

National Coordination of Data Steward Education in Denmark






Final report to the National Forum
for Research Data Management (DM Forum)

Results and recommendations
January 2020






Title: National Coordination of Data Steward Education in Denmark:
Final report to the National Forum for Research Data Management (DM Forum)

Publishing date: 31.01.2020

This report is authored by:

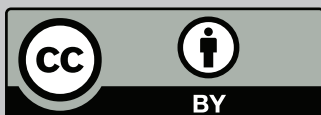
Lorna Wildgaard, Copenhagen University Library, Royal Danish Library  0000-0002-3900-5058
Evgenios Vlachos, University Library of Southern Denmark  0000-0001-8235-0423
Lars Nondal, Copenhagen Business School  0000-0001-5030-9813
Asger Væring Larsen, University Library of Southern Denmark  0000-0002-9967-2637
Michael Svendsen, Copenhagen University Library, Royal Danish Library  0000-0002-5807-5326

Contributors and participating members of the project are [in alphabetical order]:

Christian B. Knudsen, Copenhagen University Library, Royal Danish Library  0000-0002-9816-768X
Ellen Vibeke Knudsen, Aalborg University Library  0000-0002-5099-351X
Falco Hüser, Copenhagen University Library, Royal Danish Library  0000-0001-9645-6691
Gertrud Stougaard Thomsen, Aarhus University Library, Royal Danish Library  0000-0002-6090-9801
Lene Kaad, University College Copenhagen  0000-0002-7249-763X
Mette Hall-Andersen, Danish National Archives
Søren Davidsen, Roskilde University Library

Layout: Morten Bygvraa Rasmussen. DeIC

This work is licensed under a [Creative Commons Attribution \[CC BY 4.0\] License](https://creativecommons.org/licenses/by/4.0/)



DOI: 10.5282/zenodo.3609516

32 pages

All results from the final report of the DM Forum sponsored project: National Coordination of Data Steward Education in Denmark, [Jan. 2020] is available at:

<https://www.zenodo.org/communities/ds-edu-dk>

Copenhagen, January 31, 2020

National Forum for Research Data Management (DM Forum)

The project group would like to thank everyone who contributed their knowledge and time to this project. The interviewees, survey respondents, workshop participants, reviewers, and advisors. Thank you for your contributions and interest in our work.

The participating institutions would like to thank:

National Forum for Research Data Management and DeIC for financial support and especially Anders Sparre Conrad and Diba Marcus for their valuable assistance.

Read more about the project: <https://www.deic.dk/datamanagement/aktiviteter/tvaerinstitutionelle>

Index

EXECUTIVE SUMMARY	3
1 INTRODUCTION	4
1.1 MOTIVATION AND BACKGROUND OF THE PROJECT	5
2 REVIEW OF EXISTING DATA STEWARD EDUCATIONS IN DENMARK AND ABROAD	6
2.1.1 FINDINGS FROM THE REVIEW OF EXISTING DATA STEWARD EDUCATIONS IN DENMARK AND ABROAD	6
3 LINKEDIN ANALYSIS	8
3.1 DATA STEWARDSHIP SKILLS	8
3.2 FINDINGS FROM THE LINKEDIN ANALYSIS	9
4 JOB VACANCIES ANALYSIS	10
4.1 FINDINGS FROM THE LINKEDIN PROFILES AND THE JOB VACANCIES	10
5 QUESTIONNAIRE	12
5.1 FINDINGS FROM THE QUESTIONNAIRE	12
5.1.1 TERMINOLOGY SURROUNDING DATA STEWARDS	12
5.1.2 HARD AND SOFT SKILLS	13
6 INTERVIEWS	17
6.1 FINDINGS FROM THE FOUR INTERVIEWS	17
7 BUILDING ON EVIDENCE: MODELS FOR A DATA STEWARD EDUCATION	19
7.1 STUDENTS WITH A BACHELOR'S DEGREE	19
7.2 STUDENTS WITH A PHD OR EQUIVALENT RESEARCH EXPERIENCE	19
7.3 CONTINUING AND PROFESSIONAL EDUCATION	20
7.4 A RECOMMENDED MODEL FOR DATA STEWARD EDUCATION	20
7.5 PERSONAS	21
8 RECOMMENDATIONS	26
OTHER COMMUNICATIONS FROM THE PROJECT	29
REFERENCES	30
LIST OF ABBREVIATIONS AND ACRONYMS	31
LIST OF TABLES AND FIGURES	31
APPENDIX	32

Executive summary

This report provides an analysis of current educations in data stewardship and an evaluation of the needs and expectations to the role of a Data Steward (DS) across the private and public sector. The aim is to inform future data stewardship educations in Denmark. Methods of analysis included a combination of quantitative and qualitative methods such as literature review, text mining, questionnaires, correlation analyses and interviews. All supporting materials produced throughout the report are found in [Zenodo](#).

Results of the analyses show that the DS provides a multitude of services, that differ in the context of which he or she is employed. Primarily, the DS functions as a bridge between the organisation, infrastructures and the end user/customer. However, we identified four common roles that appear across all sectors and are supported to varying degrees in current educations. These are the roles of the DS as an Administrator, Analyst, Developer and Agent of Change.

In particular, an educational model is recommended that supports: experiential learning and internships; flexibility to encompass the activities and commitment that successful internships entail and, has a balance of practice in “hard” technical skills such as programming and work-flow optimization, soft skills in teaching, communication, networking and project management as well as disciplinary knowledge and good data management practices (ethics, policy and accountability).

The report finds the current position of data stewardship education in academia to be in its infancy, whereas in the corporate world education is much more advanced. Future education in data stewardship can benefit from a strong collaboration between academia and the corporate world, to ensure the education has a stronger recognition in stakeholder communities, including potential employers, thus improving the student's employability. The major weakness of current academic educations in DS, in the lack of opportunity for the graduate to up-skill through flexible, short term education. Further education should be offered to ensure skill maintenance.

Recommendations discussed include the following.

- Offering a two year Candidatus in data stewardship, with courses available through the Open University to support mature students with a professional background.
- A 1 year master programme in data stewardship (BA + 1), on the condition it is open to students with a minimum of a bachelor. Thus, the programme could also be open to master students, PhD's and through the Open University to mature students. On the aforementioned condition, the 1 year education provides a dynamic learning environment and relevant upskilling programme that can rapidly supply DS to the hungry job market.

- Clearly define what a DS is in the curriculum.
- Using teaching and learning activities that have a degree of flexibility to support both internship programmes and the student's academic course work.
- Using an innovative combination of on-campus teaching and online learning, bootcamps and workshops, preferably developed and taught in collaboration with local industry, departments and organisations.

The report also investigates if a national coordinated education in Denmark should be developed between universities and industry or if competition between these parties is a healthy competition, resulting in diverse educational programmes appealing to different audiences. Data stewardship is rapidly evolving and holds a key role in digital society in the public sector, the corporate world, as well as the research and education (R&E) sectors. We strongly recommend therefore a common national approach to university education in data stewardship and engage with national, European and international initiatives in a coordinated manner to ensure both an education aligned with policy, initiatives and directions in data stewardship and job mobility for our graduates.

The analyses used in the report have limitations. The investigation into the skills and job profiles of practicing DS in Denmark is biased towards industry. Only a few DS employed in academia were included in the study. This we consider to be caused by source-bias, in that the data used in the analysis was harvested from Linked In. The study could have been more representative if we had included other sources, to ensure an equal representation of DS employed in academia and those employed in industry. We only managed to undertake four interviews, thus we achieved cultural probing rather than a representative sample. The study was conducted within a short period of time, therefore we had to limit our input and also the scope of the project as well as the analysis of our results.

1 Introduction

Data stewardship has an important and rapidly evolving future role to play in many parts of the digital society encompassing the public sector, the corporate world, as well as the research and education (R&E) sectors.

As business processes and decision-making in the private and public sectors along today's R&E become increasingly data-driven, national and international frameworks for management of data have emerged with requirements of responsible handling and use of data, data protection, and terms of customer data services. Concepts of Public Sector Information (PSI), Open Science (OS) and FAIR (Findable, Accessible, Interoperable, Reuseable) data are accompanied by a political vision of a more open and transparent society in the future. Research communities, business enterprises and the public sector need to live up to new standards of data management to strengthen both accountability and competition in the digital single market.

A new labour of data expertise, "Data Steward" (DS), will take on diverse roles, tasks and responsibilities documenting, curating and structuring data across scientific, public or enterprise organisations. A DS will be responsible for aligning data processes and applications in the development and enforcement of data governance and compliance in relation to GDPR (General Data Protection Regulation), IT-security, data ethics and will facilitate the production of FAIR data (Vollertsen, 2019). These activities are key to the future success of data-driven organisations; helping the organisation avoid missed opportunities or ineffective or poor business decisions. Employing a DS to engage in the aforementioned tasks, is a step in the right direction towards responsible datafication of an organisation, ensuring informed decisions based on quality data, leading organisational and research strategies and priorities, developing relevant services and ultimately resulting in the organisation having a competitive edge in a market already leveraging massive amounts of data.

According to experts (European Commission & Directorate-General for Research and Innovation, 2016), the world needs DSs – and a lot of them! The shortage of data experts with intermediary expertise requires a new breed of people with DS skills. Estimates say that in Europe alone several hundred thousand DSs will be needed over the coming years to bridge the gap between e-infrastructure, organisations, data and users or customers, (Versweyveld, 2016).

The need for DS expertise was echoed at the workshop we held in October 2019, with stakeholders from Danish enterprises, research organisations and the public sector. Yet the question was raised as to what a DS actually was. Many companies had people undertaking the responsibilities of a DS, but as part of different job profiles, for example as a Data Librarian, Data Analyst or DPO. Many simply had not heard the term "data steward" before, yet the description of the tasks of a DS was recognisable and a profile much needed in the represented companies. (DM Forum [DeIC], 2019).

DSs, are the "humanware" that can connect vast amounts of data with people who want to access, harvest, protect or exploit the potential of the data. People involved in working with data in the public and private sector (e.g. employed in providing data infrastructures, as network providers, IT-developers, researchers, administrators and in libraries) all have a shared responsibility to promote high-quality data by being well-curated, -documented and -structured. But documenting, organising and preserving data is labour intensive and time consuming and the temptation is to use machine-learning and artificial intelligence to make these processes less labour exhaustive. We argue that best practices for quality assurance of data can't be solved using machine learning or artificial intelligence alone. People are still needed to work with and support the complexity of producing and governing quality data. This is where the DS can step into organisations to help develop and establish responsible practice of stewarding data in close collaboration with specialised competencies of data managers, data scientists, security specialists, IT staff, researchers and administrators across private and public domains.

The concept of a DS is not a new one. They have existed for more than 20 years in business and enterprises where the roles of a DS are many, depending on the organisation and the context in which they work. Kimball (1998, p. 35) describes the role of a DS in businesses as:

"...responsible for driving organizational agreement on definitions, business rules, and permissible domain values for the warehouse data, and then publishing and reinforcing these definitions and rules",

and continues defining the DS as a new role within the organizational structure:

"Historically, this role was referred to as data administration, a function within the IT organization. However, it's much better if the data steward role is staffed by the subject matter experts from the business community".

As Kimball points out, subject expertise gives key insights in how data should responsibly be processed and governed, and is as such also expected as the basic competence of a DS in academic and public organisations. Yet still, in the academic and public sector, data stewardship is in its infancy. At the same time as the roles and functions of the DS are actively being investigated (Sempreviva 2019; Börner et

al, 2018] European directives and mandates are pressing for competent DSs now - to act as organisational representatives responsible for handling, administering and facilitating the interoperability of data. For example, requirements from the European Commission related to Open data and FAIR data, rules of participation in the European Open Science Cloud (EOSC) for research performing organisations in the implementation of the Public Sector Information (PSI) Directive in EU member countries.

Data is important. For businesses, governments, health sectors, education, research, and society. It is important that all possible measures are taken to ensure that data is responsibly collected, created, produced, processed, stored and shared and decisions based on data are valid, reliable and effective. A DS education will allow individuals to become experts in exactly this. It will allow them to explore “datafication” issues within organisations on a local, national and international level, engage in problem solving, and take action to improve how data is processed and governed. As a result, we as a society will all reap the benefits of a workforce equipped with the necessary knowledge to ensure transparency and responsibility in handling our data, and skills to liaise with organisations and clients on data issues across boundaries of practice.

1.1 Motivation and background of the project

Despite the shortage of DS expertise in the public and private sector, international and national DS educational initiatives in academia are few and at a master or diploma level perhaps non-existent. Currently, there is no formal university education offered in Denmark to train the vital position of a Data Steward. This project has been initiated and coordinated from the National Forum for Data Management (DeIC) with the aim to answer the following fundamental questions.

- Is a national Data Steward education a feasible solution?
- Is there a rationale in terms of rewards and benefits for a national coordinated effort?
- Should the development of Data Steward education in Denmark be in collaboration or competition between universities and between businesses?

In this report, we point towards a possible DS education in Denmark. We provide ideas and recommendations that universities and other educational institutions can use as a starting point in the prequalification of a new education in data stewardship. Some of our findings can also be applicable in other, non-academic settings and alternative educational environments.

The report is structured after the methodological strategy we used to investigate the current status of DS educations and the data stewardship. Each section presents a short summary of the analyses and the main findings. More details of the analyses are available in the [Zenodo community of the project](#).

