Uddannelses- og Forskningsministeriet

Prækvalifikation af videregående uddannelser - Data science

Udskrevet 13. november 2026

Bachelor - Data science - IT Universitetet i København

Institutionsnavn: IT Universitetet i København
Indsendt: 03/10-2016 08:05
Ansøgningsrunde: 2016-2
Status på ansøgning: Godkendt
Afgørelsesbilag
Samlet godkendelsesbrev
Download den samlede ansøgning
Læs hele ansøgningen

Ansøgningstype
Ny uddannelse

Udbudssted
København

Kontaktperson for ansøgningen på uddannelsesinstitutionen
Mette Holm Smith meho@itu.dk 72185087

Er institutionen institutionsakkrediteret?
Betinget

Er der tidligere søgt om godkendelse af uddannelsen eller udbuddet?
Nej

Uddannelsestype
Bachelor

Uddannelsens fagbetegnelse på dansk fx. kemi
Data science

Uddannelsens fagbetegnelse på engelsk fx. chemistry
Data Science

Den uddannedes titel på dansk
Bachelor (BSc) i data science

Den uddannedes titel på engelsk
Bachelor of Science (BSc) in Data Science

Hvilket hovedområde hører uddannelsen under?

Naturvidenskab

Hvilke adgangskrav gælder til uddannelsen?

Hovedområde naturvidenskab: Uddannelsens konstituerende fagelementer (150 ECTS) hører på nær 7,5 ECTS ind under naturvidenskab. Se under *Uddannelsens konstituerende fagområder*.

Uddannelsesspecifikke adgangskrav:
Matematik A med 6 i gennemsnit
Engelsk B med 6 i gennemsnit.

Bacheloruddannelsen i data science kræver solide færdigheder i og forståelse for matematik, der er forudsætning for de matematisk/tekniske dele af uddannelsen, hvilket matematik A med 6 i gennemsnit vil sikre. Da uddannelsen udbydes på engelsk kræves et rimeligt højt niveau både i tale og skrift, hvilket engelsk B med 6 i gennemsnit vil sikre.

Er det et internationalt uddannelsessamarbejde?

Nej

Hvis ja, hvilket samarbejde?

Hvilket sprog udbydes uddannelsen på?

Engelsk

Er uddannelsen primært baseret på e-læring?

Nej

ECTS-omfang

180

Beskrivelse af uddannelsen

Vi mennesker har truffet "datadrevne" beslutninger baseret på kvantitativ information i hundreder af år. Men i dag er det blevet muligt at basere beslutninger på objektiv information i en helt ny skala. Det har aldrig været så let og billigt at indsamle og gemme store datamængder. Og vi har aldrig før haft adgang til så kraftfulde computere til at hjælpe med at processere og analysere data. Harvard Business Review har kaldt *data scientist* det "*mest sexede job i det 21. Århundrede*" (Harvard Business Review 2012). I bogen "*Data - virksomhedens nye grundstof*" skriver Mikkel Holm Sørensen og Simon Bentholm: "*Data er ikke længere blot et analytisk instrument, men i stigende grad selve grundstoffet i virksomhedens værdiskabelse på linje med teknologi, talenter og patenter.*" (Gyldendal 2013)

Men der er mangel på mennesker med de rette kvalifikationer til at udnytte disse muligheder, og i særdeleshed på mennesker der mestrer den palette af tekniske, matematiske, analytiske og kommunikative evner, der kræves for at være en velkvalificeret "data scientist" (vi undlader at oversætte det engelske begreb fordi der ikke findes en etableret dansk oversættelse). Kompetencerne er efterspurgt af organisationer, såvel offentlige som private, der ønsker at basere konklusioner og beslutninger på data. Universiteter over hele verden har derfor i de seneste år skabt nye uddannelser i data science på tværs af eksisterende faggrænser, på såvel bachelor- som kandidatniveau.

Data science er på vej til at etablere sig som et eget fagområde, og det er derfor naturligt at have data science uddannelse på bachelorniveau (en trend man også kan se internationalt). Mikkel Holm Sørensen fra KL/7, deltager i aftagermøde, udtaler:
- "I mine øjne giver det virkelig god mening, at data science er en grunduddannelse, og at man kan læse den som en bachelorgrad. Så har studerende tre år til at fordybe sig i de grundlæggende begreber og snuse til hele spektret af kompetencer, man skal mestre, som en god data scientist. Herfra kan de enten gå direkte på arbejdsmarkedet og specialisere sig inden for en branche eller tage en overbygning med fokus på forretning eller endnu større teknisk specialisering." (Mikkel Holm Sørensen, Direktør, KL7. 2016)

På en kandidatuddannelse i forlængelse af en BSc i data science vurderer vi, at man i høj grad vil kunne bruge de kompetencer, der er opnået i bacheloruddannelsen. Det giver en anden og dybere data science profil end man ville opnå med en kandidatoverbygning i data science.
Vi mener desuden, at de bedste forudsætninger for at rekruttere flere unge til en IT-karriere, herunder målgrupper der ikke traditionelt ville søge en teknisk uddannelse, findes på bachelorniveau. Derfor har vi målrettet uddannelsen til studerende med interesse for matematik, analyse og anvendelser. Der bør ikke forudsættes nogen eksisterende tekniske interesser eller kompetencer (fx programmeringserfaring), men derimod en interesse i at bruge matematisk baserede metoder til at gøre en forskel i en bred vifte af anvendelsesområder.

Data science på ITU. ITU ønsker at starte Danmarks første bacheloruddannelse i data science, som desuden bliver den mest komplette og bredt anvendelige baggrund for fremtidens data scientists. Uddannelsen giver omfattende analytiske og tekniske kompetencer, der dækker alle aspekter af at håndtere, behandle og analysere data. De studerende bliver trænet i at anvende disse kompetencer i realistiske situationer; herunder at interagere med eksperter i et givet anvendelsesområde for at formulere relevante målsætninger for de data-drevne beslutninger. Specielt får de kommunikationskompetencer, der gør det muligt for dem at hjælpe beslutningstagere uden teknisk baggrund med at operationalisere den viden, der ligger i data.

De fleste, der i dag arbejder som data scientist, har uddannelse i et beslægtet teknisk, samfundsfagligt eller naturvidenskabeligt område, hvor databearbejdning og analyse indgår. Det betyder at de som udgangspunkt mangler en

del af paletten af relevante kompetencer, og på egen hånd må forsøge at skaffe sig yderligere kompetencer. Vi ønsker derimod at uddanne bachelorer med en samlet pakke af de kompetencer, der kræves for at blive en kompetent, værdiskabende data scientist. Det inkluderer kompetencer på tværs af traditionelle faggrænser fra datalogi, statistik, kvantitativ sociologi, såvel som komplementære kompetencer i teknisk kommunikation og etiske aspekter af databehandling (*Se uddannelsens samlede mål for læringsudbytte i The Academic Profile and Objectives for the programme i Dokumentation af efterspørgsel på uddannelsesprofil.*)

Relevans for jobmarkedet. Uddannelsen sigter på at uddanne alsidige data scientists med en stærk kernefaglighed, som er direkte attraktive på jobmarkedet. Bachelorerne vil kunne specialisere sig i et anvendelsesområde på kandidatniveau. Kombinationerne gør det muligt at bidrage til at lukke flere "huller" i det danske udbud af universitetsgrader, som diskuteret ovenfor og i afsnit *Kort redegørelse for behovet for den nye uddannelse*, f.eks.:
- Bachelorer i Data Science, der arbejder med generelle data science problemstillinger.
- Bachelorer i Data Science med en specialisering på kandidatniveau i softwareudvikling (opnået fx ved en kandidatgrad i softwareudvikling på ITU).
- Bachelorer i Data Science med en specialisering på kandidatniveau i forretnings-IT (opnået fx ved en kandidatgrad i Digital Innovation and Management på ITU).

