Salminen, J., Milenković, M. & Jansen, B.J. (2017). Problems of data science in organizations: An explorative qualitative analysis of business professionals' concerns. In *Proceedings of The 17th International Conference on Electronic Business* (pp. 192-201). ICEB, Dubai, UAE, December 4-8.



# Problems of Data Science in Organizations: An Explorative Qualitative Analysis of Business Professionals' Concerns

Joni Salminen*, Qatar Computing Research Institute, HBKU; and Turku School of Economics, State of Qatar, jsalminen@hbku.edu.qa
Milica Milencović, micagmekic@gmail.com
Bernard J. Jansen, Qatar Computing Research Institute, HBKU, State of Qatar, bjansen@hbku.edu.qa

## ABSTRACT

In this exploratory study, we analyze 150 comments from 79 participants, the roles ranging from top management and other business professionals to software developers, to identify key problems of employing data science in organizations. The comments are retrieved from a publicly available LinkedIn discussion thread in which the participants are discussing the problems relating to data science implementation and management. We use qualitative coding to analyze the comments and find issues from several management-related categories, including (a) job descriptions and recruitment, (b) leadership, (c) economical aspects, and (d) clarity about data use and goals. The findings also highlight that 'data scientist' is not just a one role, but combination of many different roles, including analyst, scientist, programmer, and business person. The multiplicity of skills required hinders the recruitment of such individuals, and the existing organizational structures are not always compatible with the multidisciplinary nature of data scientists. We conclude with recommendations to address these issues.

*Keywords*: Data science, analytics, problems, management

*Corresponding author

## INTRODUCTION

There is tremendous interest in data science as organizations seek to gain value from the increased ability to gather consumer information. Yet, many organizations struggle to implement it in practice. Often, the problems are not related to knowledge and expertise of the data scientists, but to organizational issues, which hinder the full exploitation and use of data science in achieving business goals. Understanding about these problems, however, is limited in the current research literature. By leveraging the comments of a highly relevant LinkedIn discussion thread, we aim to provide useful insights into the nature of the organizational problems relating to data science. In particular, we ask: What are the key themes, according to business professionals and other stakeholders, hindering the full use of data science in organizations? We answer this question through a qualitative analysis of 150 social media comments from 71 participants.

## RELATED LITERATURE

Overall, the topic of organizational issues and data science or analytics has not been explored in detail. In particular, how organizations use analytics in practice, is an open research question (Acito & Khatri, 2014; Carlsson, 2017). While many organizations have adopted analytics, strategic planning hinders behind the operational issues, risking the profitability of analytics investments (Hong, 2007). In particular, three sub-themes in prior literature on organizational issues with analytics can be identified, those being 1) issues with analytics adoption, 2) transforming data into insights, and 3) management of analytics (Järvinen, 2016). Here, we use "analytics" here as a synonym term for "data science," because analytics is a more widely used concept but thematically analytics and data science, defined as the use of computational techniques and big data to solve real-world analytical problems, are highly similar. Therefore, it makes sense to position our study in the analytics stream of literature, instead of focusing on the highly scarce literature that particularly deals with data science and data scientists. Data scientist, originating as a practitioner-created concept and job role (Davenport & Patil, 2012), is more rarely used in academic literature than the term "analytics" that has years of research tradition with sub-fields such as Web analytics and customer analytics. In the following sections, we briefly summarize the three abovementioned sub-themes of organizational issues with analytics.

### Issues with Analytics Adoption

There are potential adoption issues related to analytics systems. First, analytics systems, such as balanced scorecard measurement, can bring with it change resistance, especially when it is revealing under-performance by individuals that would rather leave performance details unexplored (Neely *et al*., 2000). Lapointe and Rivard (2005) identify this factor, prohibitive to information system adoption, as threatening consequences of the adoption and use of a system. Carlsson (2017) argues that digitalization does not only bring potential advantages, but that the proliferation of data and analytics also increases competition, reduces margins for productivity and profitability, and adds pressure for effective planning and decision making. An example of such adverse effects is digital analytics driving down marketing budgets in overall media spend (Honkaniemi, 2015), which has possibly curbed the traditional publishers' willingness to adopt analytics as part of their sales processes. Lapointe and Rivard (2005) also observe power-related issues between staff and system advocates which seems to suggest that lack of user participation is prohibitive to system adoption. Neely and Bourne (2000) distinguish between political and infrastructural reasons for failure of adopting measurement systems, which seems to suggest a division between human and technological issues. Two major barriers that Järvinen (2016) mentions include lack of skills, and lack of time – since analytics adoption may involve unpredictability and socio-technical unknowns, getting analytics investments right poses some learning curve.