In the following section, Section 2, we present the summary and findings from the review of current DS educational programmes. This is followed in Section 3 by an analysis of DS job profiles on LinkedIn. In Sections 4 and 5 the investigation of the required skills and expectations to a DS are presented, respectively through an analysis of job vacancies and a national questionnaire. Section 6 summarizes interviews addressing the expectations to DS in an organisational context. Finally, in Sections 7 through 8, we present our recommendations followed by a brief afterword backed with advisory next steps. The sections follow the project’s timeplan and indicate the steps taken in chronological order.

For further interest in our work created by the project, we wish to highlight our collected output and products, that are openly available in the [Zenodo community of the project](#).



2 Review of existing Data Steward educations in Denmark and abroad

The review aimed to describe the differences and commonalities between current DS educational programmes in Denmark and abroad. 24 data steward programmes were reviewed: 8 from academia and 16 from the corporate world. The background of the study, methodology, results and conclusions of the analyses carried out are described in the report: “Reframing Data Stewardship Education in Denmark and abroad” [Wildgaard, 2020] which is freely available in the DS community pages on Zenodo. We present a summary of the findings below.

The high level observation is that educational possibilities as a DS are few. The lack of educational opportunities for undergraduates present a golden opportunity for the development of innovative and relevant programmes that could educate dedicated DS professionals.

Corporate bodies offer on-the-job training to fulfill DS functions relevant for the specific enterprise and commercial system and platform developers offer courses in data stewardship using their product[s]. The training often comes in an innovative format that suits the busy workday of the professional - webinars, online on-demand training, close collaboration with industry experts in data stewardship, virtual instructor-led courses and boot-camps.

At the university level, data stewardship education is approached with a traditional teaching pedagogy that require the participant’s physical presence on campus. The course typically fills a single module on data scientist education, or as in-house training for PhD-graduates who wish to function as research support and at the same time put their own research experience to good use [Kurapati, 2019].

The majority of the academic data stewardship programmes were on a post graduate level and bound to research support and the research data lifecycle. The application of data stewardship in the corporate programmes was more diverse and the learning aims can be grouped into three perspectives of data stewardship: data manager, data support and data governance.

Specific findings were discovered in through a content analysis of the programmes [Vlachos & Wildgaard, 2020]. Using Nvivo, we coded each programmes’ description of outreach, audience, domain, admission requirements, course format and the expected job titles a graduate of the programme could fulfill. Next, a thematic analysis of the coded content was completed. We investigated if input from corporate and university influences the content and structure of the programmes, the aims of the programmes, the skills and competencies a participant is expected to master as well as the demands and expectations to them before and during the programme, and finally how the programme is certified and requirements to continued certification and employability as a DS.

2.1 Findings from the review of existing Data Steward educations in Denmark and abroad

The results of this comprehensive analysis gave rise to a set of recommendations within 5 themes.

- **Collaboration.** Universities and industry that unite to develop curricula and course modules are able to facilitate a sustainable collaboration with disciplinary actors and other stakeholders, including the data stewardship community. Collaboration creates alignment between education and the job market, ensures the relevancy of learning outcomes and production of competent [employable] DS.
- **Student profile and course content.** The majority of the educations, analysed in the review were marketed as post-graduate opportunities. The outset are persons with well-established disciplinary knowledge and experience on a professional level, preferably with existing data and programming skills. Course content is mediated through in-house training as DS at specific faculties or in local businesses, where the disciplinary effects and aims of data management are taught. To establish an education for undergraduates requires a different approach. A flexible study programme, that frame-sets the individual course within relevant professional perspectives is recommended, with activities that support experiential learning yet maintain a structure that the less mature student can rely on for grounding and guidance. Students need to experience, not just learn about how the presented theories, methods and skills are relatable to current workplace practices and challenges.
- **Skills and competencies.** The following mandatory skills were identified across academic and corporate programmes.
 - a. Technological expertise: programming, analytics, and data integration.
 - b. Value for end users: business intelligence, customer insights, protection of information assets, policy and strategy development & implementation, research support & research data lifecycle, ability to leverage data for maximum value.
 - c. Data management, FAIR data management, data governance, data security, ethical and legal data stewardship.
 - d. Understanding of data stewardship as a profession, the function of data stewards, ability to assess the effect of data stewardship within and outside an organization.
 - e. Communication, teaching, liaison skills, knowledge and contact with surrounding communities.
 - f. Disciplinary knowledge: specialization in specific area of data stewardship. This could be as broad as choosing a research or corporate direction or investigating the research and corporate requirements to stewardship within a specific field, fx the life sciences.
 - g. Project management and the ability to work in multidisciplinary teams.

Please note: Renewal of skills is recommended, and an obligation to the candidate to maintain and develop their knowledge through active participation in networks and training events. Commitment to skill maintenance also is an opportunity to network, facilitate connections, and define levels of proficiency (as because candidates will have different degrees of knowledge about different disciplines, and acquire different skill sets throughout the education). Thus, it is advised to consider further education, as workshops, online training programmes, or short courses in collaboration with local partners from industry, professional and trade unions.

- **Certification.** The analysis emphasised the importance of peer recognition of the certification. By collaborating with industry partners the university education has the potential to receive the seal of approval in the surrounding community through association accreditation and consequently increase the visibility of candidate for potential employers.
- **Employability.** Employability encourages the student's motivation, and how organised and prepared they are for class and the maturity in which they will interact with learning activities. Overall, the term employability is seen as the inducement for the degree of motivation, organisedness and preparedness. The candidate's chances of employment are increased in educations that coordinate training and internships with stakeholder sectors and match job and policy requirements to curricula.

"A data steward can have many roles. For example an analytical role where the data steward is focused on data quality, customer support and crunching data for the industry or doing research for the university; a more administrative role, focused on data service support and guidance on policies or perhaps the role of the developer, who is project oriented, emphasizing data management, procedures and testing with technical and analytical skills."

Anonymous quote from the survey



3 LinkedIn analysis

LinkedIn was scraped for data steward profiles in March 2019. The aim was to gain insight into how data stewardship is currently implemented and positioned in Danish organisations, both private and public, and learn more about the professionalization of research data management and data stewardship functions.

- Where are DS employed?
- What functions do they fulfill and which skills do they need?
- Are there commonalities across organisations regarding the role and skill-sets of the data steward that could inform a data stewardship education?

We identified 85 DSs with a LinkedIn profile using the search terms “data steward” and Denmark. Two profiles were not publically available and along with 9 persons identified as not DSs, were removed from our sample. This left us with 74 DSs, 39 currently active as DSs, the remainder listing DS as a previous position. 35 of the DSs were female, and 39 male. Their educational background was listed as university [42%], business school [40%], professional college [10%] and on-the-job training [7%], while 3 profiles did not list their education. The majority were currently employed in the capital region of Denmark [76%], 15% employed in Region Zealand and the remainder in the Region of Southern Denmark. The organisations in which DSs are active, represent a cross-section of Danish industry, both commercial and research industries, though universities were not represented. From the analysis of profiles it is clear that DS have multiple roles in providing data support with organisations, and are employed in many diverse areas [Table 1].

TABLE 1. Organisations employing data stewards and the titles applied to data steward function.

Sort of organisation	Role	Nr.
Business solutions, Engineering & computing, Food, Healthcare, Lifestyle, Pharma, Telecommunications	Data steward, including junior, senior, associate, master, global	35
Insurance, Law, Ministry, Pharma	Other including buyer, planner, clerk, consultant, librarian, information specialist	15
Engineering, Finance, Lifestyle, Pharma	Customer-data steward	9
Engineering & computing, Finance, Food, Pharma	Data system developer	7
Food, Technology & computing, Pharma	Business intelligence	6
Finance, Pharma, Engineering & computing, Telecommunications	Data analyst	5

The analysis of the DS' LinkedIn profiles allowed us to gain an overview of the qualities and skills needed in their jobs and accordingly suggested which competencies, knowledge and skills need to be nurtured in a DS education. We categorized their skills in two groups: the Hard skills, such as pro-

gramming, software and system knowledge, policy awareness; and the Soft skills such as the ability to communicate, teach, or work in a team to name a few.

“...It is a big thrill to excel in technical disciplines, but the feeling of bringing business and IT[-people] together and making them understand each others needs and constraints are even better. In my experience, many projects fail because of misunderstandings and lack of communication, and I therefore always strive to be both precise and fulfilling in my own communication and to take into account who the receivers are, and what they are interested in knowing.”

Citation from Data Steward profile publically available on LinkedIn

3.1 Data stewardship skills

The profiles of the DS in our sample indicate the necessity of mastering both hard and soft skills in the workplace to provide professional data stewardship.

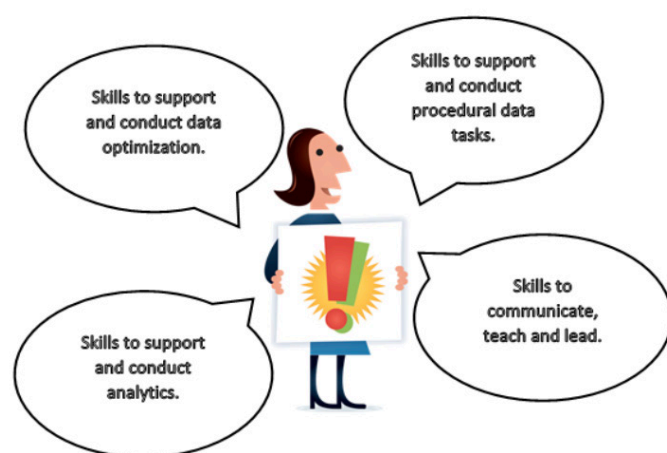


Figure 1. Data stewardship skills. Image adapted from Jørgen Stamp, <http://www.digitalbevaring.dk>

We applied text mining and machine learning techniques with the use of R [free software environment] for statistical computing to filter, process and analyse the corpus of the collected Hard and Soft skills descriptions as presented in the scraped LinkedIn profiles [Feinerer & Hornik, 2015; Bouchet-Valat, 2014; Hennig, 2015; Maechler et al., 2016]. Text in danish was translated to english. Figure 2a presents a frequency plot of the 57 terms [out of the 720 unique terms] that appeared more than 5 times in the profiles when describing their Hard skills. Figure 2b presents a frequency plot of the 21 terms [out of the 302 unique terms] that appeared more than 3 times in the profiles when describing their Soft skills. For more details on the method see [Vlachos & Tan, 2018].

While indicating the skills of DSs in industry, the analysis of the LinkedIn profiles failed to identify the extent the DS was embedded in the activities that produced data or if the DS provided generic services and guidance; if one DS was serving the entire company or more than one DSs were employed, each serving a different department thus having different functions and hence skills. However, the analysis does show us that DS functions consist of a mix of diverse tasks dependent on the domain of application and flexibility in the role of the DS is essential. Yet despite this multiplicity in DS profiles, there are some commonalities in the described hard and soft skills sets that could beneficially be included in a generic DS education: computer science skills, didactic and pedagogy, project and workflow management, user and domain analysis, data governance and project management.

3.2 Findings from the LinkedIn analysis

Based on the previous analysis we have identified the following Hard skills and Soft skills for people in existing or previous data stewardship positions.

Hard skills (computer science skills, user and domain analysis and data governance)

- **Conduct and support procedural data tasks:** data collection, data cleaning, data integration, data migration, data tracking, data modelling, error management, data quality assurance, mapping and linking data, data validation and data storage, technical aspects of curation and preservation.
- **Conduct and support analytical tasks:** crunch the data to gain new insights from the data itself or about the organisation, their products, customers or their competition (programming languages, software expertise, machine learning, testing and methodologies, user experience, business intelligence and statistics).

- **Conduct and support data optimization tasks:** maximize the comprehensiveness and speed of data extraction and analysis through the development of systems and practices such as data enrichment, sharing, documentation, process revision and streamlining, mapping and linking data, data architecture, and the enhancement and use of ontologies and vocabularies.

Soft skills (didactic and pedagogy, project management and data governance)

Importantly, the data steward has the necessary didactic and pedagogic skills to teach and guide others. This includes the following.

- **Communication:** ability to communicate across different platforms, using different media to connect with different user-groups.
- **Mediation:** proficiency to act as a “bridge” between data infrastructure, users/staff/customers and the organisation.
- **Facilitation:** data steward as “the go-to-guy” and as such is assertive, social, a good listener, a networker and works well in a team as well as a team-leader.
- **Learning:** commitment to be a curious, a life-long learner who enjoys challenges and sharing knowledge.

All skill-sets require a strong foundation of domain, organisational and user (customer) knowledge supplemented with aptitude in policy implementation, best practices, data management, governance and ethics, data security and privacy. This craves a personality that is enthusiastic about the details, yet structured and has the political skills and writing skills to develop procedures and guidelines.