Vores kvalitative og kvantitative undersøgelser bekræfter et stærkt behov for de nævnte profiler, og bekræfter således at konklusionerne fra en lang række internationale undersøgelser også kan overføres til Danmark.

- En uafhængig rapport udfærdiget af Rambøll (Rambøll-Markedsundersøgelse. 2016. I upload: *Dokumentation af efterspørgsel på uddannelsesprofil*.) peger på et stort udækket behov, der ikke tilfredsstillende opfyldes af nuværende uddannelser.
- Det specifikke indhold i uddannelsen er diskuteret med et bredt og højt kvalificeret panel af ledere, der er potentielle aftagere til vores data scientists. Af mødereferatet (Minutes Employers Meeting. 2016. I upload: *Dokumentation af efterspørgsel på uddannelsesprofil*) fremgår det at alle tre ovennævnte profiler vil være efterspurgte.
- En kvantitativ undersøgelse af jobannoncer fra danske virksomheder (Job Listings Analysis. 2016. I upload: *Dokumentation af efterspørgsel på uddannelsesprofil*) påviser at der allerede er en betydelig efterspørgsel efter data science kompetencer. I en 3-måneders periode i foråret 2016 var der således 270 jobannoncer, hvor en bachelor i data science (men ikke fx en uddannelse i statistik eller datalogi) ville give det fornødne fundament.

Uddannelsens konstituerende faglige elementer

The programme comprises mandatory study activities worth 135 ECTS points, track specific study activities worth 15 ECTS, optional study activities worth 15 ECTS points, and a bachelor project worth 15 ECTS points.
The courses listed below constitute the core elements of the programme and are worth 150 ECTS points.

- Introduction to Data Science and Programming (15 ECTS)
- Applied Statistics (7,5 ECTS)
- Data Science in Research, Business and Society (7,5 ECTS)
- First Year Project (15 ECTS)
- Algorithms and Data Structures (7,5 ECTS)
- Linear Algebra and Optimization (7,5 ECTS)
- Machine Learning (15 ECTS)
- Data Management (7,5 ECTS)
- Network Analysis (7,5 ECTS)
- Second Year Project (15 ECTS)
- Data Visualization and Data-driven Decision Making (7,5 ECTS)
- Technical Communication (7,5 ECTS)
- Security and Privacy (7.5 ECTS)
- Reflections on Data Science (Videnskabsteori)(7,5 ECTS points)
- Bachelor Project (15 ECTS points)

There will additionally be two elective courses, and mandatory courses depending on the student's "track".
For a graphical overview of the programme please see *Course plan* in upload: *Dokumentation af efterspørgsel på uddannelsesprofil*.

Assisted collaboration with industry. Projects that connect to real data and questions are an integral part of the programme, developed in collaboration with external partners. The design of the projects was discussed with employers at the employers' meeting 30 May at ITU and all the employers present expressed great interest in participating and contributing in collaborative projects (Minutes Employers Meeting. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*).

Qualification for ITU's MSc programmes. There are two specialisations so called "tracks" with slightly different mandatory courses: Technical and Business. The Technical track gives right to admission to the MSc in Software Development and the Business track gives right to admission to the MSc in Digital Innovation & Management.
Technical track: Programming and Software Engineering (7,5 ECTS) and Large-scale Data Analysis (7,5 ECTS).
Business track: Organisation Theory (7,5 ECTS) and Process Management (7,5 ECTS).

Exchange opportunities. The 5th semester is well suited for a stay abroad. We expect that it will be possible to find courses matching the two mandatory courses and one track course in most universities.

Study environment. The programme is accompanied by academic support activities such as the study café and mentoring programmes that promote a culture of engagement and a deep approach to learning among students.

Short course descriptions:

- *Introduction to Data Science and Programming.* An introduction to Data Science, including basic data processing, visualisation, data sampling, prediction, inference, and probability, coupled with a hands-on introductory programming class. After the course, students should be able to do basic data processing and analysis using Python as a tool. Possible books: (1a) Computational and inferential thinking [1] or (1b) Data Science from Scratch [2], together with (2) Introduction to Programming in Python [3].
- *Applied Statistics.* An introduction to statistical learning, including basic probability theory, statistical modeling and inference. Examples are implemented in the functional statistical programming language R. After the course students are able to make valid statistical analyses, and identify statistical flaws in data analysis. Possible book: An Introduction to Statistical Learning with Applications in R [4].
- *Data Science in Research, Business, and Society.* A substantive discussion of applications and opportunities for Data Science with examples of real world problems coupled with an overview of field-specific research approaches for addressing these problems. After the course students are able to develop empirical research questions and define empirical research problems, discuss organizational processes and societal concerns relevant to Data Science. Possible books: Big data, little data, no data: scholarship in the networked world [5] and The data revolution: Big data, open data, data infrastructures and their consequences [6].
- *First Year Project*. Data Science project with data sets and objectives proposed by (in-house) domain experts. Composed, for example, of a series of seven 2-week mini-projects. A coordinator will run the course managing domain experts in business, games, digital humanities, and IoT. After the project students are able to translate diverse problem settings into a well-defined data analysis problem, implement a solution to the problem, and translate the findings back to the problem domain.
- *Algorithms and Data Structures.* The course introduces structures that allow efficient organization and retrieval of data, frequently used algorithms, and generic techniques for modeling, understanding, and solving algorithmic problems. After the course students are able to make informed choices of algorithms and data structures for data analysis based on knowledge of their resource requirements. Possible book: Algorithms and data structures: The Basic Toolbox [7].
- *Linear Algebra and Optimization.* The goal of this course is to provide firm foundations in linear algebra and optimization techniques. After the course students are able to analyze and solve optimization problems arising in areas such as machine learning. Possible book: Fundamentals of Linear Algebra and Optimization [8].
- *Machine Learning.* Comprehensive overview of machine learning methods and their application in practice. After the course students are able to choose appropriate machine learning methods for a given setting, and make simple modifications to machine learning methods to improve effectiveness. Possible book: Pattern Recognition and Machine Learning [9].
- *Data Management.* Introduction to fundamental concepts of data modelling, storage, organization, querying and transactions. After the course students are able to choose a data model and a database system that fits a given problem definition, and implement analytics solutions (e.g. in SQL) that pull data from a database system. Possible book: Patterns in Data Management [10].
- *Second Year Project*. Data Science project with topic, data sets and domain expertise proposed by external stakeholders (e.g., analysis of patterns of traffic around schools in the city of Copenhagen, in partnership with Københavns Kommune). After the project students are able to combine various tools and methods to get new insights in a complex domain. This can also be a chance to get experience with widely used data analysis tools such as SAS.
- *Reflections on Data Science.* Epistemology and ethics of quantitative reasoning. After the course students will be able to analyze a problem of interest that touches upon the relationships between Data Science and its context (may it be of political, ethical, philosophical, historical or societal nature). Reading material: a compendium of components of different books/articles to cover content relevant specifically to Data Science.
- *Technical Communication.* Skills for communicating analytics findings to non-technical stakeholders. After the course