**Transforming Data into Insights**

Moreover, after adoption, the main problem of analytics systems seems to relate to value extraction. Acito and Khatri (2014, p. 565) note that "At its core, business analytics is about leveraging value from data." Despite its enormous potential, extracting valuable insights from data represents a persistent and continuous practical problem for organizations (Carlsson, 2017). Järvinen (2016, p. 60) argues that "all data gathered is useless if not understood and refined into meaningful conclusions that are capable of driving future actions." The point of actionability of also stressed by Edwards and Taborda (2016) who postulate that unless there are new or different actions by some stakeholder, the analytical insight has been in vain. As analytics adoption represents a considerable commitment in terms of resources, its ineffective application represents a considerable risk to the organization (Edwards & Taborda, 2016). Järvinen (2016) mentions the metrics selection problem, so that organizations struggle to select a collection of metrics that is both comprehensive and manageable; choosing the most suitable metrics among the hundreds of metrics that are typically available via online platforms, and then prioritizing their use for decision making, represents a true challenge. As such, transformation of data into insights can be seen as one of the great problems in analytics.

**Management of Data Scientists**

Relating to both of the above issues, adoption and value extraction, the management of data scientists shows itself as important. As Vidgen *et al*. (2017, p. 626) note, organizations need to have "the right people to effect a data-driven cultural change." In addition to the management issues pertaining to organizational culture being highly multi-dimensional and complicated (Barney, 1986; Denison, 1990; Schein, 1990), there is very little research on the context of data science and analytics. Most literature can be found in the cross-section of practice and academia. For example, in their Harvard Business Review article, Davenport and Patil (2012, p. 70) summarize their practical experience of managing data scientists: "Data scientists don't do well on a short leash." In particular, Davenport and Patil (2012) advocate providing the data scientists with room to experiment with the data and systems.

The disconnect with this investigative work, crucial for coming up with innovative solutions, and practical applicability to solving business problems is akin to the dilemma between basic and applied research – without basic research, ground-breaking discoveries are left unmade, and without applied research their full value is not appropriated (cf. Nelson, 1959). As a solution, Davenport and Patil (2012) suggest not only close connection with the data, but with the organization's customer-facing products and processes. Another form of organizational disconnect between the data scientists and other departments is that, despite considerable lip service to the virtue of being "data-driven," the reality is that such a state of analytics maturity is extremely difficult to infuse throughout the organizations, including the traditionally creative-driven departments, such as marketing (Järvinen, 2016). This creativity-drive versus data-drive dilemma represents a management issue in the sense that the two alternative paradigms, intuition and personal experience based on one hand, and data-driven on the other hand, are incommensurable drivers for decision making, as well as competing over internal resources in the organization. The issue is also that data scientists are not isolated but embedded in the organizational context, dependent on what Järvinen (2016) refers to as measurement resources, which include analytical skills, IT infrastructure, commitment of the executive branch, strategy and leadership, and the "soft" organizational culture. Evidently, management of data scientists represents a complex organizational issue, of which there is not yet enough research to form a complete understanding, partly due to the novelty of this role in the organizational reality (Davenport & Patil, 2012).

Even though broad and general characterizations of data science problems have been presented, as illustrated above, a detailed and granular consideration of these issues, especially from a management perspective, is scarce. We aim at addressing this research gap by a detailed classification of the problems, shedding light to the management aspect of data scientists, as well as discussing the nature of their role in the organization.

## METHOD

**Research Question**

The primary research question of this work is: What are the key themes, according to business professionals and other stakeholders of data science, that are hindering the full use and exploitation of data science in organizations? Moreover, we ask: How do professionals from different fields of practice (e.g., managers, engineers, entrepreneurs) perceive these challenges? The following sections will describe our data collection technique, data, and the analytical method applied.

**Data Collection**

The analysis focuses on 150 comments in a publicly available discussion thread[1] on LinkedIn, the popular business-oriented social network. The post was made by Andi Shehu, CEO of Byteflow Dynamics, a company which provides intelligent data-driven solutions to organizations. The original post, stressing the issue of employing data science in organizations, can be seen in Figure 1.