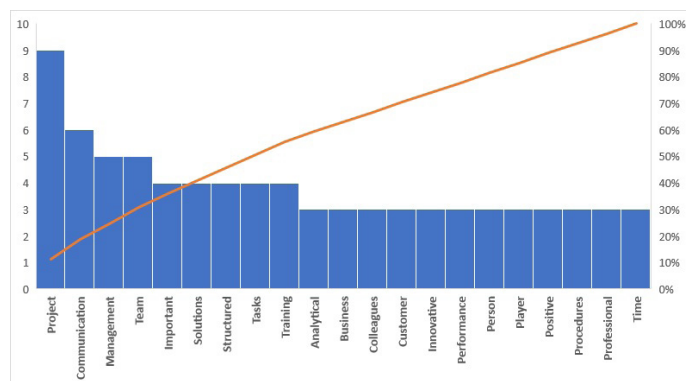
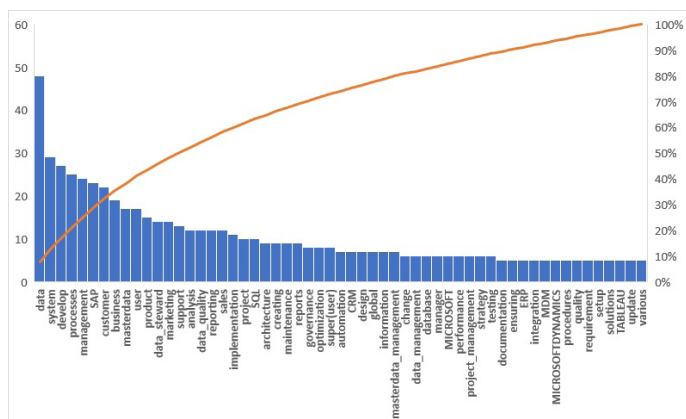


Figure 2. Pareto chart of the most frequent terms appearing in the description of the Hard Skills [a] and Soft Skills [b] from the LinkedIn profiles in descending order of frequency with a cumulative line as a percentage of the total.

4 Job vacancies analysis

To investigate further how data stewardship is organized through task and skill definition, job vacancies were collected from multiple job-seeking databases both within and outside of Denmark. The list of job-seeking databases, the vacancies retrieved and the methodology describing the data collection can be found in [Zenodo](#). The criteria for the search were the following.

- The intervention (Steward) or equivalent and the context (Data) had to appear either in the title of the position or in the text of the application.
- The language had to be English.

Due to a strict time window, the results had to be limited to vacancies retrieved on the 25th of March– 3. April 2019. Each job-seeking database, usually, has their own search criteria. The danish databases [only] had few jobs on Data Stewards, therefore depending on their search criteria we got proposals for relevant vacancies that included the terms Data Scientist, Data Consultant, Data Architect etc which were included in the study. We screened the vacancies on title and text manually for relevance and the resulting corpus finally consisted of 119 vacancies from three geographic locations: 18 from Denmark, 41 from the European Union, and 60 from the United States of America.

“Not one vacancy required applicants to perform tasks from purely one task area”.

Observation from collecting the corpus
for the analysis of job vacancies

We performed a content analysis of these 119 job vacancies, again using text mining in R. In order to prepare the corpus for analysis, we needed to preprocess it. Initially the corpus contained 5697 terms with 51% sparsity (meaning that about 50% of the terms were connected to similar and related terms forming a structured semantic matrix). Afterwards, we filtered the corpus, removing white space, symbols, punctuation, English stop words such as “the”, “a”, “in” etc. and words that have no analytical value and would bring noise to the results such as “email”, “apply”, “job”, “work”, “vacancy”, “application”, “opportunity”, “etc”, “based”, “can”, “english”, “people”, “read”, “years” and so on. Lastly, we assigned abbreviations to terms appearing together, for example when the term “data” was followed by the term “quality” the term “data quality” was registered instead. We ended up with 828 meaningful terms.

In Figure 3 we present the terms that appeared 50 or more times in the text corpus.

4.1 Findings from the LinkedIn profiles and the job vacancies

We created a combined corpora of the job vacancies, LinkedIn Hard skills and LinkedIn Soft skills text corpora, and applied machine learning techniques to interpret our findings in a meaningful way. We actually used topic modelling to discover the hidden semantic structures in our combined corpora. A topic is a distribution over a fixed vocabulary containing words with high probability, therefore a topic is made by words that usually appear together in the same phrase or sentence. We used the Latent Dirichlet Allocation probabilistic model and the variation expectation–maximization algorithm for modeling our text corpora and presenting it as a mixture of topics comprised of terms with certain probabilities. For more details on the method see [Vlachos & Tan, 2018].

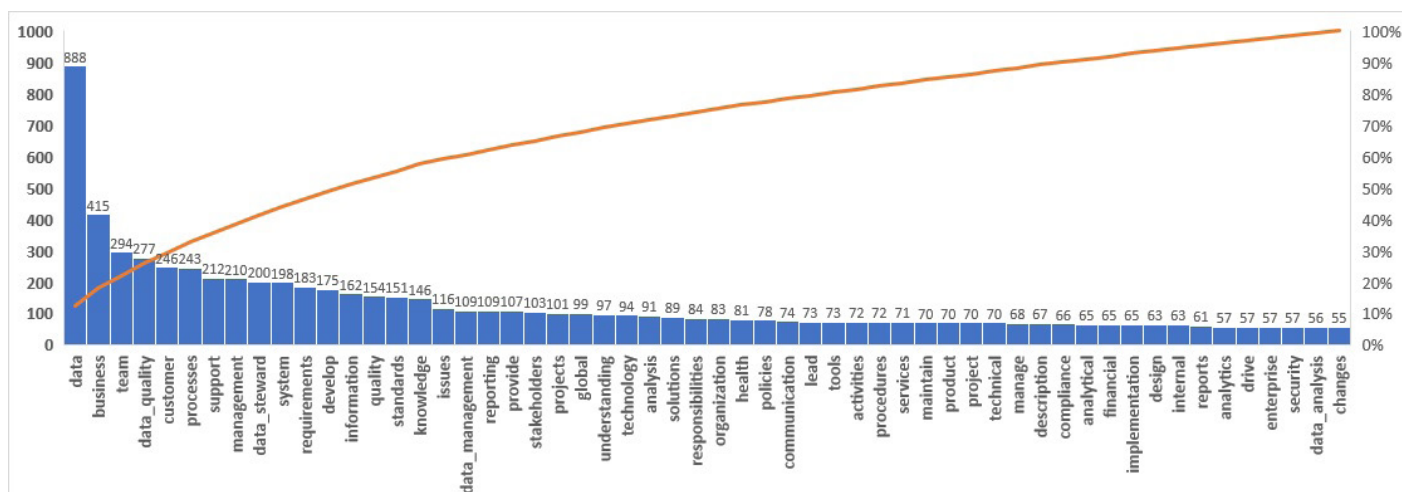


Figure 3. Pareto chart of the terms from the job vacancies analysis in descending order of frequency with a cumulative line as a percentage of the total.

Based on the aforementioned probabilistic model, Table 2 presents the most frequent terms in ascending order of specificity. Terms typeset in boldface are unique terms for that column [corpus] and help define it. Terms in green are common in all three facets of a suggested role of a data steward, see below, while terms in orange are common in only two facets of a role. Terms in blue indicate tools and software.

Reading through the terms in Table 2, we can make some inferences on what the three primary, and equally important roles of a data steward in the private or public sector could be:

- **The Administrator:** focused on business intelligence, implementing solutions and knowledgeable in data-service support, providing management with key performance indicators
- **The Analyst:** focused on data quality, understanding customer needs, and crunching data for reports to the management.
- **The Developer:** project oriented, emphasis on data management, procedures and testing with technical and analytical skills.

TABLE 2. Suggested roles of a data stewards and identified topics, common in the combined corpora of Job Vacancies,

The Administrator			The Analyst			The Developer		
Job Vacancies	LinkedIn Hard Skills	LinkedIn Soft Skills	Job Vacancies	LinkedIn Hard Skills	LinkedIn Soft Skills	Job Vacancies	LinkedIn Hard Skills	LinkedIn Soft Skills
business	data	project	data	processes	communication	data	data	project
management	system	communication	data quality	sap	management	business	develop	team
global	sap	innovative	business	management	training	data quality	system	tasks
team	business	team	team	customer	customer	team	masterdata	management
develop	customer	professional	support	develop	structured	customer	management	important
information	marketing	solutions	customer	business	tool	data steward	super [user]	close
organization	data steward	positive	data steward	marketing	important	requirements	processes	structured
reporting	user	player	description	masterdata	collaboration	quality	customer	training
system	sales	business	system	product	outgoing	management	data steward	player
security	product	colleagues	knowledge	reporting	solutions	support	sales	Procedures-processes
support	analysis	performance	international	data quality	professional	information	support	life
solutions	implementation	importance	monitoring	user	simple	system	project	time
implementation	reporting	analytical	analytical	maintenance	various	standards	manager	performance
compliance	data quality	developing	reporting	governance	tasks	develop	product	human
relevant	database	making	ensuring	architecture	focus	data management	implementation	challenge
customer	sql	great	projects	change	understanding	processes	analysis	games
risk	architecture	extra	finance	ensuring	procedures-processes	issues	global	create
requirements	information	data	tools	reports	extra	knowledge	data management	seek
financial	management	execution	activities	tableau	user	analysis	data quality	knowledge
standards	develop		compliance	microsoft		provide	design	
policy	masterdata management		research	creating		stakeholders	project management	
stakeholders	creating		management	project		understanding	performance	
strategic	mdm		employees	sql		product	reports	
privacy	microsoft dynamics		global	automation		responsibilities	governance	
knowledge	automation		looking	flexplm		projects	crm	
services	reports		changes	analysis		procedures	creating	
technology	processes		processeses	data steward		policies	maintenance	
health	optimization		understanding	various		technology	optimization	
areas	production		quality	optimization		manage	sql	
healthcare	global		operations	system		technical	microsoft	

5 Questionnaire

The review of DS educations, LinkedIn and job vacancy analyses informed the content of a questionnaire. The aim was to gain stakeholders perceptions of the range of activities, skills and practice for DS in the Danish academic and corporate settings. Accordingly, the recipients were questioned on four themes:

- familiarity with the term “data steward”;
- the function and tasks a DS fills at their company;
- the skills and competencies of a DS; and
- the need for DSs.

In a pilot study, the linguistic, face and content validity of the questionnaire was evaluated following the principles described in Kim, 2009. After adjustments based on comments from the evaluators, the questionnaire was made available in August 2019. The target audience for this questionnaire were people employed in all types of data driven environments and at all levels across an organisation. This included people working with the practicalities of data, including, but not limited to, producing data, data analysis and data management, and people who work in administrative and managerial roles in organisations that work with data.

The questionnaire was closed one month later in September 2019. As we used a snowballing technique of distribution, asking the recipient to send the questionnaire on to other potentially interested parties, we have no knowledge of the response ratio to sample size. However, as the questionnaire was available through SurveyXact, we do know that 350 recipients opened the survey, 86 recipients completed the survey fully and 62 provided partial answers. 125 responders came from universities, libraries, hospitals and the public sector in general (84%), 19 came from private companies (13%) and 1 from an NGO (1%). 66 (47%) respondents have encountered the term DS for the first time in this questionnaire, while 73 (53%) were already familiar with the term. In the question regarding the likelihood a DS [or equivalent] to be hired in the organisation within the next couple of years, 45 (41%) responded positively.

In the following subsections we present a brief description of the main findings from the questionnaire study. All the questions along with the report of the responses produced from SurveyXact can be found in Zenodo.

5.1 Findings from the questionnaire

5.1.1 Terminology surrounding Data Stewards

The concept, function and tasks of a “data steward” are highly dependent on the place of employment. However, the common theme is that the DS is the person that acts as a bridge between the organisation, the infrastructure and the users (including customers). Accordingly, title given to a role of a data steward varies greatly - from ‘data manager’, ‘core data expert’ to ‘data science researcher’ as seen in Table 3.

TABLE 3. Titles given to the role of a DS.

Preferred term	<i>n</i>
Don't know	37
Data manager	29
DPO	13
Data Steward	10
Data librarian	6
Information Management Specialist	4
Arkivar / Digitaliseringskonsulent	3
Data scientist	2
Dataansvarlig	2
IT specialists	2
Records manager	2
Consultant	2
Colleague	2
Administrative medarbejder - IT	2
Project managers / dept. coordinator	2
Business intelligence analyst	1
Data Engineers	1
Data handler	1
Data koordinator (coordinator)	1
Data specialists	1
Data / system owners	1
Datadisciple	1
IT Risk Management	1
IT-koordinator	1
Systemansvarlig og digitaliseringskonsulent	1
Projektmedarbejder	1
Research Data Officer	1
Research supporter	1

These titles have a different connotation outside of the organisations in which the person is employed. Some titles imply a strong focus on compliance control, for example, which can be confusing. It can cause potential job applicants to have inaccurate expectations and frustrate researchers who need data support. Figure 4 illustrates the status of awareness of DS in their places of employment.

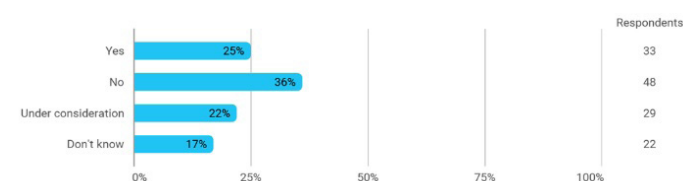


Figure 4. Are DS [or equivalent] a specific function in your organisation?

“...so many of the terms are still completely interchangeable or incomprehensible to me. I actually think this might be your biggest challenge to get this project flying... finding a common language.”