students will be able to choose and use effective means for communicating, orally and in writing, technical findings at all levels of understanding. Possible book: Technical Communication [11].
- *Large-scale Data Analysis* (Technical track). Overview of the models, systems and algorithms for analyzing large volumes of data, both in batch and in real-time. After the course students are able to use frameworks for large-scale data analysis and reason about the correctness and efficiency of their implementations. Possible books: (1) Map-Reduce Design Patterns [12], and (2) Big Data Analytics Beyond Hadoop: Real-Time Applications with Storm, Spark and More Hadoop Alternatives [13].
- *Network Analysis.* Basics of graph theory, network analysis methodology and social network analytics approaches. After the course students are able to describe, measure and analyze social networks, extract and visualize network data and apply this knowledge to real-world problems. Possible book: Introduction to Social Network Analysis with R [14], supplemented by Networks, Crowds, and Markets [15].
- *Data Visualization and Data-driven Decision Making*. Basic introduction to data visualisation techniques for exploratory, analytical and presentation purposes. Introduction to data-driven decision making strategies, demands and concerns in public and private organizations. After the course students are able to define necessary resources for a particular decision-making task, identify the avenues for obtaining data and insights and discuss potential legal and societal concerns with this process, to select appropriate data visualisation techniques and to produce visualisation of relevant data for stakeholders. Possible books: (1) Visual Display of Quantitative Information [16], and (2) Data Science for Business [17].
- *Software Engineering and Programming* (Technical track). Introduction to the Java language and modular software development aimed at students that already know a programming language (such as Python). After the course, students are fluent in Java and modular software development.
- *Security and Privacy.* Techniques for information security and privacy in a networked world. The course will include a basic overview of the General Data Protection Regulation as well as an overview of legal, technical and societal concerns around data collection, data processing and data-driven decision making. After the course students are able to design solutions that comply with regulations, based on security and privacy best practices. Possible books: (1) Elementary Information Security [18] and (2) Privacy in context: Technology, policy, and the integrity of social life [19].
- *Organization Theory* (Business track). Introduction to the field of organization theory, addressing topics such as work, action, practice and complexity and how to deal with such topics analytically. After the course, students will be able to demonstrate how information technologies and data affect organizational activities and practices. Possible reading material: Images of Organization [20] or Organization theory [21], together with a selection of journal papers.
- *Process Management* (Business track). Introduction to the analysis, design, implementation, and improvement of organizational work processes and supporting IT systems. After the course, students will master the phases of the process management lifecycle, from process identification over modeling, analysis, redesign, automation, to process monitoring and performance management. Students will have engaged with real-world case studies related to process-oriented management methodologies like Six Sigma and Lean Management. Possible book: Fundamentals of BPM [22].

Examples of electives:
- Deep learning
- Natural language processing
- Multimedia analytics
- Business foundations
- Algorithm design
- Network society
- Digital material and social media
- Functional programming

Course book references

[1] Adhikari, Ani and John DeNero. Computational and Inferential Thinking. https://www.gitbook.com/book/ds8/textbook/details
[2] Grus, Joel. (2015). Data Science from Scratch. First Principles with Python. O'Reilly. http://shop.oreilly.com/product/0636920033400.do
[3] Sedgewick, Robert, Kevin Wayne and Robert Dondero. (2015). Introduction to Programming in Python. Pearson. http://introcs.cs.princeton.edu/python/home/
[4] James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. (2013). An Introduction to Statistical Learning with Applications in R. Springer. http://www-bcf.usc.edu/~gareth/ISL/
[5] Borgman, C. L. (2015). Big data, little data, no data: scholarship in the networked world. MIT Press. https://mitpress.mit.edu/big-data-little-data-no-data
[6] Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructures and their consequences. Sage. https://thedatarevolutionbook.wordpress.com/
[7] Melhorn, Kurt and Peter Sanders. (2008). Algorithms and Data Structures. Springer. http://www.springer.com/gp/book/9783540779773
[8] Gallier, Jean. (2016). Fundamentals of Linear Algebra and Optimization. http://www.seas.upenn.edu/~cis515/linalg.pdf
[9] Bishop, Christopher. (2006). Pattern Recognition and Machine Learning. Springer. http://www.springer.com/gp/book/9780387310732
[10] Dittrich, Jens. (2015). Patterns in Data Management. https://infosys.uni-saarland.de/datenbankenlernen/Patterns_In_Data_Management_Preview.pdf
[11] Markel, Mike. (2012). Technical Communication. 10th ed. Bedford/St. Martin's.
[12] Shook, Adam and Donald Miner. (2012). MapReduce Design Patterns. O'Reilly. https://www.safaribooksonline.com/library/view/mapreduce-design-patterns/9781449341954/
[13] Srinivas Agneeswaran, Vijay. (2014). Big Data Analytics Beyond Hadoop: Real-Time Applications with Storm, Spark, and More Hadoop Alternatives. Pearson FT. https://www.safaribooksonline.com/library/view/big-data-analytics/9780133838268/
[14] Bojanowski, Michal. (2017). Introduction to Social Network Analysis with R. Wiley. http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1118456041.html
[15] Easley, David and Jon Kleinberg. (2010). Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press. https://www.cs.cornell.edu/home/kleinber/networks-book/
[16] Tufte, Edward R. The Visual Display of Quantitative Information. https://www.edwardtufte.com/tufte/books_vdqi
[17] Provost, Foster and Fawcett, Tom. (2013). Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media. http://shop.oreilly.com/product/0636920028918.do
[18] Smith, Richard E. (2016). Elementary Information Security. 2nd ed. Jones & Bartlett. http://www.jblearning.com/catalog/9781284055931/
[19] Nissenbaum, Helen. (2009). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press.http://www.sup.org/books/title/?id=8862
[20] Morgan, Gareth. (2007). Images of Organization. Sage. https://uk.sagepub.com/en-gb/eur/images-of-organization/book229704
[21] Hatch, Mary Jo and Cunliffe, Ann L. (2013). Organization Theory. Oxford University Press. http://www.oupcanada.com/catalog/9780199640379.html
[22] Dumas, Maron, Marcello La Rosa, Jan Mendling, and Hajo A. Reijers. (2013). Fundamentals of Business Process Management. Springer. http://shop.oreilly.com/product/0636920033400.do

Begrundet forslag til taxameterindplacering

As the BSc in Data Science is located within natural sciences in line with e.g. BSc in Computer Science and BSc in Software Development ITU proposes taxameter 3.