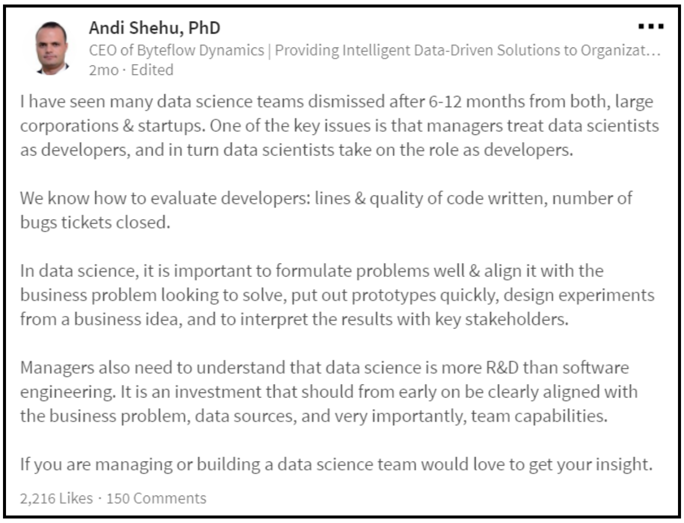


Figure 1: The original post on LinkedIn.

As of the date of data collection (September 4th, 2017), the post had gathered 2,216 likes and 150 comments. Clearly, it is a "viral post" among the LinkedIn business audience, the numbers indicating an unconventionally high diffusion (typically, LinkedIn posts tend to accumulate a few likes). Since the comments are relevant for the topic and originate from business professionals who have first-hand experience in the matter at hand, they represent a good opportunity for a systematic qualitative analysis.

**Description of Data**

We manually extracted all the comments answering to the original post. We then discarded those that did not include useful insight about the topic. By 'useful insight,' we mean that the comment has something original to say, e.g. "based on my experience, the major problems include 1) knowledge and 2) skills of the managers." An example of non-insightful comment is, e.g., "I agree, you are right" – as such comments do not add any insight, they were excluded from the analysis. This left us with 79 comments (~53% of all comments) from 71 participants, 67 (94%) being male and 4 (6%) female.

The participants were classified by visiting their public LinkedIn profiles, deriving their currently displayed job positions. This showed that the participants form a diverse pool of business professionals, including entrepreneurs (14), Top Management (16), Middle Management (12), Lower Management (3), Data Scientists/Analysts (19), and Software Developers and Consultants. Figure 2 shows a breakdown of the participants by role.

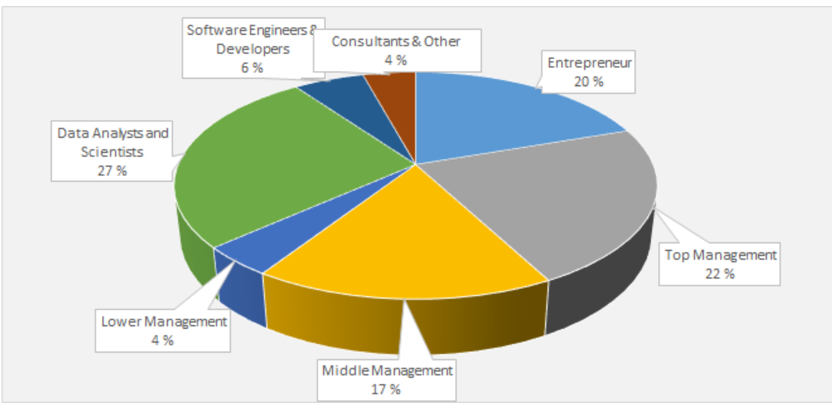


Figure 2: Participants by role. We grouped the 71 job roles into seven representative categories.

[1] https://www.linkedin.com/feed/update/urn:li:activity:6295317004786110464/

**Analysis Method**

After transferring each relevant comment to Microsoft Excel, we analyzed all comments to identify common themes. These themes were in another column of the sheet. For example, the theme of the previously mentioned example would be "Knowledge and skills," representing the meaning of the comment. This approach, referred to as open coding (see Glaser & Strauss, 1967; Strauss & Corbin, 1998), is a commonly used technique in qualitative analysis. The approach includes having no specific categories in mind, rather letting them emerge from the material, as emphasized by Glaser and Strauss (1967) in their original approach. After the open coding, we discussed the found themes, and decided to make some fine-tuning, including merging and addition of sub-themes. The resulting theme framework can be seen in the findings. It includes a combination of main themes and sub-themes. To make possible more detailed analyses, we tied the coded comments to each author, so that comparisons by different criteria, including job positions and gender, becomes possible.