Free-text comment from questionnaire respondent

5.1.2 Hard and Soft skills

The respondents were asked to select 10 Hard skills, technical and didactical, from a list with 32 options considered relevant for a DS. Likewise, they were asked to select 10 Soft and communication skills from a list with 18 options. These skills were all identified in our previous analyses. Discovering the skills needed for a DS to operate in Denmark will help dictate the curriculum requirements. To be able to explore the differences and commonalities between employment sectors, we posed the question: “Are you employed at a university?” where 55% (n:81) of the respondents replied “Yes” and 45% (n:65) of the respondents replied “No”, meaning that they were employed either in the private or public sector excluding academia. This division’s aim is twofold: it will assist in comparing the findings from the previous se-

ctions - the three roles of a DS - to the results from the survey for those who responded with a “No”, and create a requirement list with the skills and needs of the DS who will follow a career in academia.

Using R, we identified hidden structures and patterns of Hard and Soft skills through a hierarchical clustering method. Formally this means we created four correlation matrices with the responses reordered according to the correlation coefficient [Figures 5-8]. The black boxes in the figures indicate the main clusters. Positive correlations are displayed with blue dots and negative correlations with orange dots. The more intense the colour and the bigger the size of the dot, the stronger the correlation between the skills, which means that these skills appear together (or are needed together) most of the times. Weak correlations, correlations with p-value > 0.01, are considered insignificant and are left blank.

We also calculated the standard deviations per skill and per participant. A skill with zero standard deviation indicates a skill that got no responses in the questionnaire, meaning that the responders did not find this skill important at all. Likewise, zero standard deviation for a participant means

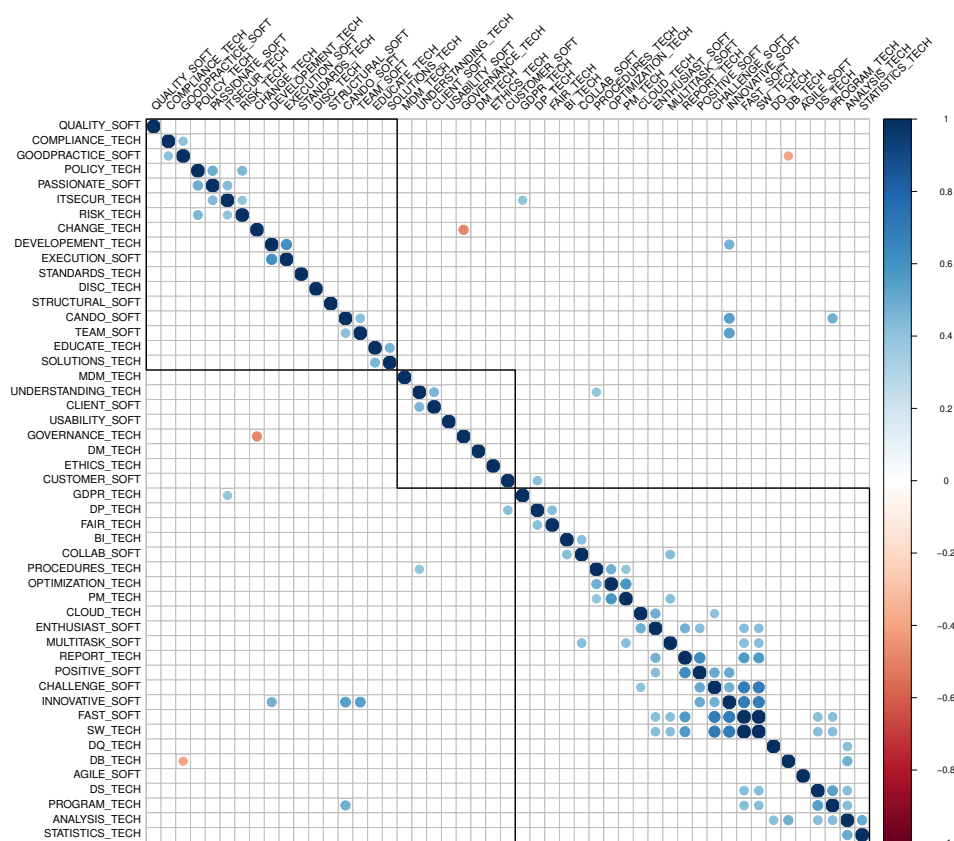


Figure 5. Correlogram with the most correlated Hard and Soft skills for those employed in the private or public sector excluding academia. See explanation of abbreviations in the Appendix.

that the participant did not select any skill, and therefore was omitted from the analysis. The list of hard and soft skills can be found in the Appendix.

Accordingly, we use our data in the following subsections to indicate the role of a DS and expectations to which tasks the DS could be engaged with in the context of different career paths and perspectives.

Not employed at university

Only the Hard skill of “product and technological expertise” was deemed not important. Figure 5 shows how different skills are correlated for DSs not employed at a university, and by reading the correlations we suggest the following roles.

- **Administrator** [first cluster], team player with can-do attitude focused on execution in strategic development, passionate about policy and IT security, establishing good practices in compliance and data privacy, thriving in risk assessments, implementing solutions and educating end-users about them.
- **Agent of Change** [second cluster], client oriented with a greater understanding of processes and operations.

- Combination of **Analyst and Developer** [third cluster], fast learner and innovative on building custom software and databases, performing data analysis and ensuring data quality, enthusiast of cloud solutions, seeking challenges with a positive attitude towards reporting, capable of optimizing processes via good project management, expert in FAIR principles and data planning, attaching great importance to collaboration and knowledge sharing to raise business awareness.

Employed at university

To make our results even more relevant for universities wanting to create a curriculum for DSs, we grouped the answers to faculty level. The way the grouping was made provides sufficient numbers of participants in each group to identify strong faculty patterns in skill clusters.

Working in Engineering, Natural Sciences and Health Sciences

The Hard skills of “product and technological expertise” and “business awareness [domain specific knowledge]” was deemed of no importance. Reading through Figure 6, we suggest the following roles.

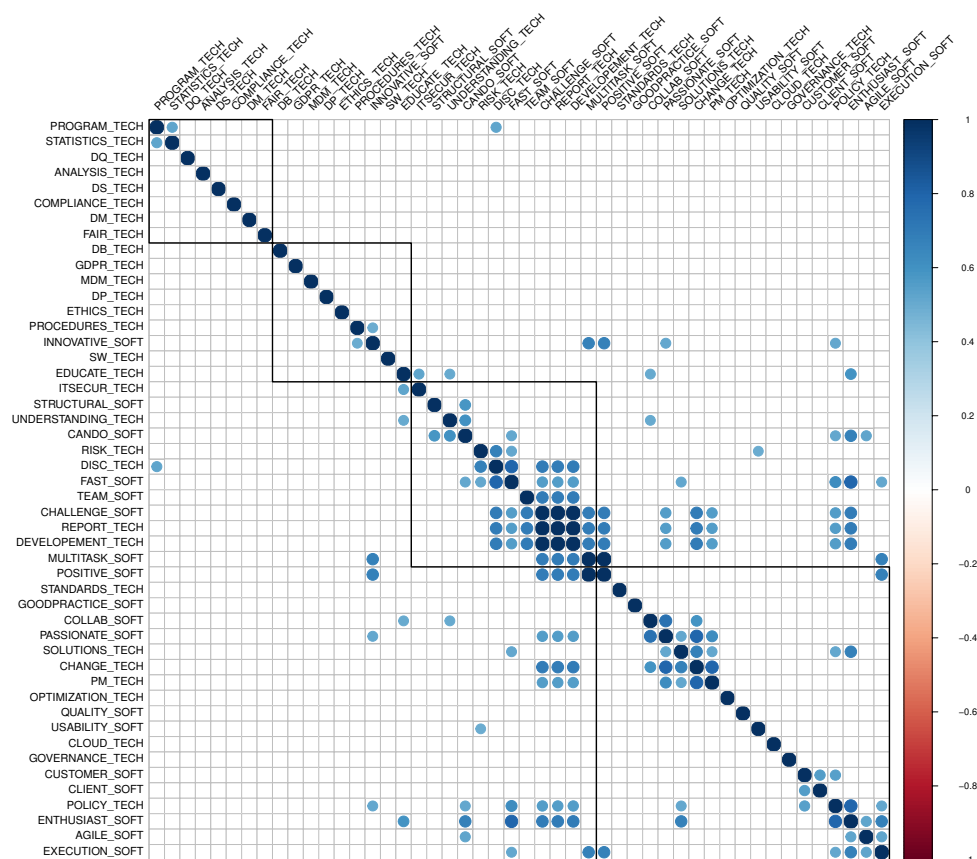


Figure 6. Correlogram with the most correlated Hard and Soft skills for those employed in academia [NAT, TEK and SUND]. See explanation of abbreviations in the Appendix.

- **Analyst** (first cluster), with good programming skills in statistical analysis.
- **Developer** (second cluster), innovative thinker who develops procedures and guidelines.
- **Administrator** (third cluster), a fast learner with a structured and analytical mindset and a can-do attitude towards processes and operations, able to make risk assessments while having disciplinary knowledge, good at multitasking within a team that makes reports and seeks challenges for strategic development.
- **Agent of Change** (fourth cluster), passionate to implement solutions via project and change management, customer oriented with a deep understanding of users, and focused on execution regarding policy and strategy awareness, approaching the tasks with an agile mindset and enthusiasm.
- **Administrator** (first cluster), with a positive attitude on cloud solutions, with a structured and analytical mindset towards data science, establishing good practices and supporting the customers with matters concerning compliance and data privacy.
- Combination of **Agent of Change and Developer** (second cluster), with innovative thinking concerning master data management, seeking challenges while having disciplinary knowledge, developing user friendly procedures and guidelines, understanding the clients enough to educate them on ethics and the responsible conduct of research.
- Combination of **Analyst and Developer** (third cluster), a fast learner with expertise in statistical analysis and in building databases and data warehouses and generally good at multitasking, having business awareness [domain specific knowledge] in implementing solutions, skilled in data planning and data governance, focused on quality and execution while maintaining enthusiasm for collaborations and knowledge sharing.

Working in Humanities and Social Sciences

The Hard skills of “project management”, “strategic development”, “optimization of processes”, “change management” and “software development” were deemed of no importance. The suggested roles of a DS as illustrated at the three clusters formed in Figure 7 are the following.

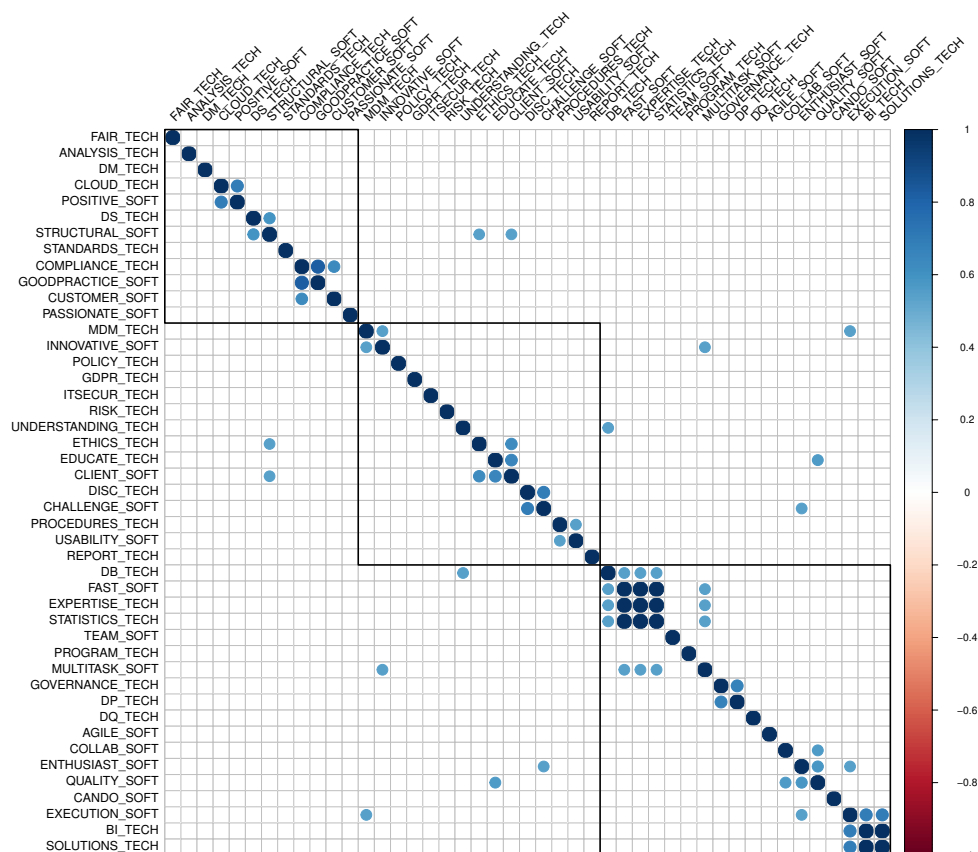


Figure 7. Correlogram with the most correlated Hard and Soft skills for those employed in academia [HUM and SAMF]. See explanation of abbreviations in the Appendix.

Working in Library and IT

The Hard skills of “project management”, “strategic development”, “reporting”, “statistical analysis”, “implement solutions” and “software development”, and the Soft skills of “focus on execution”, “multitasking”, “enthusiastic” and “positive attitude” were deemed of no importance. The clustering pattern in Figure 8 indicates the following roles.