Forslag til censorkorps

The censors will be taken from the Datalogisk censorkorps, which is currently the most relevant. However we acknowledge that due to the unique composition of Data Science competencies and corresponding courses, not all courses on the programme will find an appropriate external examiner (censor) in this censorkorps. Censors for the Data Science, in Research, Business and Society course may be difficult to find in the Datalogisk censorkorps in its current make-up.

Dokumentation af efterspørgsel på uddannelsesprofil - Upload PDF-fil på max 30 sider. Der kan kun uploades én fil.

Dokumentation af efterspørgsel på uddannelsesprofil Bachelor i data science.pdf

Kort redegørelse for behovet for den nye uddannelse

There is widespread and growing demand for data scientists on the Danish job market, as well as worldwide (particularly in USA and Europe). Currently the demand for data scientists is so high that companies must seek candidates abroad (particularly, in the USA and the UK). Federal reports in Denmark warn of this current and expected growth, with the lack of expertise in Data Science (also sometimes referred to as Big Data) consistently named as a central area for concern. And it is an explicit recommendation from Danmarks Vækstråd to educate more data scientists (Danmarks Vækstråd, Anbefalinger vedr. datadreven udvikling og vækst. 2015).

Denmark is behind in the data game (emphasis is ours):
- "*[...] Danmark er det mest digitale samfund i EU, og at virksomhedernes digitalisering på mange områder er helt i front. Fremadrettet er der dog også en række udfordringer, som skal tages alvorligt. [...] Samtidig* ***er virksomhederne bagud på anvendelsen inden for nyere digitale vækstområder – som fx dataanalyse***" (Erhvervs og Vækstministeriet: Redegørelse om Danmarks digitale vækst 2016 (a))

- ***"Sammenlignet med de bedste i Europa, herunder også øvrige nordeuropæiske lande, indsamler og analyserer danske virksomheder i mindre grad kundedata.*** *I en digital tidsalder er data et afgørende råstof, der skaber nye forretningsmuligheder for erhvervslivet, og som virksomheder kan bruge til at optimere deres forretningsgang"* (Erhvervs og Vækstministeriet: Redegørelse om Danmarks digitale vækst 2016 (b), p11)

- *"[D]er inden for både energi, fødevarer og handel er store potentialer inden for big data. En mindre gruppe af frontløbervirksomheder er allerede begyndt at høste gevinsterne. Men generelt er potentialerne langt større end resultaterne i alle tre sektorer"* (IRIS Group, Datadreven vækst i Danmark. 2015, p3)

One source for this lagging behind is the lack of Data Science competencies in Denmark, as well as educational resources to produce candidates with these (emphasis is ours):

- *"Der kan også [...] konstateres en række fælles træk i de barrierer, der gør det vanskeligt for virksomhederne at høste potentialerne. Det drejer sig bl.a. om usikkerhed,* ***manglende indsigt i anvendelsesmuligheder, kompetencemangler,*** *uensartede standarder, adgang til offentlige data og tilbageholdenhed ved at dele data med andre i værdikæden. [...] udbredelsen af big data og datadreven forretningsudvikling kan stimuleres gennem målrettede forbedringer i en række offentlige vækstbetingelser. Det gælder på områder som* ***forskning, uddannelse****, regulering, prissætning på offentlige data og adgang til uvildig sparring."* (IRIS Group, Datadreven vækst i Danmark. 2015, p3)

- *"The* ***demand for data specialist skills will exceed the current supply of the labour market and the current capacity of education and training systems,*** *requiring rapid adjustments in curricula and the skill sets of teachers and on-the-job workers."* (OECD, An OECD horizon scan of megatrends and technology trends in the context of future research policy. 2016, p53)

- *"Det er bekymrende, at vi er bagud med at bruge digitale værktøjer til at producere smartere eller til at analysere data – som der er et meget stort forretningspotentiale i – bl.a. fordi* ***vi mangler de rigtige medarbejdere.****"* --Adam Lebech, Director, CEO, DI Digital (DI: Stort behov for øget dansk digitalisering. 2016)

- *"Lige nu har vi en masse dygtige folk, der sidder og arbejder med det data, vi indhenter. Men det er folk, som er uddannet i noget andet, og det vil sige, at vi ikke har nogen uddannet i den faglighed, som en data scientist vil have – og den*

faglighed vil vi gerne have." -- (Lisbeth Nielsen, Direktør, Sundhedsdatastyrelsen. 2016)

ITU's BSc in Data Science programme would provide an important and much needed contribution to Danish Data Science competences. The contribution will be unique: the first Bachelor of its kind, strong on Data Science related technical and mathematical skills, but with a central real-world applications focus. ITU stands uniquely capable to make this contribution to the comprehensive development of Data Science competencies in Denmark, due to its de facto interdisciplinary nature, especially on the frontier between technology and society.

Outline of ITU's research into the need for a BSc in Data Science.
ITU's research into the need for the new BSc in Data Science has been both qualitative and quantitative.
1. Qualitative research.
- ***The Rambøll report.*** ITU enlisted Rambøll Management Consulting to conduct a thorough independent qualitative investigation into the education demand within Denmark both in terms of the following:
a) The current and foreseeable required skills and competencies of data scientists as described by industry leaders in Denmark. (Cf. Section *Hvilke aftagere/aftagerorganisationer har været inddraget i behovsundersøgelsen* for a list of responants.)
b) The education landscape in Denmark with respect to these required skills and competencies, and how well this landscape meets the needs of the demand.
- Employers panel. ITU has assembled its own panel of industry leaders in Denmark to help form the content of the BSc in Data Science programme. Documentation from these meetings takes the form of minutes.

2. Quantitative research.
- Job listing analysis. ITU conducted a quantitative analysis of the Danish job market in terms of actual existing jobs for which Bachelors would be qualified.
All corresponding reports are attached as supplementary documentation. In addition, all data extracted and analysed for the quantitative report is available at https://owncloud.itu.dk/index.php/s/7ju4GeEnw2k5zPN . We now provide a short summary of our findings.

❖ ***The Rambøll Report (May 2016)***. In order to confirm that the global trend of increasing demand for data scientists is also manifest in Denmark, in February 2016, ITU enlisted Rambøll Management Consulting (Rambøll) to carry out an independent comprehensive qualitative investigation - including intense study of related current education programmes in Denmark, official federal reports, and direct employer feedback from a selection of leaders from the most influential companies within Denmark - into:
1. the current educational offerings within Data Science in Denmark (Section 3: "Delopgave I"), and
2. the expected size of the Danish labour within Data Science, and in particular with a 5-10 year horizon (Section 4: "Delopgave II").
The resulting report, *Markedsundersøgelse for data science* is uploaded with this application. We now present a summary of the relevant main conclusions, and explain how ITU's proposed Bachelor's programme relates (Cf. Rambøll-Markedsundersøgelse. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*, p21).