## FINDINGS

**Themes Identified from Qualitative Analysis**

We made several interesting findings. Table 1 summarizes the themes emerging from the qualitative coding.

Table 1: Themes emerging from the qualitative analysis

| Theme | Explanation |
|---|---|
| Knowledge and Skills | Anything about skills (including comments about lack of skills) that data scientists, and data analysts need to have and develop. Also, the value, and impact of the knowledge. |
| Multidisciplinarity | The connection with other technical sciences, also with some social sciences, when it comes to interpreting and representing the data. Using the methods of other disciplines in order to conduct analysis, make assessments and conclusions. |
| Market Trends | Companies' actions are dictated by new trends, like AI/ML, and so on. Thus, the behavior needs to be adapted to the market requirements. |
| Turning to Programming | Some data scientists also think of themselves as programmers, both with a good reason (having skills), or without, so they apply and do the jobs that do not include data science only. |
| Software Development | Stating skills, needs, and activities of software developers, and talking about software development in general. |
| Seasonal Job | Data scientists are not constantly needed in companies, so their work is not necessary during the whole year. |
| Suggesting the sector for implementation | Suggesting in which industries and lines of businesses data science as a discipline should be implemented and applied. |
| Job Title | Statements relating to job title. |
| Management Approach | The view and opinion the leadership has towards the position of data scientists in the organizational structure of the company, and the way the leadership conducts the process of achieving goals and bringing the company closer to fulfilling those goals. In that process, the position of data scientists can be compromised due to imbalance between management requests and expectations, and competences data scientists should have, along with other external factors that can influence decisions of the top management. |

Each theme represents a specific way of determining what needs to be done in order to make data scientists fit better into the corporate context. Figure 3 summarizes the prevalence of the identified concerns among the participants.

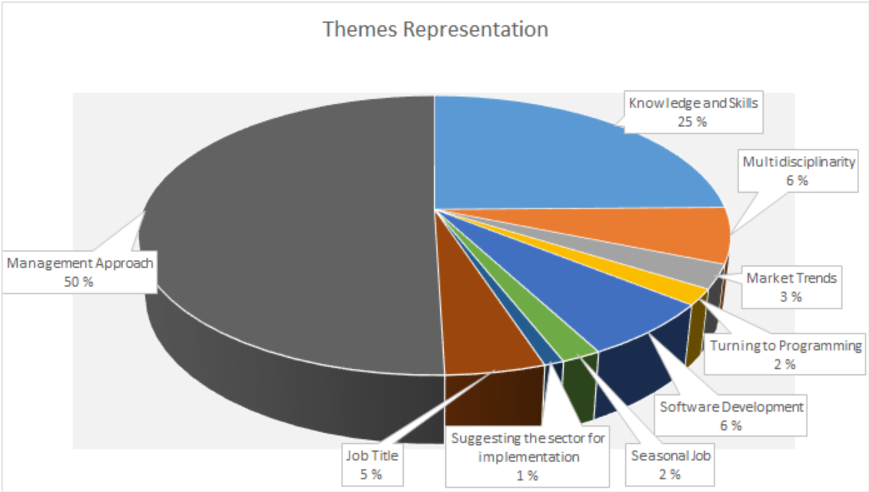


Figure 3: Distribution of data science concerns. Percentage is calculated by dividing the number annotations in a given category of the number of all annotations.

Themes that were mentioned the most are Management Approach, and Knowledge and Skills, where 77% of comments contained ether complaints of praises (not many), and suggestions about leadership, job descriptions (41%), leadership (37%), clear goal setting (15%), economical side of the decision making (10%). The following sections summarize the findings from each theme.

***Knowledge and Skills***

The comments in this category point out the set of skills that data scientists and analysts need to acquire to succeed and stay relevant in their field.

> *They should have software development skills but just high level. Why? Bcz, they should conceptualise the implementation part, while using their algorithms and functions. Some data scientists methods are just theoretical. Its implementation becomes impossible or poor outcome than expected.* (Participant 9)

Crucial element for fixing this problem is, according to the participants, education. However, because of a tremendous amount of learning material, some key subjects may become overlooked, due to the freedom of choice when it comes to picking courses in data science and analytics field.

