Kaplan, "Artificial Intelligence: Risks to Privacy and Democracy," 21 Yale J.L. & Tech., vol. 21, 2019.

[10] K. Schwalbe, Information Technology Project Management, Eighth Edition, Boston: Cengage Learning, 2016.

[11] E. Raguseo, "Big data technologies: An empirical investigation on their adoption, benefits and risks for companies," International Journal of Information Management, pp. 187-195, 2018.
[12] S. Ozdemir, Principles of data science, Birmingham - Mumbai: Packt Publishing, 2016.

[13] PMI, A Guide to the Project Management Body of Knowledge (PMBOK® Guide) Sixth Edition, Pensilvânia: Project Management Institute, Inc., 2017.
[14] R. d. Graaf, Managing Your Data Science Projects: Learn Salesmanship, Presentation, and Maintenance of Completed Models 1st ed. Edition, Australia: Apress, 2019.
[15] D. K. Jadhav, "Big Data: The New Challenges in Data Mining," International Journal of Innovative Research in Computer Science & Technology, pp. 2347-5552, 2013.
[16] B. Gärtner e M. R. Hiebl, "Issues with Big Data," The Routledge Companion to Accounting Information Systems, 2018.
[17] N. A. Hasan, A. A. Rahman e N. A. Lahad, "ISSUES AND CHALLENGES IN BUSINESS INTELLIGENCE CASE STUDIES," Jurnal Teknologi, pp. 171-178, 2015.
[18] B. Cheatham, K. Javanmardian e H. Samandari, "Confronting the risks of artificial intelligence," McKinsey Quarterly,, pp. 1-9, 2019.
[19] H. R. Tizhoosh e L. Pantanowitz, "Artificial intelligence and digital pathology: Challenges and opportunities," Journal of pathology Informatics, pp. 1-6, 2018.
[20]D. H. Freedman, "A Reality Check for IBM's AI Ambitions," 27 June 2017. [Online]. Available: https://www.technologyreview.com/2017/06/27/4462/a-reality-check-for-ibms-ai-ambitions/.
[21] J. Osborne, S. Collins, M. Ratcliffe, R. Millar e R. Duschl, "What "Ideas-about-Science" should be taught in school science? A Delphi study of the expert community," Journal of Research in science teaching, pp. 692-720, 2002.
[22] R. Schmidt, K. Lyytinen, M. Keil e P. Cule, "Identifying Software Project Risks An International Delphi Study," Journal of Management Information System, pp. 5-36, 2001.
[23] V. Kasi, M. Keil, L. Mathiassen e K. Pedersen, "The post mortem paradox: a Delphi study of IT specialist perceptions," European Journal of Information Systems, pp. 67-78, 2008.
[24] S. D. Haes e W. V. Grembergen, "An exploratory study into IT governance implementations and its impact on business/IT alignment," Information Systems Management, pp. 123-137, 2009.
[25] C. Okoli e S. D. Pawlowski, "The Delphi method as a research tool: an example, design considerations and applications," Information & Management, pp. 15-29, 2004.
[26] K. W. Brooks, "Delphi techniques: Expanding applications," North Central Association Quarterly, 1979.
[27] L. Frei, "Why The Data Science Venn Diagram Is Misleading," Towards data science, 27 March 2019. [Online]. Available: https://towardsdatascience.com/why-the-data-science-venn-diagram-is-misleading-16751f852063. [Accessed on October 18th, 2020].