1. ***Few Danish educations comprise the right combination of competencies for becoming a data scientist.*** Rambøll conducted an exhaustive study of the educations in Denmark, finding that one Bachelor programme at DTU (Mathematics and Technology) could provide Bachelors with a sufficient cross-disciplinary understanding to serve as a basis for Data Science. Further, Bachelor students of, for example, Technology, could tailor their programmes so that they could carry

out Data Science. However, there are no existing comprehensive educations leading to becoming a full-fledged data scientist in Denmark.
➤ ITU's BSc in Data Science would be the first degree of its kind in Denmark, and we have been assured that Bachelors of this programme would be hireable directly out of the BSc programme both by leaders in Danish industry at our Employers Panel Meeting (Minutes Employers Meeting. 2016. Upload in *Dokumentation af efterspørgsel på uddannelsesprofil*) as well as via our own quantitative job listings analysis (Job Listings Analysis. 2016. Upload in *Dokumentation af efterspørgsel på uddannelsesprofil*). (See also Example Job Profiles, p. 14 in upload: *Dokumentation af efterspørgsel på uddannelsesprofil*.) Furthermore, the competencies that the programme will develop are aligned those laid out by the Rambøll report (Cf. Course Plan, p2 in upload: *Dokumentation af efterspørgsel på uddannelsesprofil*).

2. ***Currently Data Science is a specialisation at the Master's level.*** The most relevant Danish educations within Data Science typically consist of a Bachelor education providing students with general Mathematics or Technology competencies, after which they individually can choose to specialise in Data Science at the Master's level.
➤ Bachelor level studies focussing on Data Science will ensure a deeper, more versatile, and more comprehensive data scientist profile, which could not achieved by just a specialisation at the MSc level. Moreover, ITU's data scientists will have a strong analytical, results-oriented profile, with a view to solving real-world problems rather studying pure science for science. The job prospectives for the resulting graduates correspond to Big Data versions of "Small Data" Market Analysts, Business Analysts, Games Analysts, or other Quantitative Social Scientists, where real-world questions are prioritised, and techniques to solve them are a means rather than an ends. Still, the content of the programme will be high on sophisticated mathematical and technological skills required for Data Science, and it will be possible to complete a more technology-focussed Data Science Master's (on the Bachelor's Technical track, together with the MSc in Software Development).

3. ***The demand for data scientists is large and increasing.*** This signal is clear, both nationally and internationally. Typically data scientists have been employed within the Technology sector, but there is increasing demand for Data Science with in the Finance, Marketing, and Public sectors. The demand is so great that many companies are forced to hire abroad. All industry leaders surveyed agreed that this demand will continue in the 5-10 year spectrum, and that educating future data scientists cannot go fast enough.

❖ ***Employers Meeting Minutes.*** Besides the feedback from relevant employers in the Rambøll Market investigation report ITU has also received direct feedback from other relevant employers at a half day meeting May 30, 2016 at the university. The employers present were industry and public sector leaders in Denmark. (Cf. *Hvilke aftagere/aftagerorganisationer har været inddraget i behovsundersøgelsen*.) See upload: *Dokumentation af efterspørgsel på uddannelsesprofil* for the full meeting minutes (Minutes Employers Meeting. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*).
Before the meeting the employers received a 9-page programme proposal which included a description of the programme, graphic course plan, short course descriptions, examples of learning outcomes, possible MSc combinations together with examples of job profiles in relation to existing job positions, as well as the market investigation report *Markedsundersøgelse for data science* made by Rambøll (Rambøll-Markedsundersøgelse. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*). At the meeting the demand for data scientists was discussed based on the demand in the industry and public sector represented by the employers present at the meeting. For more description of employers, see Section *Hvilke aftagere/aftagerorganisationer har været inddraget i behovsundersøgelsen?*
The employers also provided feedback on the the content and structure of the programme in relation to the demand, which helped produced the programme contents and structure in its current form. For examples see Section *Hvordan er det*

sikret, at den nye uddannelse matcher det påviste behov?

The employers stressed that they were positive towards the programme design, found it important that the programme is developed, and would like to hire both Bachelors and Graduates.
➤ Our Bachelors will have received a comprehensive foundation for practising Data Science, with focussed Data Science related scientific and technical skills and extensive work in their application. The programme is an ideal fit for students whose main interest is data driven solutions and analysis of real-world problems, rather than the isolated technical skills. We treat Data Science as a discipline in its own right, rather than as a specialisation in order to provide a complete Data Science education for versatile generalist data scientists that are capable of applying to a wide variety of real-world problems. In addition, because of the unique profile of our intended student base for IT, the programme should be successful at attracting new young students to IT who would not have originally sought out a more technical or scientific education.

❖ ***Job listings analysis***. ITU carried out quantitative studies regarding the current demand for data scientists in Denmark. The resulting technical report is provided (Job Listings Analysis. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil)*. The report includes two studies on Danish Data Science job listings data collected over 3-6 month periods in 2016:
1. a purely data-driven quantitative analysis Data Science related jobs advertisements available over 6 months of 2016, and
2. a hand-annotated quantitative analysis on Data Science related jobs advertisements **for which graduates from the proposed program would be immediately qualified**, over a period of 3 months in 2016. The description of the 3 month study is given in Section *Underbygget skøn over det samlede behov for dimittender* as an estimate for the current demand in Denmark for data scientists.
The studies also provide concrete numbers the public and industry sectors in which the data scientist demand is manifested. The main contribution of the studies is to show how the current demand for data scientists well exceeds the prospective number of graduates from the ITU's proposed data scientist BSc programme, by several magnitudes, and that the BSc-MSc progressions proposed by ITU here are well-founded: there are a large majority of job advertisements related to business, and a solid presence of more technical Data Science jobs related to software.

Underbygget skøn over det samlede behov for dimittender

How many graduates in Data Science should ITU produce? There is ongoing and increasing demand for competent data scientists (Cf. Section *Kort redegørelse for behovet for den nye uddannelse*). According to our the panel of Employers, educating data scientists can only go too slow (Minutes Employers Meeting. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*).

ITU carried out a quantitative study to gain an idea of the current demand for data scientists in Denmark. Job advertisements from a central Danish job site were collected over a three month period (March, April, May of 2016), using a variety of keywords to ensure high recall on Data Science related jobs. The advertisements were then hand-filtered for relevance to graduates from the proposed programme and annotated as (1) business-related, (2) software-related, or (3) other. In total 270 were found to be job advertisements where a BSc in Data Science would be relevant (but a BSc in, say, software development would not be a good enough background). About 65% of these are in business intelligence and 27% are software related.

The Rambøll report also provides some indicative figures for the future to this end.