- **Ethics Expert** (first cluster), with a can-do attitude towards ethics and the responsible conduct of research.
- **Administrator** (second cluster), fast learner, being challenged with understanding processes and procedures, and good at risk assessing cloud solutions.
- **Developer** (third cluster), passionate about optimizing processes and systems, working in a team with compliance and data privacy experts trying to establish good practices.

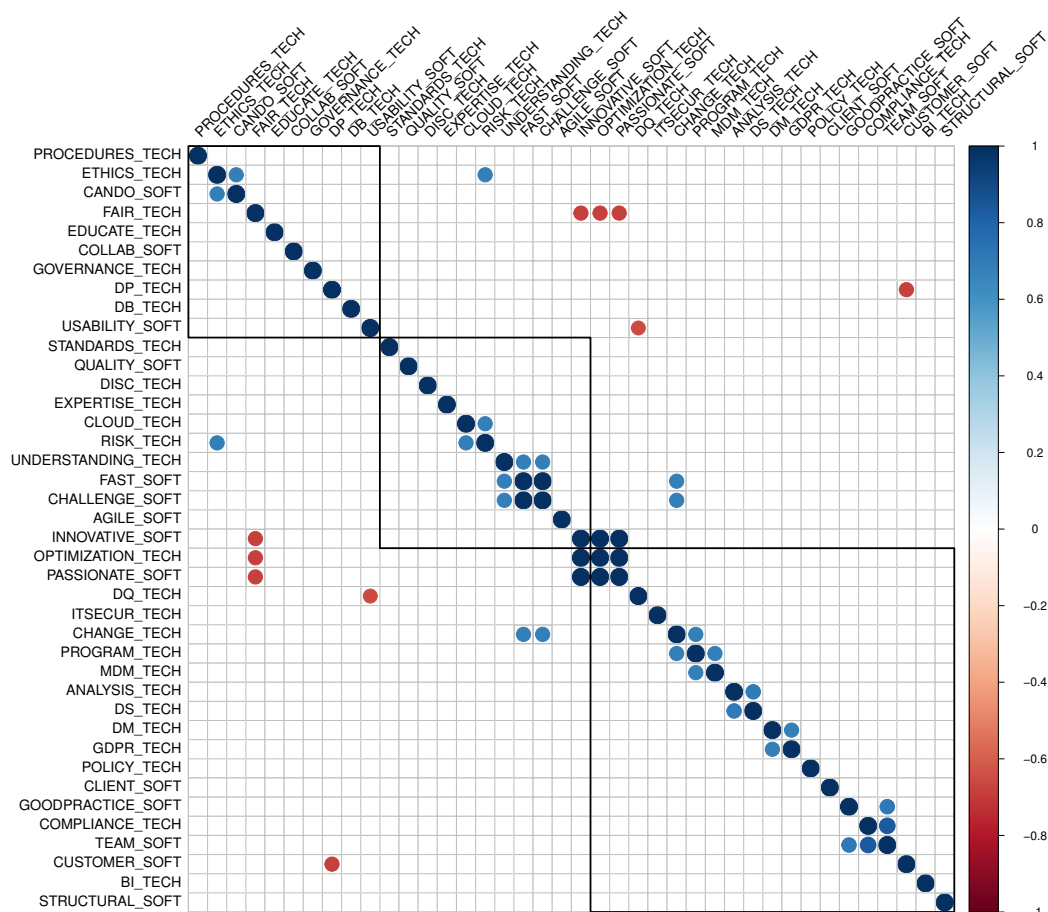


Figure 8. Correlogram with the most correlated Hard and Soft skills for those employed in academia (BIB and IT). See explanation of abbreviations in the Appendix.

6 Interviews

Interviews were held to further assess differences between academic and industry expectations to DS. The interview persons represented different places of work and were in a position to potentially hire a DS. Through the interviews, we explored the extent DSs are embedded in projects and provide generic services, and the rough classification of task areas: “administrative”, “analyst”, “developer” and “agent of change”.

We conducted four interviews with interested parties from the private and public sector, who – in one way or another – have significant experience in working with or related to data stewardship:

- research support: (**interview 1**: university administration);
- government administration (**interview 2**: public agency);
- academic research (**interview 3**: data domain specialist); and
- industry (**interview 4**: pharmaceutical company).

The interviews did not follow a structured guideline, but comprised more informal and loose conversations based around the following questions.

- Which tasks do employees in the organization perform that correspond to data stewardship?
- How are these tasks embedded within the organization?
- Which competencies, skills and levels of experience are required for this work?
- What value does data stewardship have for the organization?

A summary of each interview can be found in Zenodo. The summaries are condensed and provide a clear statement of the commonalities and differences in data stewardship across the institutions, rather than a verbatim transcription. Opinions and ideas presented in the summaries reflect the personal experiences and views of the interviewees and are not official statements of their respective institutions. The main findings from the interviews are presented in the following section.

6.1 Findings from the four interviews

None of the interviewees held the official title of a DS or had a DS employed within their own organisation. However, all could recognise tasks that overlapped with existing job

assignments in their organisation and that of the tasks and roles performed within data stewardship. These tasks included: structure and organize data, secure and access data and metadata quality, being a point of contact for data-related between producers and users of data.

The term “data steward” was not well defined or even frequently used by our interview persons. The interview persons discussed tasks that they saw could belong to the domain of a DS. These tasks the interview persons told us, are allocated to other functions and colleagues in their organisations and workflows such as a sub-task of data science, data analytics, data governance or data librarianship.

In Table 4 we illustrate the different expectations to a DS across the four interviews. When summarizing the commonalities and differences in the perception of a DS from the interviews (persons 1-4) there are some findings that support our previous observations from the other analyses regarding the four roles of data stewardship. The roles of the Administrator and the Developer were present in more interviews as tasks of data governance/data analytics and data librarianship, respectively. Roles of the Developer in combination with that of the Analyst was identified across all four interviews. Data analytics, identified in Table 4, under the role of Developer/Analyst was a key task in interview 4. Finally, the role of the Agent of Change was a needed task of being the data contact or go-to-person both within (interview 1) and out of academia (interview 4).

However, common across all interviews is the ambition that a DS in their organisations, could be the person who “provides a link between the data and the user”. The DS would be expected to make data accessible to others at the right time and for a given purpose. The DS would have the ability to identify and enhance the potential of data for application and communicate responsible exploitation of data to data-users. Accordingly, a thorough understanding of the organization and their data producing activities is crucial for the work of a potential DS, just as Kimball discussed back in 1998. Specifically for research institutions (interviews 1 & 3), a working knowledge of how research is produced, methodologies applied, and ethical handling of personal data and discipline-specific domain knowledge would be essential.

TABLE 4. Data stewardship roles and tasks identified from interviews.

Administrator	Developer/Analyst	Administrator/Developer	Agent of Change
Data Governance Ensure legal compliance Secure & assess data quality Assist decision-making	Data Analytics Design processes Automate workflows Apply new technologies	Data Librarian Structure & organize data Ensure metadata quality Sharing knowledge	Data Contact Communicate w. stakeholders Understand user needs Advise on compliance
(interview 1+4)	(interview 1+2+3+4)	(interview 2+3)	(interview 1+4)

Data stewardship activities in universities (interviews 1 & 3) has a highly interdisciplinary character and reaches across many different departments within the institution. A DS employed in universities would be assigned to participate in specific research projects and included early on in the project-planning phase. Specifically the DS would be expected to have solid knowledge and skills in securing compliance, ensuring data quality, and applying the FAIR principles to the data. Please remember here, that the findings are based on a very small sample.

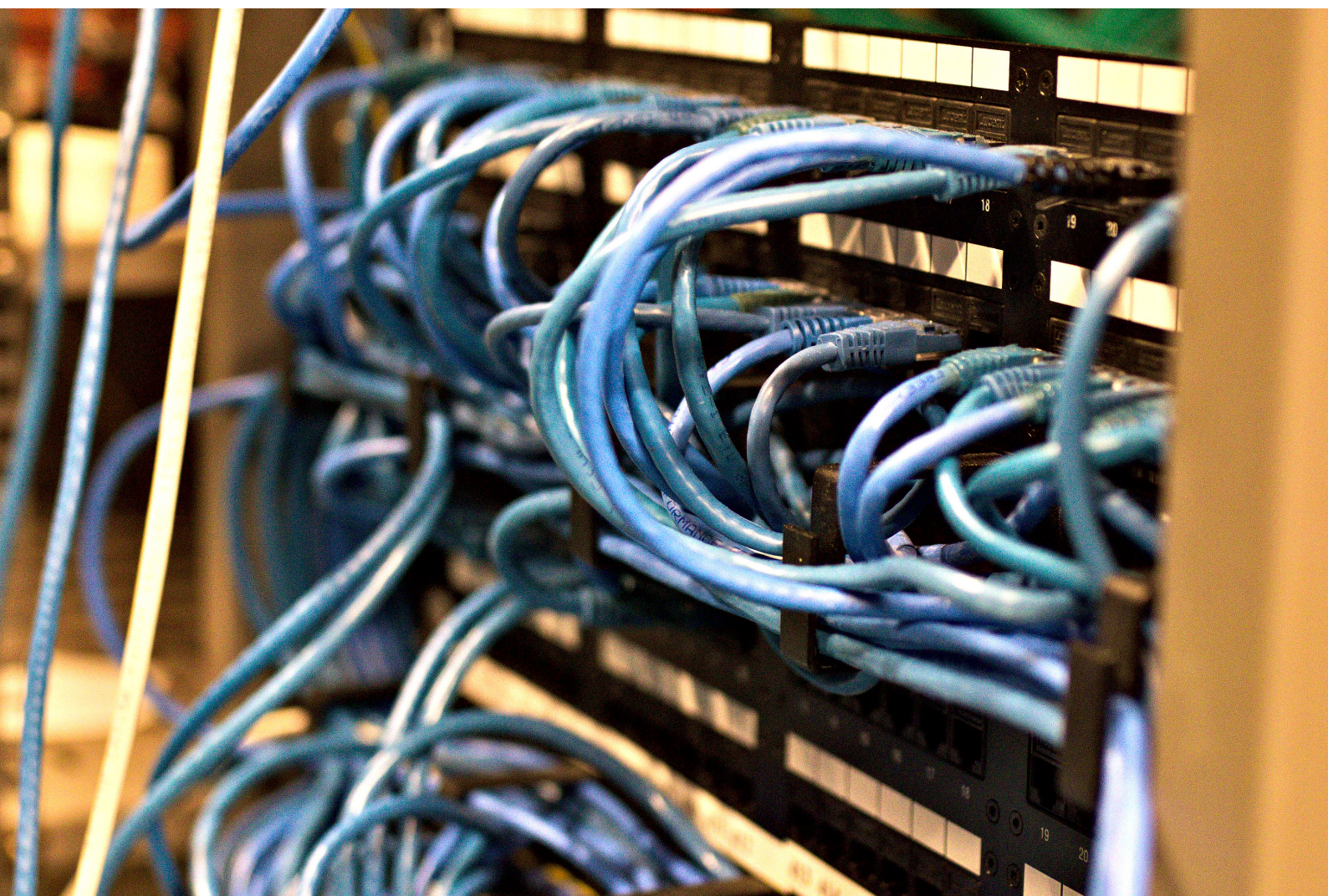
Similarly, the DS working in the corporate sector (interview 4) would also be expected to be embedded in data processes from the very beginning, to be able to master workflows and data optimization, as in the role of a e.g. customer data steward.

Our four interview persons agree that data stewardship is an area under rapid development, and is currently gaining more and more attention both in the workplace and in the media. However, all the interviewees suggested that the need for a specific DS role in their organisations is still somewhat unclear. Perhaps the unclarity is due to a lack of consensus regarding what a DS actually is. Confirming our

observation from the questionnaire, that the many different conceptualizations and understandings of a DS creates confusion and the foundational work of a DS education is to clearly define what a DS is in the curriculum.

However, the interviewees agreed that there is an increasing demand for developing competencies (upskilling) in data stewardship among existing employees to ensure good data practices and better use of data. Pointing towards the need for flexible further and continuing educational opportunities. Interestingly, the challenges and requirements for making better use of data, as discussed in the interviews, are very similar across industry, public administration and academia. This strengthens our belief that generic competences taught on a DS education would have a broad application with a diverse range of employers.

In the roles identified in Table 4, the categorisation of the areas of “administrator”, “analyst”, “developer” and “agent of change” overlap to a certain degree. If we had interviewed more people, the boundaries between the roles might have been more clearly defined.



7 Building on evidence: Models for a Data Steward education

Candidates for a DS education will have different educational and professional backgrounds and their purpose for taking a DS education will also differ. Taking this into consideration the following models for a DS education build on our findings. Together with the up-coming section “Personas”, the models illustrate the need for differently structured educations.

Conceivable candidates are:

- students with a bachelor’s degree in any field;
- students with a PhD or equivalent experience in any field;
- mature students, need flexible continuing education in DS that will enable them to improve or supplement their knowledge and skills.

7.1 Students with a Bachelor’s degree

A full-time DS education should primarily be targeted towards students with the minimum of a Bachelor’s degree and marketed primarily employability in the corporate sector. There are three programmes to consider.

A one-year master programme

Aimed at students already equipped with a bachelor, who fulfill requirements to basic programming skills, study skills, subject knowledge and academic language level. They are eager to move on from formal education and to the next stage of their career.