***Multidisciplinarity***

Multidisciplinarity is closely related to previous theme, because it requires insights in terms of knowledge from other sciences.

> *I think another issue here is that machine learning and data science at a startup is multi-disciplinary. I would call it "Full Stack", and there is often a need to produce results by going up and down the full problem. Unfortunately, this is a brutally tough problem at a startup because almost no one has the skills to understand the entire problem. Almost reminds of the 1990's internet days when you had to: 1. compile linux, 2. setup DNS, Configure Apache, etc AND write the code.* (Participant 1)

Diversity of approaches and methods has its benefits, and provides value during both data gathering and analysis.

***Market Trends***

As progress is the goal of every company, market trends tend to dictate the direction towards that goal. Adapting and developing the necessary technology and expertise in order to follow demands of the market is something also data scientists and analysts need to undergo. However, adaptation needs to be done properly, otherwise problems like the one Participant 7 stated, may occur.

> *Data Science is part science and part art. ....organizations are trying to tackle this new trend of AI/ML like they tackled other disruptions before. Make your current BI Manager as your Analytics Manager and voila! the problem is solved.* (Participant 7)

***Turning to Programming***

Another issue is that data scientists, somewhat often, think of themselves primarily as programmers, so they would leave their area of expertise in order to achieve a specific result. That can happen due to wrong assumptions of management, and data scientist role, which we will discuss later.

> *One of the challenges I face is that the DS want to be a developer until their stuff breaks and they get called at 2am to fix it. I think a DS should build prototypes but production is better left to developers/engineers.* (Participant 8)

***Software Development***

Since the biggest problem was a distinction between software development and programming on one side, and data science and analysis on the other, some participants outlined the purpose and duties of software developers, in order to point out the difference, along with synergy between business management, data science, and software development.

> *Lines and quality of code are not the only parameters to judge software development. If these lines and quality do not align to business objective it is a failure and to add to your statement if the software products that are gathering and providing you with the data on which data scientist do number crunching, are not build in the right way I am sure neither will be any data data science nor will be any software product. It all has to happen in harmony and as a unified system not fragmented and unstructured.* (Participant 6)

***Seasonal Job***

We found that participants think of data science as a more of a seasonal job, since after the analysis are done, data scientists are being moved to different positions that include programming and development. So, perhaps a better option for data scientists is to work part-time for multiple companies, than to be transferred to a position for which they do not have the necessary qualifications.

> *Just playing devils advocate here... So, what happens when a "data scientist" has revealed no useful (or additional) insights? Are the businesses obligated to still hold on to them when no value is being demonstrated? I can't see that a team is let go whilst they are adding value... My personal view is that many "data scientists" have skipped the general data analytics phase, which is why they are let go once the job is done. Your seasoned data analysts are adept at finding general problems and coming up with solutions. Machine learning and AI sounds great but most companies don't need it! Most value from an analytics perspective is still generated through "supervised" and deterministic methods. There is big gap between "General Data Analysts" and "Data Scientists".* (Participant 21)

***Suggesting the Sector of Implementation***

One comment in particular focused on applying data science and analysis to other types of industries, not just to business services.

> *Our annual reviews for schools are built in our shop, for presentation to client systems that lack capacity and can't afford in-house expertise. That's a kind of niche that needs to be served more often, even in the private sector.* (Participant 26)

***Job Title***

Relating to this theme, the participants expressed the question that is it just the title "data scientist" all that needs changing, or elaborating, or is it the whole job description, along with educational phase of developing an individual as a data scientist.

> *They [data scientists] need to be a resource, like engineers, rather than a project hire. Perhaps the name is the problem, most companies don't see themselves in the "science" business. So maybe Data Engineer would be a better name.* (Participant 37)

**Management Approach**

As Management Approach was highly prevalent in the coding, we split it into several subcategories in order to describe it in more detail. The subcategories can be seen in Table 2.