- "[...] *ITEK6 under Dansk Industri [opstiller] en prognose, som viser et udækket behov i Danmark i 2020 på ca. 3.000 kandidater generelt i ITEK-branchen. ITEK konkluderer, at der skal tillægges en generel vækst inden for teknologibranchen, bl.a. i form af fremdrift i Big Data, hvilket skønnes at udgøre op til yderligere 3.000 kandidater (ITEK, 2015). I den tidligere nævnte undersøgelse af IDC tilkendegiver 13 pct. danske virksomheder, at de har planer om at ansætte en specialist inden for Big Data, mens 21 pct. udtrykker et behov for kompetencerne, men mangel på ressourcer (IDC, 2011). Ligeledes peger regeringens redegørelse for den digitale vækstplan på, at på trods af en stigning i antallet af it-uddannede, er det fortsat en udfordring at få dækket virksomhedernes behov inden for it, især i forhold til adgangen til specialiserede it-kompetencer, som fx it-udviklere og dataanalytikere (Regeringen, 2015)."*
(Rambøll-Markedsundersøgelse. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*, pp35-36).

Expected intake into the BSc in Data Science. The ITU will start by admitting 50 students on the BSc in Data Science in 2017 and thereafter expects to increase the number of students according to the demand. Based on input from employers, Bachelors in Data Science may be able to get a job immediately following their BSc (Minutes Employers Meeting. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*). They may also choose to remain and combine their BSc with a relevant MSc, specialising for a particular field of application at the graduate level. (See the combination of MSc and job profiles in Section *Sammenhæng med eksisterende uddannelser* and in *Example Job Profiles* in upload: *Dokumentation af efterspørgsel på uddannelsesprofil*.)

Hvilke aftagere/aftagerorganisationer har været inddraget i behovsundersøgelsen?

The employers who participated in in the market investigation both at ITU and in the Rambøll report were carefully selected. The selection process took place in February and March 2016. ITU listed 60 potential employers including information on their affiliation, linkedin profile, area, and educational background and the list was used as basis for the selection for both the Rambøll report and the employers' meeting at ITU.
In order to have the employer's view represented in the selection, a Chairman for the coming group of employers was appointed. Kaare Brandt Petersen, Nordic Director for for Education & Academic from SAS Institute, assisted ITU in selecting a sub-group of employers to represent a broad range of employers for Data Science related positions in relevant industries and organisations. The selection further ensured that both small and large companies were represented as well as both private and public sectors. Furthermore the educational background of the employers were also taken into account in order to ensure a broad perspective on Data Science.
The employers were approached in April 2016 with an initial description of the programme and invited to attend a meeting in May. The employers approached were very positive and seven out of the eight invited attended the meeting on May 30, 2016.
Since the May meeting employers are receiving updates on the continued development of the programme. At the end of August, ITU sent an update on which adjustments have been made based on the May meeting (see Section *Hvordan er det sikret, at den nye uddannelse matcher det påviste behov?* for a more detailed description). After the submission of the application for pre-qualification ITU will continue the dialogue with the employers and invite them to be part of a coming employers panel for the programme.

For the Markedundersøgelse for data science report ITU provided the above mentioned list of relevant employers to Rambøll and was in dialogue with Rambøll for the selection. In the appendices in the report it is stated: *...Da data science er et område, der er relevante indenfor flere og flere sektorer, har det været en prioritet at afdække et bredt felt potentielle aftagere indenfor forskellige brancher og sektorer… Respondenterne har primært haft stilling som chef eller direktør i virksomhedens teknologiske og analytiske afdelinger, mens respondenterne i de resterende interview ligeledes har været chefer, partnere eller direktører samt haft kendskab til data science på anden måde.* (Rambøll-Markedsundersøgelse. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil* p.57).

Please see *Hvilke aftagere/aftagerorganisationer har været inddraget i behovsundersøgelsen* in upload: *Dokumentation af efterspørgsel på uddannelsesprofil* for a full list of employers and their background.

Hvordan er det sikret, at den nye uddannelse matcher det påviste behov?

In order to determine the relevant contents of the proposed BSc in Data Science programme, ITU carried out the research summarised in Section *Kort redegørelse for behovet for den nye uddannelse* and placed in upload: *Dokumentation af efterspørgsel på uddannelsesprofil* (i.e., the market investigation report made by Rambøll (Rambøll-Markedsundersøgelse. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*), an employers meeting with ongoing feedback (Minutes Employers Meeting. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*), and a data scientist job listings analysis (Job Listings Analysis. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*). The Rambøll report provided verification that the developed Data Science programme corresponded to industry needs in Denmark. For active feedback on the alignment of the proposed Data Science programme with the needs of the employers, ITU gathered an employers panel for a half day meeting on May 30, 2016 at the university. The employers present were industry and public sector leaders in Denmark. (See Section *Hvilke aftagere/aftagerorganisationer har været inddraget i behovsundersøgelsen.*)

Before the meeting the employers received a programme proposal and the Rambøll report. At the meeting the demand for data scientists was discussed based on the demand in the industry and public sector represented by the employers present at the meeting. The employers also provided feedback on the the content and structure of the programme in relation to the demand, which helped produced the programme contents and structure in its current form.

The employers stressed that they were positive towards the programme design, found it important that the programme is developed, and would like to hire the graduates. The employers from the meeting are continuously updated on the progress of the application and the structure of the programme, and are encouraged to provide as much feedback as possible to ensure success.

The feedback for suggested revisions from the employers was taken into account, and the programme contents and structure optimised for this. We now give the optimisations made based on some of the issues raised by employers at the meeting:

1. The employers suggested we create a programme with two branches: one for data scientists that wish to achieve even stronger technical skills, and the other for those that focus on preparing to engage with businesses and organisations by understanding underlying business processes. → The BSc in Data Science now consists of these two corresponding tracks: Technical and Business.
2. The employers expressed concern over students being exposed to real-world problems and learning the possible pitfalls or real data. → The Data Science programme consists of several projects where students are meant to collaborate with industry. This collaboration will be facilitated by ITU; ITU has successfully facilitated such collaboration in its other degree programmes for many years. Moreover, the employers expressed interested in taking part in this collaboration. The programme structure (See *Course Plan* in upload: *Dokumentation af efterspørgsel på uddannelsesprofil* p2) now includes icons to indicate the courses where students will be challenged by real-world problems and associated data as part of the curriculum.
3. The employers expressed concern that the mathematical and statistical content could be at least 1/3 of the programme. → The programme does indeed contain at least 1/3 mathematics and statistics content. This also more than delivers on the required mathematics competencies laid out by the Rambøll report (Rambøll-Markedsundersøgelse. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*, p22. The programme structure (*Course Plan* p2, in upload: *Dokumentation af efterspørgsel på uddannelsesprofil*) now includes icons to indicate the courses heavy on mathematics/statistics content, to make this more explicit. There are also icons to indicate heavy programming content. Bachelors will become highly skilled in Data Science-related mathematics, statistics and programming.

Sammenhæng med eksisterende uddannelser

The BSc in Data Science will be the first of its kind in Denmark and will set a standard for comprehensive Data Science education, as pointed out in Section *Kort redegørelse for behovet for den nye uddannelse.* As laid out in Section *Beskrivelse af uddannelsen*, the study programme aims to educate versatile, strongly competent data scientists, ready for the job market. Bachelors will have the opportunity to further specialise for a particular field of application at the graduate level. Hence the full study programme uniquely aims to fill several gaps in the education landscape. For example:
- Generalist data scientist Bachelors.
- Graduates with deep expertise in Data Science and a specialisation in software development (obtained for example by combining with the MSc in Software Development at ITU, to which Bachelors will have right to admission).
- Graduates with deep expertise in Data Science and a specialisation in business informatics (obtained for example by combining with the MSc in Digital Innovation and Management at ITU, to which Bachelors will have right to admission).