The one year study programme combines theory with practice. The student gains a thorough introduction to data stewardship, including skills in pedagogical, communicative and project leadership skills as well as GDPR, tools and principles for [FAIR] data management exemplified through domain specific cases and electives.

Internship should be part of the programme. It gives the student a chance to exchange knowledge and test skills and values in a business environment of their choice, with relevance for the next stage of their career.

A pre-master’s year plus one year master

Aimed at students with a bachelor from any field, who do not meet the requirement for programming and study skills. As such, this master’s course includes a pre-master’s year before the master’s to develop study skills, subject knowledge and academic language level.

This course will be suitable if:

- the undergraduate’s qualifications do not meet the level required for direct entry to the one year master;
- the undergraduates wishes to improve their programming skills, study skills, subject knowledge and academic language level before starting the master’s course;
- the student has a bachelor, but has had a break from education and wants to build up their confidence before starting the one year master.

Year 1

In the first year the student will improve their programming, communication (language) and study skills to the level needed for postgraduate study. They will also take domain specific courses to improve their disciplinary knowledge.

Year 2

The second year, the student joins the one year master programme, described above. On successful completion of the two year programme, the student will be awarded a master’s degree in data stewardship.

A two year Candidatus

Aimed at students with a Bachelor in any field. The programme is research-based and gives students a theoretical knowledge of data stewardship combined with the ability to apply this practically.

The programme aims:

- to provide the foundation for a professional career in data stewardship, including programming for data wrangling, generic data analysis methods, user studies, system and user design, and theories and knowledge of organisations;
- to provide a foundation in pedagogical, communicative and project-management skills;
- to enhance the students’ professional skills through internships, cases and disciplinary specific electives;
- to present knowledge, experience, reasoning methods and design and implementation of data optimization and storage with regards to GDPR, tools and principles for data management.

To earn the candidatus in data stewardship, the student must complete an internship and a short project investigating an aspect of data stewardship through this internship. A dissertation in the area of data stewardship is required.

7.2 Students with a PhD or equivalent research experience

Candidates aspiring to work as a Data Steward employed at a University Faculty or embedded Data Steward in research projects at universities or other research-heavy organisations require research experience, preferably a PhD.

Aimed at PhD students or Post Docs from any given field, an educational programme in data stewardship would not need to be full-time. The role of the DS provides an alternative career for the junior researcher who still wishes to work in research but is not looking for a career as a researcher.

Data stewardship education of PhD’s is a collaborative endeavour between the faculties, the Library or Knowledge Center, and other stakeholders, such as Centers for Information Security, Data Steward Community, Research coordinators, Project PI’s, System Developers, Communication and Teaching teams, the Graduate School (for PhD training), Human Research Ethics Committees, and more.

Suggested educational formats:

Short courses [also single subjects available through the Open University] workshops, mentorships and online [on demand] modules, e.g. summer school programmes and MOOCs, that immerse the PhD student in the Data Steward Community and encourage the PhD to identify and participate in DS networks, fora and educational activities.

Educational content should include, but not be limited to, research ethics, project management, policy and directives, GDPR, tools and principles for data management and FAIR data.

7.3 Continuing and professional education

Aimed at students who have professional experience and wish to improve their data stewardship skills but keep working full-time, or have already acquired a master's degree in any field. A flexible master programme is recommended. A flexible master's programme is a continuing education programme; a part-time education with a prescribed period of study and a total of 60 ECTS-points. The purpose is to allow candidates with specific needs to augment their competence within certain disciplines. The flexible master is an individually planned master's programme. The student can combine elements from various established programmes to custom-design their own flexible programme.

A flexible master programme requires that data stewardship courses, such as the one year master and two year candidatus, are available through the Open University.

7.4 A Recommended model for Data Steward education

Based on our combined findings, a full-time DS education is recommended as a two year Candidatus, with courses available through the Open University to support mature students with a professional background.

The two year Candidatus study programme is the programme most likely to increase the employability of the student. Using experiential learning and cases, the Candidatus can provide the student with theory and methods in data stewardship and one or more internships at a private company. The teaching and learning activities will need to have a degree of flexibility to support both affiliated companies internship programmes and the student's academic course work. The internship is a chance for the student to make the connection between theory and practice while still studying. It is also a chance to exchange knowledge and test skills and values in a business environment of relevance to study programme. As a result, the student will be a much more attractive employee once he or she graduates. In addition, the host company will get to know them.

An innovative combination of on-campus teaching and online learning, bootcamps and workshops is recommended, preferably developed in collaboration with local industry, departments and organisations. A flexible and professionally

oriented learning structure will potentially equip the student with theoretical and practical, communicative and analytical skills to gain an understanding of the particular domain of the company, which will be needed in case of employment there. Most of the job vacancies we reviewed during this study required at least one year of work experience on top of a degree, therefore a 2 year master degree including an internship could provide sufficient corporate understanding to the inexperienced student.

Additionally, we recommend the 1 year master programme in data stewardship [BA + 1], on the condition it is open to students with a minimum of a bachelor. Thus, the programme is open to master students, PhD's and through the Open University to mature students [V.E.T]. On the aforementioned condition, the 1 year education provides a dynamic learning environment and relevant upskilling programme that can rapidly supply DS to the hungry job market.

Our concern with the one-year master programme, is if places are primarily offered to applicants with only a bachelor degree and number of study places and access for students with other educational profiles is restricted. Based on the evidence in our study, we fear the Bachelor student will not be sufficiently equipped or mature enough to enter the job market without substantial training after completing the one-year education. We have found that working as a DS requires a maturity to engage in organisations as the Agent of Change, motivating a cultural change in data practices and workflows. Thus, demanding the graduate has project management skills and liaison and networking abilities, that build on their knowledge of the domain in which they are employed.

However, we acknowledge that a current trend in the corporate world is to hire Candidatus students before graduation. As Tim Cook from Apple said, "... about half of our U.S. employment last year were people that did not have a four-year degree. And we're very proud of that, but we want to go further..." [Remarks, 2019].

The above two models are primarily marketed employment in the corporate sector. Of course, if the participant on the BA+1 has experience in research, the possibility is there, for employment in academia. Working as a DS in academic research or other research-heavy environments is a special case, that requires the candidate has a PhD or equivalent experience as well an understanding of how research operates and how data and software underpins reproducible research. As exemplified at TU Delft, this understanding is a necessary prerequisite to be able to act as a spokesperson for a faculty, advise researchers and develop the faculty's data management policies and advancements and tailoring training to fit the researcher's needs.

There is an oversupply of PhDs. Although a doctorate is designed as training for a job in academia, the number of PhD positions is unrelated to the number of job openings. Me-

anwhile, business leaders complain about shortage of high-level skills, suggesting PhD's are not taught the right things for a career in the corporate world (Smaglik, 2014). Training as a DS provides PhDs and postdocs with alternative job possibilities and career paths in academia and research. If, however, the tasks needed to be carried out by a research team, university department or faculty do not require the above level of domain expertise, then a DS is probably not what they are looking for, but rather someone with a bachelor's or master's degree in computer science or equivalent to ensure data optimization, data crunching and quality assurance, etc. Here again, the short one-year master programme in DS would be-

nefit the computer scientist to qualify for both the technical and the mediary role of a DS in the organisation.

7.5 Personas

The personas add a human touch to our recommendations and findings. They illustrate how the collected findings from this report, especially the four roles of a DS (Administrator, Analyst, Developer and Agent of Change) identified in the correlation analyses, could be operationalised in an education, and represent how different student types that might use the DS educational programmes.



Zander V. Carlsen*

“Conversations around data stewardship are happening at my workplace, but must be given more urgency.”

Age: 40.

Family: Father of 3, married.

Private: Enjoys building robots in his garage with his three sons. Volunteers at the local library at a programming club for children aged 7-14.

Education: Bachelor degree as Biomedical equipment technician from his home country of England in 2001.

Personality: Curious - Extrovert - A natural teacher.

DS roles Zander already can fill:

Analyst	✓	Developer	✓
Administrator	✗	Agent of Change	✗

Frustrations: Zander's expertise is in instrument and equipment making. In his current position he produces one-off, bespoke equipment for medical research. In his job as a Lab Technician, Zander assists researchers in developing innovative equipment using his knowledge and experience of instrumentation developments. He feels he can optimize research workflows in his organisation by combining his ability as a teacher with his software and innovation skills. However, he admits he doesn't know enough about policies and the administrative side of data management to confidently act in an advisory position for researchers or intermediary position between researchers and the organisation.

Goals:

- Change culture in data management.
- Ensure compliance.
- Provide a single point of contact.
- Share knowledge and experience.

Ambition: As a busy family man with financial responsibilities, he only has the economic freedom to take a short-term education. He is looking for an education that fits with his family and work-life and is tempted by **the one year master programme in data stewardship** offering a deeper knowledge on management of data and policy compliance. His employer has agreed to give him up to one years study leave and he is eligible for benefits from his A-Kasse.

Jesper L. Andersen*

"I read an article that all the major companies in Greater Copenhagen have a need for data stewards."

Age: 21.

Family: Single.

Private: Plays handball in a local team and has run his first Nordic Race (obstacle course racing).

Education: Attending his fourth semester at a bachelor (BSc) in data science at IT University of Copenhagen.

Personality: Structured - Outgoing - Keeps things simple - Fast learner.

DS roles Jesper already can fill:

Analyst



Developer



Administrator



Agent of Change



Frustrations: Even though Jesper's degree equips him with skills and knowledge that are in high demand across both the private and the public sector, the competition is increasing dramatically. Jesper wants to differentiate his profile and focus exclusively on business intelligence in the financial sector.

Goals:

- Be able to ensure data quality.
- Become skilled on building custom software and databases.
- Enhance his programming skills in statistical and data analysis.

Ambition: Jesper wants to be able to assess the quality of shared enterprise datasets, ensure accurate and consistent data across the enterprise and understand the business value of data. He would like to grab that chance and become a DS himself, so he is already considering to build on top of his bachelor with a **two year full-time master degree in data stewardship (Candidatus)**. Currently, he is trying to get a data stewardship internship at one of these companies that would assist him greatly in clarifying his future steps.

Lillian Thomsen*

Age: 47.

Family: Mother of 2, married.

Private: She volunteers in the National Society for Women's Crisis Centre (Landsforening for Kvindekrisecentre) and loves gardening.

Education: Master degree in Library & Information Science from the Royal School of Library & Information Science in Copenhagen in 2005.

Personality: Creative - Loves challenges - Communicative - Team player.

DS roles Lillian already can fill:

Analyst



Developer



Administrator



Agent of Change



Frustrations: Lillian would like to gain more competencies and skills in the domain of data stewardship, but she does not want to pause her career to take a full-time masters degree. Her boss sees the return on investment in supporting Lillian in her further education.

Goals: Being already tech-savvy, she would like to focus more on:

- Data privacy and protection.
- Laws concerning GDPR compliance.
- The FAIR principles.
- Copyright, licenses and ownership.
- Develop procedures and guidelines.
- Responsible conduct of research.

Ambition: Lillian would like to enrol in education by attending the Continuing and professional programme. Luckily for her, there is such a programme at her local university, where as a part-time master student she can select courses equal to 60 ECTS over 3 years spanning across different faculties and departments.

Nancy Rutherford*

"I have set my life in Denmark, I don't want to pursue a post-doc abroad, I love research, but I also know that only 0.45% of the PhDs manage to become Professors."

Age: 29.

Family: Mother of 1, single.

Private: Plays the piano and sings in the Odense Gospel Choir.

Education: Currently finishing her PhD in physics specialized in ambient aerosols from the Centre for Cosmology and Particle Physics Phenomenology at University of Southern Denmark.

Personality: Positive attitude - Solutions oriented - Analytical mindset - Good at multitasking.

DS roles Lillian already can fill:

Analyst	✓	Developer	✗
Administrator	✗	Agent of Change	✗

Frustrations: Nancy has a deep disciplinary knowledge of her domain and she is very good in advanced analytical and statistical methods as well as in programming and in creating prediction models. She would like to use all her knowledge and skills to help young researchers create better datasets in her department, respecting the FAIR principles and the code of conduct for research integrity, as well as help senior researchers in writing better grant applications by focusing on Open Science practices.

Goals:

- Focus on project and change management.
- Get better at risk assessments, writing reports and guidelines.
- Enhance her knowledge on user experience and usability.
- Focus on execution and strategy awareness.

Ambition: The head of her department is willing to hire her as a DS provided that she enriches her knowledge with **tutorials, attendance in DS field specific conferences, participation in relevant workshops, and PhD courses.**

8 Recommendations

Combining and interpreting the results

In order to provide informed input and point towards a possible prequalification of a DS education in Denmark we used a combination of quantitative and qualitative methods including literature search, interviews, text mining, content analysis and topic modelling. We:

- reviewed 24 DS national and international educational programmes;
- scraped the profiles of 74 Danish DSs on LinkedIn;
- analysed 119 national and international DS (or equivalent) job vacancies;
- created and forwarded a national questionnaire to 350 stakeholders from Danish industry, public sector and universities; and
- conducted 4 interviews with stakeholders from the Danish public and private sector.

Based on our combined findings from the above analyses we recommend any DS education to support the four roles of the DS, identified as the Administrator, the Analyst, the Developer and the Agent of Change, preferably in an educational model that allows for experiential learning, collaboration with industry and innovative pedagogies that will enable a flexible education that fits with problem-based learning strategies and increase the employability of the graduates. In summary, the interpretation of our results combining all our findings can be illustrated by the four roles of a DS and the three models for data stewardship education in Figure 9 below.