Table 2: Sub-themes of Management Approach.

| Sub-theme | Explanation |
|---|---|
| JD issues | Job descriptions (JD) are not properly set and are not in the balance with the skills a data scientist should have. Those JDs are either not consistent with the ones placed on the job posts, or during the interviews. Data scientists are required to have various other skills, like coding, software engineering, etc. Some of the comments are positive and include a proper distinction between JDs. |
| Leadership | The impact top management has when it comes to expectations, motivations, and organization inside the company. Also, the way they handle external requests, and how they pass those on their employees. Basically, the understanding of structures inside the company and its relation to data science. |
| Economical aspects | When there is an awareness of the requirements of the data scientists, but the lack of funds, or the need to save more or other economic reasons dictate the ways data scientist will operate, and what will be required of them. |
| Team of people | A data scientist is not just one role, one person, so, when making decisions, and planning a project, data science should be thought of as a team of individuals. |
| Clarity about data use and goals | Managers should always have a clear goal and vision of what they will use the data for, so that they can make adequate requests from data scientists. |

***JD Issues***

Most of comments under the Management Approach theme focused on Job descriptions, where comments containing negative feedback were made about companies' requests that data scientists also work as developers, and engineers, due to lack of precise definition of job responsibilities. Two data scientists point out this problem very well:

> *"One of the key issues is that managers treat data scientists as developers". Indeed, I have seen recently many positions requiring data scientists be capable to make iOs apps! I also think some corporations and startups are losing the compass of what data scientist is!* (Participant 46)

> *Data Science is part science and part art. The issue here is two fold. One, the IT recruiters are used to coders. I got sent the usual skill assessment form asking for advanced level SQL expertise. No problem with that. Problem was there was none on data science. Second, organizations are trying to tackle this new trend of AI/ML like they tackled other disruptions before. Make your current BI Manager as your Analytics Manager and voila! the problem is solved. Just browse here on Linkedin for Analytics Managers profiles in large IT/Consulting organizations. The usual skill set like SQL, OLAP, SSRS, SSIS, Power BI, and point & click data visualization tools. How would they engage a Data Scientist?* (Participant 7)

We can conclude that people feel comfortable having well-defined job description, also in the sphere of the "murky" data science. Also, one entrepreneur commented that organizations should ease up on the job descriptions, since that does not help teams become more coherent.

> *I think if we had less names and definitions for roles in a single business world, we'd do a better job as a team. I bet there is no way to define all those existing roles; data architect, data analytics, data intelligence, data scientist,... same goes to process, system,etc. And by the way, your definition of a sw developer is very interesting too. I believe a true sw developer would start with business need on data side and ends her/his work when the right data is out.* (Participant 64)

***Leadership***

Some management experts have suggested that managers do not need to be expert in the field they are managing, or leading (Wallace & Creelman, 2015). Rather, they need to have understanding and apply business knowledge and expertise. Based on the analysis, we find that the problem is in understanding of data science and its position in the organization. Participants stated that proper understanding of roles, organization's vision and having clarity are necessary when setting goals and expectations for the data science team.

> *Ironically, the terms Data Analytics, Data Science, Business Intelligence et al tend to be misnomers in the way these are practiced in the industry, and especially in the way a lot of consulting companies offer these services. Most of them seem to be oriented towards "technology implementation" of the above. Could be because they are led by technologists, or because that is the service that clients want out of these companies. In a typical Problem -> Data -> Insights -> Actions*

> *journey, consulting organizations that offer Data Analytics tend to focus on the 2nd leg of this journey i.e. Data to Insights where the focus is on data structures, sources, integrations, implementation etc. All these things make it a technical story. And therefore, a sense that data scientists/analysts are coders. However, the 1st leg of identifying the Business Problem, and decomposing it into a Data Problem seems to be out of scope. Maybe organizations need to focus on building Data Science/Analytics practices that has a healthy mix of 2 types of roles: 1) Business Analysts with relevant industry expertise and background/understanding of Statistics, and 2) Technology developers.* (Participant 18)

***Economical Aspects***

Cutting teams, and merging roles in order to cover expenses, or to create a higher profit in the short term, are the main reasons for dissatisfaction of data science and data analysis teams. Also, not realizing the need for a fully functional data science team will lead to one of the issues mentioned earlier, which is, hiring data scientists for permanent positions, where they will actually do data analysis only temporary, and then be moved to other positions.

> *When data science teams are set up just as a checkbox to prove stakeholders that we are also doing something with data... disillusionment generally happens. Once backed by strong business use cases and clearly defined expected outcomes - it can do wonders. However, very few clearly understand the truth behind the hoopla of data science!!* (Participant 10)

> *Bang on.. Most commonly faced problems by data sciences organizations.. Clients end up creating a 90 day analytics plan sometimes.. I mean come on!! Everyone needs to understand that data sciences or rather decision sciences is a slow medicine.. You just can't expect very short term outcomes.. And even worst part is that today I see a lot of software companies masquerading as data sciences companies not solving business problems rather developing black box solutions.. These put more pressure on genuine data scientist companies to deliver quick outcomes.* (Participant 71)

We can conclude that data science does not necessarily provide quick results, but rather more valuable, long term, and deeply relevant results, so patience is a key characteristic of this discipline.