In addition to the two above-mentioned MSc programmes which give right to admission examples of MSc programmes where BSc in Data Science qualify for admission based on the universities' advertised requirements.

- ITU: Games
- ITU: Digital Design and Communication
- DTU: Computer science and engineering
- AAU (Copenhagen): Operations and Innovation Management - Media Management
- AAU (Copenhagen): Operations and Innovation Management - Global Management
- DTU: Mathematical modeling and computation
- KU/AU/AAU/SDU: Computer Science
- KU: IT & Cognition

Examples of job profiles for which these education combinations prepare graduates are listed in *Example Job Profiles* in upload: *Dokumentation af efterspørgsel på uddannelsesprofil.*

Rekrutteringsgrundlag

Target group. The target group for the BSc in Data Science is graduates from all high schools (Upper secondary education) in Denmark (stx, htx, hhx, hf, as well as international schools). Applicants should have strong math skills and solid English skills, combined with an interest in statistics and computing for science, social science and/or business. ITU's Data Science program aims to be attractive to a group of students that are less likely to seek an education in a technical area, thus increasing the number of graduates with strong competences in information technology. This is because the program is designed for students with no prior programming experience and focuses on application of theoretical skills to a wide range of real-world problems.

As the programme has two tracks, the target group consists partly of potential students with a strong technical ambition and partly of potential students with an interest in the business aspect of Data Science. Regardless of which track applicants might choose, no programming skills are required at time of application.

We expect the BSc in Data Science to be popular with a range of applicants. The program will also focus on strategic and sustainable approaches for attracting and retaining female applicants. Many female high school students in Denmark are strong in mathematics and are generally high achievers. However, current research on women in computing suggests that female students are more motivated by the potential of their scientific or technical skills to having an impact and solving the world’s problems. The BSc in Data Science is a programme heavy on mathematics and technical skills for the sole purpose of enabling Bachelors, using these skills, to effectively solve complex real world problems involving data. ITU has also devoted resources, in collaboration with employers panels, on understanding how to increase female intake in its programmes and produce more females in IT. Therefore, the ITU has an ambitious goal of enrolling and retaining 30% female students per year.

Specific minimum requirements:
- English, B-level with a grade point average of at least 6.
- Mathematics, A-level with a grade point average of at least 6.

However both grades are ITU standards for technical BSc programmes and in general it is expected that Data Science students achieve a higher Mathematics grade point average.

Since the programme aims to be attractive to a group of students that are less likely to seek an education in a technical area, there will be an increase in the number of graduates with strong competences in information technology. We therefore do not foresee any direct consequences for existing technical programmes. Furthermore the three current BSc programmes at ITU have many applicants and high grade point average entry cut-offs. This summer, ITU successfully increased the number of applicants to the BSc in Software Development and admitted 50% more students without compromising the quality of the students. In addition, not all students choosing ITU as a first choice were admitted, due to capacity limits.

Forventet optag

2017: 50

Hvis relevant: forventede praktikaftaler

NA

Hermed erklæres, at ansøgning om prækvalifikation er godkendt af institutionens rektor

Ja

Status på ansøgningen

Godkendt

Ansøgningsrunde

2016-2

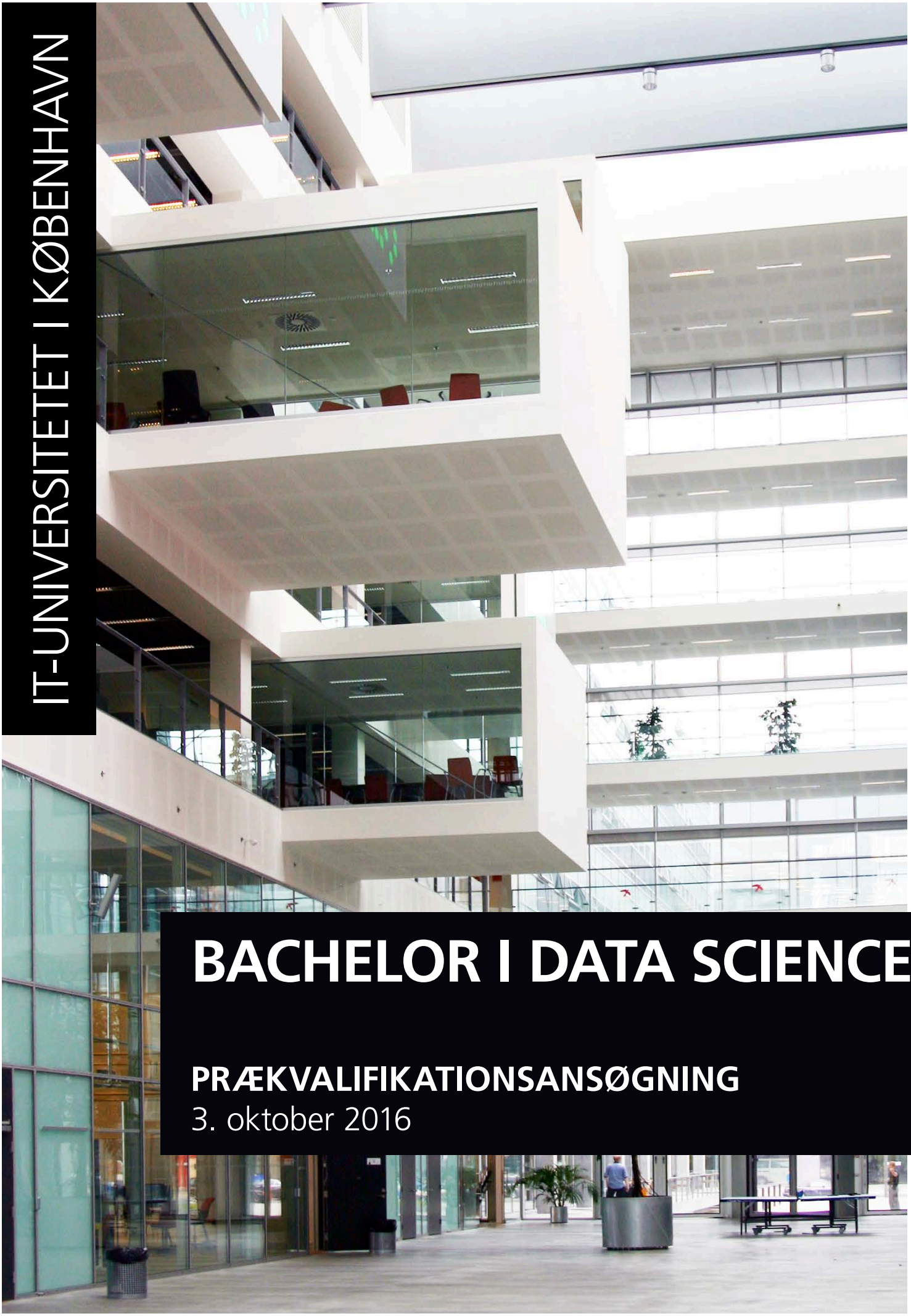
Afgørelsesbilag - Upload PDF-fil

A5 - Godkendelse af ny uddannelse - BA i Data science - ITU (revideret, januar 2017).pdf

Samlet godkendelsesbrev

ITU - Godkendelse af ny uddannelse - BA i Data science.pdf

IT-UNIVERSITETET I KØBENHAVN



BACHELOR I DATA SCIENCE

PRÆKVALIFIKATIONSANSØGNING

3. oktober 2016

Indhold

Indledning 1

Course Plan 2

The academic profile and objectives for the programme 3

Hvilke aftagere/aftagerorganisationer har været inddraget i behovsundersøgelsen 5

Minutes Employers Meeting 6

Job Listings Analysis 10

Example Job Profiles 14

Markedsundersøgelse for data science 16

Relevante sider: side 18-41 (1-24 i selve rapporten uden bilag) .....