The three roles of the Administrator, the Analyst and the Developer we identified from the combination of the LinkedIn profiles analysis and the job vacancies analysis, are aligned - and in the same order - with the outcome of the Dutch report "Towards a community-endorsed data steward profession description from life science research" [Scholtens et al., 2019] authored by the Dutch Techcentre for Life Sciences in collaboration with the universities of Utrecht, Radboud and Groningen. The report concluded that there are three implementation areas for a DS: Policy; Research; and Infrastructure. During the same year, another report from the Netherlands was published from the National Coordination Point Research Data Management task group [Verheul et al., 2019] where they identified three task areas of DSs: Embedded and operational; Generic and advisory; Policy, strategy and coordination. Worth noting is that the report by the Dutch Techcenter for Life Sciences analysed 40 international job descriptions in the life sciences domain, and the latter only 22 DS vacancies within the Netherlands.

Due to the fact that we combined many more inputs in our report, we discovered an indication for the existence of a fourth role which we called Agent of Change. This role emerged when analysing the questionnaire results in Section 5.1.2.

Afterword - next steps

This project was commissioned to address the question if DS education should be developed in a national coordination between universities and industry or if competition between these parties results in a healthy competition, resulting in diverse educational programmes appealing to different audiences. The need for educating and raising awareness of the key role of data stewardship is evolving from stakeholders both within and beyond a national setting. However, that level of awareness along with the increasing demand for DS competencies raises a question of how to agree on a healthy balance between natural competition and efforts of coordination in this emerging field.

As reviewed in Section 2, internationally different approaches to DS certification already exist within industry whilst formal educations at universities are still in the very early-stages. However, several international efforts of coordination across academic stakeholders do exist, such as the following.

- The GO-FAIR implementation network for Data Stewardship Competence Centers (DSCC) aiming for networking and practical cooperation, reuse of assets and convergence between national and international DSCCs at universities or other research performing institutions.
- The National Coordination Point Research Data Management (LCRDM) in the Netherlands to coordinate the professionalization of data stewardship [Verheul et al., 2019]. Aiming at establishing practical cooperation and building networks, and convergence between national and international initiatives around DSCCs at universities.
- The H2020-funded project FAIRsFAIR for professionalise data science and stewardship. With recognized higher education institutions as project partners, it seeks to coordinate the availability of training for data skills and stewardship. The aim is to propose a FAIR competence framework that can be translated into university education via curricula and courses for discipline-specific professional profiles, like data stewards [FAIRsFAIR, 2019a].
- The organisation around the European Open Science Cloud (EOSC) consists of different working groups who are coordinated by the EOSC secretariat. Together, they point towards specific skills and competences needed for both operationalizing federated infrastructure and curating FAIR data objects for interoperability in the cloud. Thus, future Data Stewardship training and education should also take notice of such needs for skills and competencies in academic environments. A new EOSC working group on skills and training will be established in 2020].

In the Autumn of 2019, we were invited to share our findings and knowledge of data stewardship to a EUA focus group meeting [EUA/FAIRsFAIR, 2019a] held at the University of Amsterdam [EUA/FAIRsFAIR, 2019]. The topic was on teaching [FAIR] data management and data stewardship. The focus group concluded, there is a lack of coordina-

tion regarding all the efforts, initiatives and training materials on a national, European and international level. Lack of coordination hampers the identification of the best resources to use and implement within university curricula for data stewardship [FAIRsFAIR, 2019b]. Accordingly, we also emphasize the importance for coordination of DS education in Denmark adapted geographically to fulfill the demand for DS in Denmark and adapted to both generic and discipline-specific competence frameworks for DS.

Therefore, we strongly recommend the next step is to agree on a common national approach to university education in data stewardship and engage with the above mentioned initiatives in a coordinated manner to ensure job mobility and the free movement of labour in agreement with the Bologna Process and the European HE Area [Bonjean, 2018].

We consider the development of DS education to be a shared responsibility of stakeholders in the public and private

R&E sectors, including joint networks dedicated to working with transferable skills and digital competencies. The governance of data stewardship coordination in Denmark should point to educational committees such as RUVU [The Advisory Committee for Evaluation of Provision of Higher Education] which is part of the Ministry of Higher Education and Science and education committees at Universities Denmark who are expected to have the responsibility for coordinating the different levels of education [Bachelor, Master, Candidatus and continuing and professional education and training].

We expect the majority of candidates will be employed in the corporate sector, but tasks related to good data stewardship are steadily increasing in international academia in the coming years. Documentation in the survey shows that almost 40% of the respondents will be willing to hire DS in the future [See Survey report in the Zenodo Community].

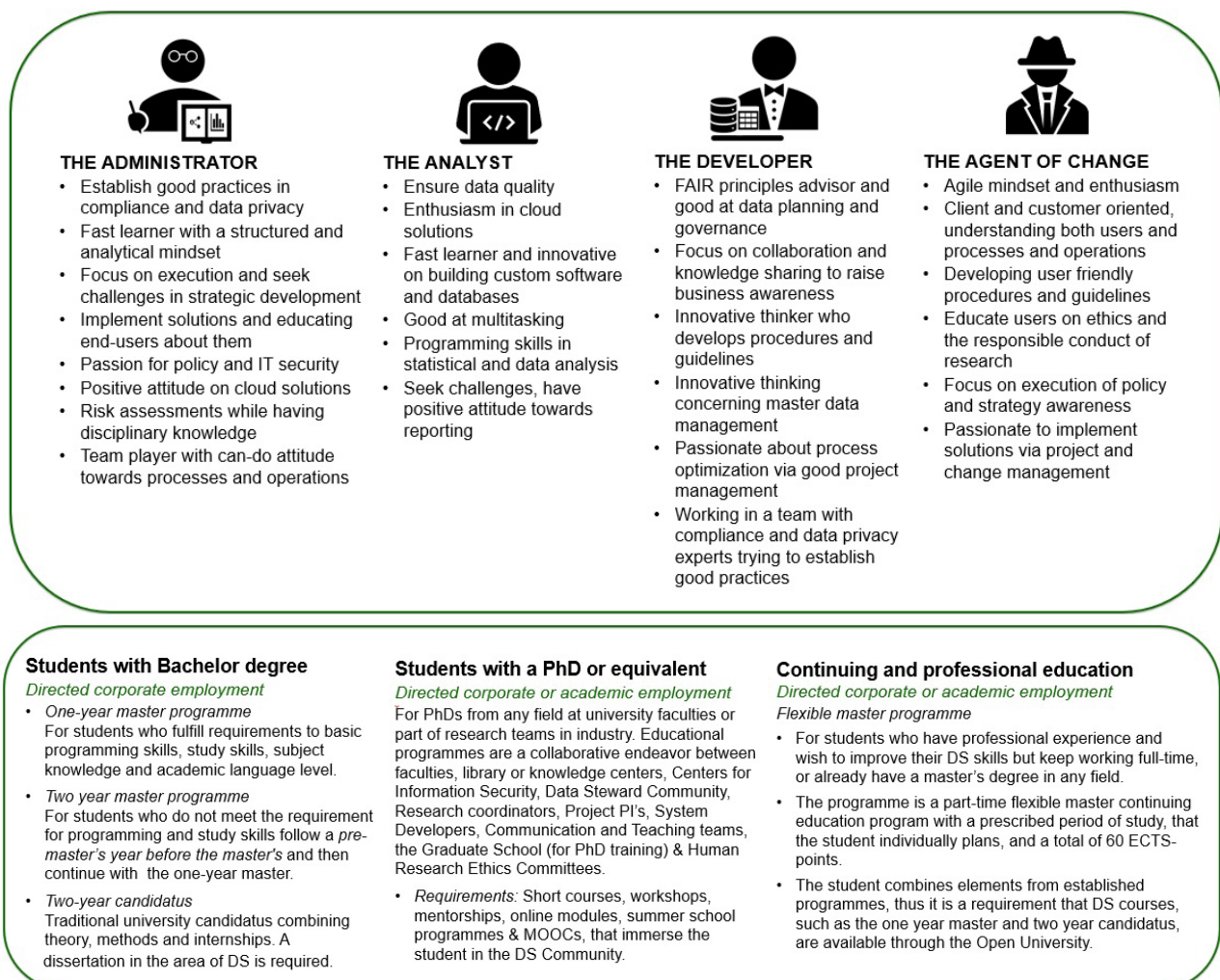


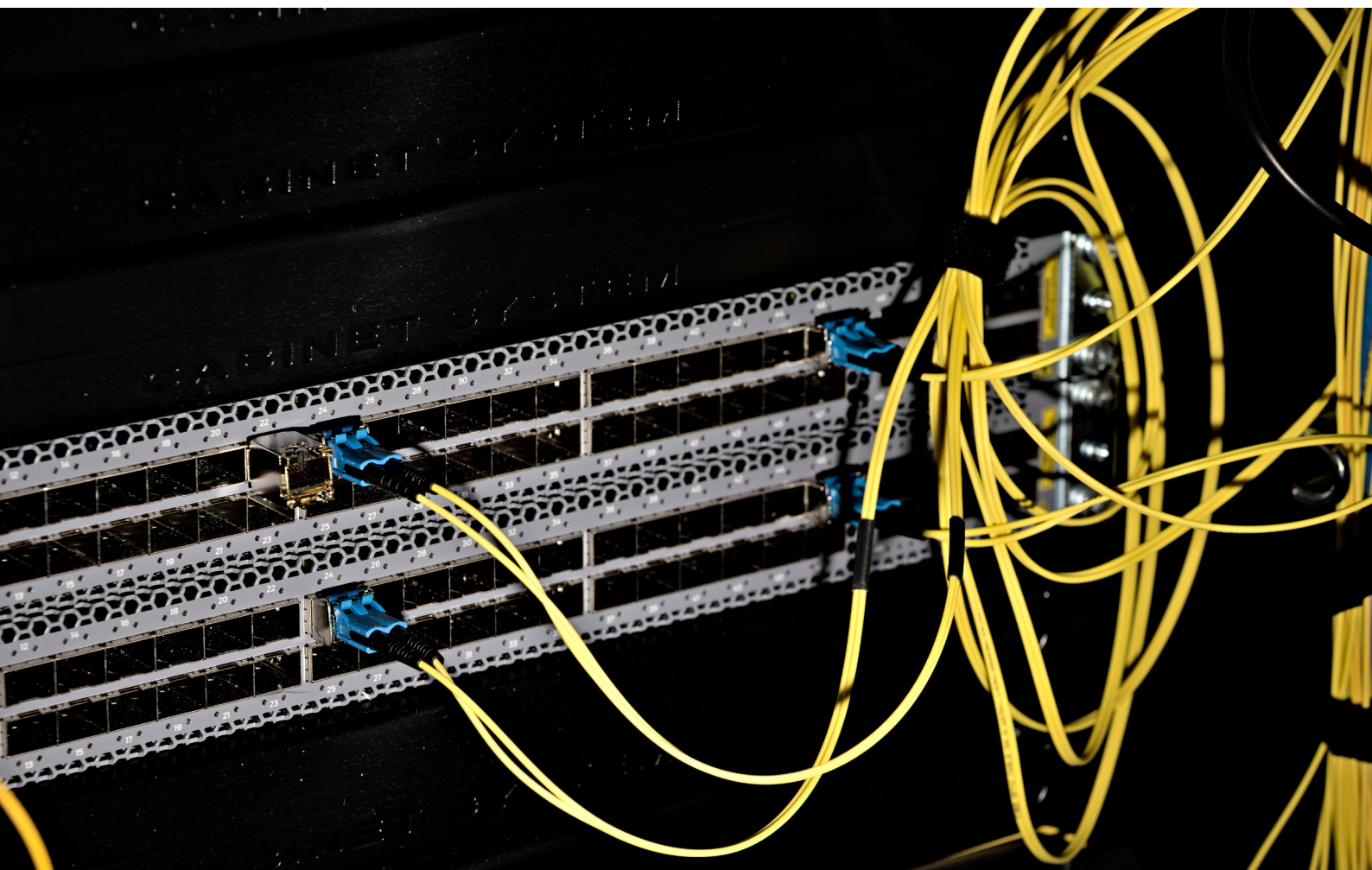
Figure 9. The 4 roles of a Data Steward and the 3 models for a data stewardship education.

Next steps towards data steward education and training in Denmark:

- Include private and public sectors in the development of education. A diverse stakeholder group will address different positions, attitudes, needs and expertise across a wide variety of data roles in a wide variety of organisations.
- Establish a Data Steward network to increase future involvement in DS training and share learning outcomes and experiences between institutions across all relevant sectors.
- Develop an adequate number of educations (to begin with one in the east and one in the west of Denmark) with respect to scalability of the demand-side, as data stewardship is still in the process of taking shape and not well-established in the public and private sector in Denmark.
- Avoid a plethora of indistinguishable educational programmes by refusing to compete. Seek to coordinate the educational perspective of data stewardship across advisory boards in the R&E sector to the benefit of students and employers.

- Give students the opportunity for 2 years of study in order to combine his or her DS education with internships and the analytical skills to apply data stewardship learnings to their coming working place.
- Develop flexible education and use innovative learning strategies such as cases, on-demand learning, boot-camps and experiential learning, to allow students with different academic profiles to combine internships and direct experiences with focused reflection.