***Team of People***

The meaning of a word “team” here is not necessary a group of data scientists, but a team of people with different roles (both horizontal and vertical) who are working together on solving a problem, in a systematic manner.

> *Tom Ordonez shared with me that Data Science is not a role but a team. You have to have a data scientist, a business leader, a UI/UX specialist, and a programmer. A lot of organizations are under the mistaken belief that hiring a Ph.D. is all that is required. Data science requires a dedicated team to deliver actionable insights.* (Participant 24)

***Clarity about Data Use and Goals***

Having a vision, mission, and clear goals are, according to the participants, crucial to having an elevated level of efficiency and productivity inside the organization. Making sure that all employees are aware of those goals is also what needs to be done, for the previous to take place.

> *I would agree with the sentiments of this post. Aligning on the business goal, quick prototypes, and failing fast is very important in this space. I would also point out that, data-scientists treated as a developer analogy is not that bizarre though. We have to keep in mind that, in industry research without results may or may not be affordable all the time. As a developer has to make sure his/her basic idea works before architecting a software, so do we in making sure the research cycles are short, followed by quick validations and alignment on the business problem. Data Scientists = Hackers + Developer + Statistician + Machine Learning Experts + Business savvy + Researcher (in my opinion).* (Participant 58)

**Breakdown by Participant**

Table 3 shows a breakdown of the concerns by participants.

Table 3: Breakdown of the data science problems by the position of the participants

| Positions | Main Theme/Sub-Theme | Total Number |
|---|---|---|
| Entrepreneur | Leadership | 8 |
| | JD issues | 10 |
| Top Management | JD issues | 16 |
| Middle Management | Leadership | 12 |
| Lower Management | Management Approach | 3 |
| Data Analysts and Scientists | JD issues | 19 |
| Software Engineers & Developers | Knowledge and Skills | 4 |
| Consultants & Other | Various | 3 |

Entrepreneurs were mostly talking about leadership and issues regarding job description. Top management and Data analysts and scientists were pointing out that companies make mistakes while creating and requesting skills that should not be in the job description of a data scientist/analyst. Also, a couple of software developers and engineers focused on the skills, trying to differentiate them from Data Scientists. Most of the comments were posted by men, 67 out of 71 (94%). Women focused on management approach, especially on the job description of the Data Science & Data Analysis positions. Hence, we can conclude that women are very little represented in this line of business, or at least less vocal about the data science problems.

## CONCLUSION AND RESULTS

### Key Takeaway

We present here our key conclusions. First, management of data science seems to lag behind the knowledge and skills, and technological advances made in the field. The senior management, or organizations at large, are not aware of the full potential of data science and therefore find it hard to define the goals and objectives for data science programs. Second, according to the participants, data scientist is not just one role, but a combination of many. This causes friction in organizations which are used to neat, well-defined slots defining any given job position. Data scientist is a different kind of piece that does not fit into the current organizational puzzle. Third, based on the analysis, we argue that most businesses are still at the technology hype stage where the prospect of the analytics' true potential has not yet crystallized, and therefore organizations are trying to define a "moving goal." Our suggestion to these organizations is, in a nutshell: remain down to earth, and focus on practical and simple business goals over technological finesse and industry buzzwords.

### Theoretical and Conceptual Implications

This study provides novel insight into the details and practice of data science in organizations. We use an innovative way of sampling to get a glimpse into dozens of organizations at once, accounting for their employees' professional opinions relating to organizational problems of data science. Regarding prior literature, we confirm and add to existing studies. For example, we confirm the claim of Vidgen *et al*. (2017, p. 626) that "becoming data-driven is not merely a technical issue and demands that organizations firstly organize their business analytics departments to comprise business analysts, data scientists, and IT personnel, and secondly align that business analytics capability with their business strategy to tackle the analytics challenge in a systemic and joined-up way." The respondents indicated many pressing, inter-departmental and inter-processual issues – for example, the importance of recruiting, mostly ignored in the prior literature on data science in organizations, came up often, tying it success outcomes at least a conceptual level. However, some of our findings are contrary to prior literature. For example, unlike Provost and Fawcett (2013) who note that defining the boundaries of data, we find that defining the job role and organizational position of the data scientists is a priority for successful practical work. The concept, based on our findings, still seems slightly "murky" to many participants, leading to the need for an acceptable definition, which can be adapted to specific organizational situations. We have defined data science as defined as the use of computational techniques and big data to solve real-world analytical problems, and see that, based on the practitioners' use of terminology, this provides an ample working definition for data science in organizations.