In this DM Forum initiated project, we attempted to gather evidence to support a prequalification for a national DS education. During the process we involved stakeholders from various Danish organisations, companies and institutions (see other communications from the project). Despite the limitations of our applied methods, such as the small sample size, the under-representation of universities in the sample and admittedly, the short project period, we feel we have succeeded in delivering informed input to a possible future education in Data Stewardship. It's now up to universities and other educational institutions to be exemplary, and offer leading and innovative DS programmes.



Other communications from the project

- [Conference & presentations, University of Copenhagen, Oct 4, 2019]: Data Steward Education: Coordination or competition?
- [Interview, In the Field Stories, Oct 14, 2019]: Data Stewards to the rescue, please!
- [Video presentation, DeIC 2019 Conference, Oct 31]: A data what? Preliminary investigations into the role and qualifications of Data Stewards
- [Interview, Perspektiv, no. 11, December 2019]: Data Steward - uddannelse i overblik
- [Presentation, Teaching [FAIR] data management and stewardship, EUA/FAIRsFAIR Focus Group, University of Amsterdam, Nov 19]: National Coordination of Data Steward Education in Denmark
- [Poster & abstract, 15th IDCC 2020, Dublin, Feb 18-21 2020]: Reframing Data Stewardship Education in Denmark . Zenodo: <https://doi.org/10.5281/zenodo.3629191> [embargoed until Feb 21, 2020]

References

For a more comprehensive bibliography compiled in the project, we refer to: https://www.zotero.org/groups/2275886/datastewardship_dm_forum/items

- Bouchet-Valat, M. [2014]. "SnowballC: Snowball stemmers based on the C libstemmer UTF-8 library", <https://cran.rproject.org/web/packages/SnowballC/index.html>
- Bonjean, D. [2018]. The Bologna Process and the European Higher Education Area. Education and Training - European Commission. https://ec.europa.eu/education/policies/higher-education/bologna-process-and-european-higher-education-area_en
- Börner, K., Scrivner, O., Gallant, M., Ma, S., Liu, X., Chewning, K., Wu, L., & Evans, J.A. [2018] Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy. Proceedings of the National Academy of Sciences, 115 [50] 12630-12637; <https://doi.org/10.1073/pnas.1804247115>
- DM Forum [DeIC]. [2019]. Data Steward education - coordination or competition? Danish e-Infrastructure Cooperation. Retrieved 22 Jan 2019, from <https://www.deic.dk/da/datamanagement/data-steward-event>
- EUA/FAIRsFAIR. [2019]. Teaching [FAIR] data management and stewardship. <https://www.eua.eu/events/106-teaching-fair-data-management-and-stewardship.html>
- European Commission, & Directorate-General for Research and Innovation [2016]. Realising the European open science cloud: First report and recommendations of the Commission high level expert group on the European open science cloud. Publications Office. <https://doi.org/10.2777/940154>
- FAIRsFAIR. [2019a]. FAIR Data Science and Professionalisation - WP7 | FAIRsFAIR. <https://www.fairsfair.eu/fair-data-science-and-professionalisation>
- FAIRsFAIR. [2019b]. FAIRsFAIR Focus Group: University of Amsterdam. <https://www.fairsfair.eu/articles-publications/fairsfair-focus-group-university-amsterdam>
- Feinerer, I. & Hornik, K. [2015]. "tm: Text Mining Package. R package version 0.6-2", <http://CRAN.R-project.org/package=tm>
- GO FAIR: Data Stewardship Competence Centers [DSCC]. GO FAIR. Retrieved Jan 21 2020, from: <https://www.go-fair.org/implementation-networks/overview/dscc/>
- Hennig, C. [2015]. "fpc: Flexible Procedures for Clustering", <https://cran.rproject.org/web/packages/fpc/index.html>
- Kim, Y.M. [2009]. "Validation of psychometric research instruments: the case of information science". Journal of the American Society for Information Science and Technology, Vol. 60 No. 6, pp. 1178-1191.
- Kimball, R. (Ed.). [1998]. The data warehouse lifecycle toolkit: Expert methods for designing, developing, and deploying data warehouses. Wiley. ISBN: 978-0-471-25547-5
- Kurapati, S. [2019]. Becoming a data steward. Retrieved 16 July 2019, from Impact of Social Sciences website: <https://blogs.lse.ac.uk/impactofsocialsciences/2019/04/30/becoming-a-data-steward/>
- Maechler, M. et al. [2016]. "cluster: Cluster Analysis Basics and Extensions. R package version 2.0.5.", <https://cran.rproject.org/web/packages/cluster/index.html>
- National Coordination Point Research Data Management [LCRDM]. Landelijk Coördinatiepunt Research Data Management [LCRDM]. Retrieved Jan 21 2020, from: <https://www.lcrdm.nl/en/homepage>
- Remarks - by President Trump at an American Workforce Policy Advisory Board Meeting [2019]. Retrieved Jan 21 2020, from <https://www.whitehouse.gov/briefings-statements/remarks-president-trump-american-workforce-policy-advisory-board-meeting/>
- Salome Scholtens, Mijke Jetten, Jasmin Böhmer, Christine Staiger, Inge Slouwerhof, Marije van der Geest, & Celia W.G. van Gelder. [2019, October 3]. Final report: Towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise. Zenodo. <http://doi.org/10.5281/zenodo.3474789>
- Sempreviva, A. M. [2019]. Introduction to the Course Data Science essentials. Part 1 and Part 2. From Open Data to Open Innovation. <https://doi.org/10.5281/zenodo.3580533>
- Smaglik, P. [2014]. Employment: PhD overdrive. Nature 511, 255-256. <https://doi.org/10.1038/nj7508-255a>
- Taylor, M., Martin, B., & Wilsdon, J. [2010]. The scientific century: securing our future prosperity. The Royal Society. https://royalsociety.org/~media/Royal_Society_Content/policy/publications/2010/4294970126.pdf
- Wilkinson, M. D., et al. [2016]. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Verheul, I., Imming, M., Ringerma, J., Mordant, A., Ploeg, J.-L. van der, & Pronk, M. [2019]. Data Stewardship on the map: A study of tasks and roles in Dutch research institutes. Zenodo. <https://doi.org/10.5281/zenodo.2669150>
- Versweyveld, L. [2016]. We need 500.000 respected data stewards to operate the European Open Science Cloud. Retrieved 22 Jan 2019, from <http://e-irg.eu/news-blog/-/blogs/we-need-500-000-respected-data-stewards-to-operate-the-european-open-science-cloud>
- Vlachos, E., & Tan, Z. H. [2018]. Public perception of android robots: Indications from an analysis of YouTube comments. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) [pp. 1255-1260]. IEEE. <https://doi.org/10.1109/IROS.2018.8594058>
- Vlachos E. & Wildgaard, L. et al. [2020] Data Files and Documentation - National Coordination of Data Steward Education in Denmark. Zenodo: <https://doi.org/10.5081/zenodo.3634756>
- Vollertsen, A. [2019]. Data Stewards to the rescue, please!, In The Field. Retrieved Jan 21 2020, from <https://www.inthefieldstories.net/data-stewards-to-the-rescue-please/>
- Wildgaard, L [2020]: Reframing data steward education in Denmark and abroad, Zenodo: <https://doi.org/10.5281/zenodo.3628375>

List of Abbreviations and Acronyms

DS	Data Steward(s)
DSCC	Data Stewardship Competence Centers
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Reuseable
GDPR	General Data Protection Regulation
PSI	Public Sector Information
R&D	Research & Development
R&E	Research & Education

List of Tables and Figures

TABLE 1.	Organisations employing data stewards and the titles applied to data steward function.
TABLE 2.	Suggested roles of a data stewards and identified topics, common in the combined corpora of Job Vacancies, LinkedIn Hard Skills and LinkedIn Soft Skills.
TABLE 3.	Titles given to the role of a DS
TABLE 4.	Data Stewardship roles and tasks identified from the interviews
Figure 1.	Data Stewardship skills.
Figure 2.	(a) Pareto chart of the Hard Skills (a) and Soft Skills (b) from the LinkedIn profiles in descending order of frequency with a cumulative line as a percentage of the total.
Figure 3.	Pareto chart of the terms from the job vacancies analysis in descending order of frequency with a cumulative line as a percentage of the total.
Figure 4.	Are DS (or equivalent) a specific function in your organisation?
Figure 5.	Correlogram with the most correlated Hard and Soft skills for those employed in the private or public sector excluding academia.
Figure 6.	Correlogram with the most correlated Hard and Soft skills for those employed in academia (NAT, TEK and SUND).
Figure 7.	Correlogram with the most correlated Hard and Soft skills for those employed in academia (HUM and SAMF).
Figure 8.	Correlogram with the most correlated Hard and Soft skills for those employed in academia (BIB and IT).
Figure 9.	The 4 roles of a Data Steward and the 3 models for a Data Stewardship Education.

Appendix

Table of Abbreviations of Hard and Soft Skills from the questionnaire

Skills as asked in the questionnaire	Abbreviations
Group A - Tech- Builds custom software and apps	SW_TECH
Group A - Tech Builds databases and data warehouses	DB_TECH
Group A - Technical and didactical skills - Business awareness (domain specific knowledge)	BI_TECH
Group A - Technical and didactical skills - Change management	CHANGE_TECH
Group A - Technical and didactical skills - Cloud solutions	CLOUD_TECH
Group A - Technical and didactical skills - Compliance and data privacy	COMPLIANCE_TECH
Group A - Technical and didactical skills - Data analysis	ANALYSIS_TECH
Group A - Technical and didactical skills - Data governance	GOVERNANCE_TECH
Group A - Technical and didactical skills - Data management	DM_TECH
Group A - Technical and didactical skills - Data planning	DP_TECH
Group A - Technical and didactical skills - Data quality	DQ_TECH
Group A - Technical and didactical skills - Data science	DS_TECH
Group A - Technical and didactical skills - Data standards	STANDARDS_TECH
Group A - Technical and didactical skills - Develop procedures and guidelines	PROCEDURES_TECH
Group A - Technical and didactical skills - Disciplinary knowledge	DISC_TECH
Group A - Technical and didactical skills - Educating end-users and pedagogical understanding	EDUCATE_TECH
Group A - Technical and didactical skills - Ethics and responsible conduct	ETHICS_TECH
Group A - Technical and didactical skills - FAIR expertise	FAIR_TECH
Group A - Technical and didactical skills - GDPR	GDPR_TECH
Group A - Technical and didactical skills - Implement solutions	SOLUTIONS_TECH
Group A - Technical and didactical skills - IT security	ITSECUR_TECH
Group A - Technical and didactical skills - Master data management	MDM_TECH
Group A - Technical and didactical skills - Optimization [of processes, system ...]	OPTIMIZATION_TECH
Group A - Technical and didactical skills - Policy and strategy awareness	POLICY_TECH
Group A - Technical and didactical skills - Programming skills	PROGRAM_TECH
Group A - Technical and didactical skills - Product/technological expertise	EXPERTISE_TECH
Group A - Technical and didactical skills - Reporting	REPORT_TECH
Group A - Technical and didactical skills - Risk assessment	RISK_TECH
Group A - Technical and didactical skills - Project management	PM_TECH
Group A - Technical and didactical skills - Strategic development	DEVELOPEMENT_TECH
Group A - Technical and didactical skills - Statistical analysis	STATISTICS_TECH
Group A - Technical and didactical skills - Understanding of processes and operations	UNDERSTANDING_TECH
Group B - Soft and communication skills - Agile mindset	AGILE_SOFT
Group B - Soft and communication skills - Attach great importance to collaboration and knowledge sharing	COLLAB_SOFT
Group B - Soft and communication skills - Can-do attitude	CANDO_SOFT
Group B - Soft and communication skills - Customer oriented/support	CUSTOMER_SOFT
Group B - Soft and communication skills - Dedicated and passionate	PASSIONATE_SOFT
Group B - Soft and communication skills - Enthusiastic	ENTHUSIAST_SOFT
Group B - Soft and communication skills - Establish good practices	GOODPRACTICE_SOFT
Group B - Soft and communication skills - Fast learner	FAST_SOFT
Group B - Soft and communication skills - Focus on execution	EXECUTION_SOFT
Group B - Soft and communication skills - Focus on quality	QUALITY_SOFT
Group B - Soft and communication skills - Focus on usability	USABILITY_SOFT
Group B - Soft and communication skills - Innovative thinker	INNOVATIVE_SOFT
Group B - Soft and communication skills - Multitasking	MULTITASK_SOFT
Group B - Soft and communication skills - Outgoing team player	TEAM_SOFT
Group B - Soft and communication skills - Positive attitude	POSITIVE_SOFT
Group B - Soft and communication skills - Seek challenges	CHALLENGE_SOFT
Group B - Soft and communication skills - Structured and analytical mindset	STRUCTURAL_SOFT
Group B - Soft and communication skills - Understanding of clients	CLIENT_SOFT

DM FORUM FINAL REPORT 2020

The project has been financially supported by the National Forum for Research Data Management (DM Forum) as part of DeIC (Danish E-Infrastructure Cooperation)

DeiC
DTU
Asmussens Alle
Bygning 305
2800 Kgs. Lyngby
Danmark

t	35 88 82 02
m	sekretariat@deic.dk
w	www.deic.dk
cvr	30 06 09 46
ean	5798000430723