### Limitations and Future Research

Every study comes with a set of limitations. In our case, the limited sample could be supplemented with more data to generalize the findings, operationalize the organizational constructs, and measure their interaction in a formal, quantitative study. However, this was an exploratory study with the purpose of discovering key issues as perceived by practitioners dealing with them in their daily professional lives. Further studies can address the shortcomings of this exploratory study. In particular, the following questions are left for further research:

- *How could the organizational readiness for data science be improved?* Based on the variety of the participants' answers, there are underlying differences in organization readiness for data science. Therefore, it is crucial for the less mature organizations to learn from the more advanced ones. To this end, more research is needed.
- *How could job expectations from a data scientist be correctly aligned with organizational resources and decision-making authority?* In our findings, we pointed out such incongruences as a part of the organizational reality. However, it is unclear what measures, communicative and otherwise, can be undertaken to avoid incongruences and fully exploit the potential of data science and data scientists in organizations.

## REFERENCES

[1] Acito, F., & Khatri, V. (2014). Business analytics: Why now and what next?. *Business Horizons*, 57(5), 565-570.
[2] Barney, J. B. (1986). Organizational culture: Can it be a source of sustained competitive advantage?. *Academy of Management Review*, 11(3), 656-665.
[3] Carlsson, C. (2017). Decision Analytics – Key to Digitalisation. *Information Sciences*, in press. Retrieved from https://www.sciencedirect.com/science/article/pii/S0020025517309192 (September 20, 2017).
[4] Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard Business Review*, 90(5), 70-76.
[5] Denison, D. R. (1990). Corporate culture and organizational effectiveness. John Wiley & Sons.
[6] Dent, E. B., & Goldberg, S. G. (1999). Challenging "resistance to change". *The Journal of Applied Behavioral Science*, 35(1), 25-41.
[7] Edwards, J. S., & Taborda, E. R. (2016). Using knowledge management to give context to analytics and big data and reduce strategic Risk. *Procedia Computer Science*, *99*, 36-49.
[8] Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Transaction Publishers.
[9] Hong, I. (2007). A survey of web site success metrics used by Internet-dependent organizations in Korea. *Internet Research*, 17(3), 272-290.
[10] Honkaniemi, M. (2015) *Markkinointiviestinnän määrä Suomessa 2014*, Luetta. Retrieved from https://mainostajat.fi/app/uploads/sites/2/2015/11/Markkinointiviestinnan_maara_Suomessa_2014_raportti.pdf (September 20, 2017).
[11] Lapointe, L., & Rivard, S. (2005). A multilevel model of resistance to information technology implementation. *MIS Quarterly*, 29(3), 461-491.
[12] Järvinen, J. (2016). *The use of digital analytics for measuring and optimizing digital marketing performance*. Jyväskylä Studies in Business and Economics 170. Retrieved from https://jyx.jyu.fi/dspace/handle/123456789/51512 (September 20, 2017).
[13] Neely, A., & Bourne, M. (2000). Why measurement initiatives fail. *Measuring Business Excellence*, 4(4), 3-7.
[14] Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M., & Kennerley, M. (2000). Performance measurement system design: Developing and testing a process-based approach. *International Journal of Operations & Production Management*, 20(10), 1119-1145.
[15] Nelson, R. R. (1959). The simple economics of basic scientific research. *Journal of Political Economy,* 67(3), 297-306.
[16] Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59.
[17] Schein, E. H. (1990). Organizational Culture: What it is and how to change it. In *Human Resource Management in International Firms* (pp. 56-82). Palgrave Macmillan, London.
[18] Strauss, A., & Corbin, J. M. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* Los Angeles: SAGE.
[19] Verhoef, P. C., & Leeflang, P. S. (2009). Understanding the marketing department's influence within the firm. *Journal of Marketing*, 73(2), 14-37.
[20] Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, 261(2), 626-639.
[21] Wallace, W. T., & Creelman, D. (2015). Leading people when they know more than you do. *Harvard Business Review*, June, 18.