Referencer 62

Ansøgning om prækvalifikation af ny uddannelse

På vegne af IT-Universitetet i København fremsendes hermed ansøgning om oprettelse af ny bacheloruddannelse i Data Science.

Dette dokument samler

1. Yderligere dokumentation for behov og relevans for bacheloruddannelsen
2. Andre bilag, der ikke kunne uploades i den digitale ansøgning.

1. Yderligere dokumentation for behov og relevans for bachelor-uddannelsen

IT-Universitetet har gennem en række forskellige undersøgelser afdækket behovet og relevansen for uddannelsen samt i udviklingen af uddannelsen været i dialog med en række relevante aftagere og organisationer – offentlige så vel som private. Se også yderligere beskrivelse i den digitale ansøgning under *Kort redegørelse for behovet for den nye uddannelse.*

- **Markedsundersøgelse for data science.** I foråret 2016 gennemførte Rambøll Management for IT-Universitetet en omfattende markedsundersøgelse, der i to dele kortlagde 1. Eksisterende uddannelser inden for data science i Danmark samt eksempler på internationale uddannelser samt 2. Afdækkede behovet gennem relevante danske og internationale eksisterende rapporter om behov for data science samt afdækkede behovet ved interview med 11 aftagere fra forskellige sektorer og organisationer, såvel offentlige som private. Der er udvalgt mest relevante afsnit i indholdsfortegnelse
- **Job Listings Analysis.** IT-Universitetet har endvidere i en 6 måneders periode i foråret/sommeren 2016 indsamlet relevante data science jobopslag og kan gennem analyse af de indsamlede data dokumentere et eksisterende behov for data science kandidater
- **Example Job Profiles.** Her er beskrevet en række relevante job profiler både på baggrund af BSc in Data Science og eksempler på tilhørende jobopslag
- **Minutes Employers' Meeting.** Referat fra aftagermøde 30. maj 2016

2. Andre bilag, der ikke kunne uploades i den digitale ansøgning.

- **Course Plan.** Grafisk oversigt over bacheloruddannelsen
- **The academic profile and objectives for the programme.** Udarbejdet i overensstemmelse med Den Danske Kvalifikationsramme
- **Hvilke aftagere / aftagerorganisationer har været inddraget i behovsundersøgelsen.** Oversigt
- **Referenceoversigt**

Course Plan

Semester 1	Introduction to data science and programming (15 ECTS)	Applied statistics (7,5 ECTS)	Data science in research, business and society (7,5 ECTS)
Semester 2	First year project (15 ECTS)	Algorithms and data structures (7,5 ECTS)	Linear algebra and optimisation (7,5 ECTS)
Semester 3	Machine learning (15 ECTS)	Data management (7,5 ECTS)	Network analysis (7,5 ECTS)
Semester 4	Second year project (15 ECTS)	Data visualisation and data-driven decision-making (7,5 ECTS)	Programming and software engineering (Technical track) (7,5 ECTS) Organisation theory (Business track) (7,5 ECTS)
Semester 5	Technical communication (7,5 ECTS) Security and privacy (7,5 ECTS)	Large-scale data analysis (Technical track) (7,5 ECTS) Process management (Business track) (7,5 ECTS)	Elective (7,5 ECTS)
Semester 6	Bachelor project (15 ECTS)	Reflections on data science (7,5 ECTS)	Elective (7,5 ECTS)

Programming content · Mathematics content · Applications to real world problems · Social science content

Track courses

The Academic Profile and Objectives for the Programme

Programme Objectives

The objectives of the Bachelor Programme in Data Science are to provide students with the scientific qualifications to independently design and execute efficient, computation intensive, data-centered, empirical analysis of real world problems. This requires bachelors to acquire broad theoretical, practical and methodological knowledge tailored specifically to data science from mathematics, computer science, statistics, and computational social science, as well as an understanding of technical communication and research ethics.

Bachelors will be able to play an independent reflective role in all aspects of the acquisition and communication of data driven intelligence. In particular, they will obtain comprehensive analytical and technical skills covering handling, analyzing, and operationalizing data. They also will become practically skilled in the application of these skills for the formulation and analysis of real-world problems, as well as the effective communication of these results.

Bachelors will be qualified for generalist data science jobs in industry. Further, bachelors will obtain admission qualifications for ITU's Master's programmes, allowing them to specialize their competencies within a particular field of application.

Objectives for Learning Output

On completion of the programme, the student must have attained the following learning output objectives. The learning output is divided into the categories knowledge, skills and competences; cf. the Danish Qualifications Framework for higher education.

Knowledge and Understanding

The bachelor must:

- have research-based knowledge of theory, methodology and practice within data-science specific areas of mathematics: optimisation, machine learning, statistical analysis, network analysis, experiment design, and algorithms.
- have research-based knowledge of theory, methodology and practice within data-science specific areas of computing: programming languages, query languages, databases, data processing, and large-scale data analysis.
- have research-based knowledge of social science, be able to reflect on relevant social theories.
- be able to understand and reflect on theories, scientific methodologies and practice of the above areas.
- be able to understand and reflect on existing data and analytical software-platforms and their adequacy to specific data science problems.
- be able to consider data within a global perspective, with respect to different contexts and cultures.

Skills

The bachelor must:

- be able to develop software in a general-purpose programming language.
- be able to implement scalable algorithms for fundamental data analysis tasks (e.g. in machine learning, statistical inference), based on technical descriptions in textbooks or the research literature.
- be able to evaluate theoretical issues of problems in order to select and apply appropriate machine learning methods, algorithms, and software tools to perform data analysis, statistical inference, or predictive analytics tasks, based on scalability and performance.
- be able to apply systems for data management in order to clean, transform, and query data.
- be able to carry out adequate empirical evaluation of model performance in order to ensure optimal accuracy.
- be able to organize, summarize, and visualize data and the outcomes of inferential processes for relevant stakeholders.
- be able to communicate the academic issues associated with inferential processes and computational requirements to relevant stakeholders.
- be able to assess the level of privacy and security ensured by a technical solution and communicate it in a way that is understandable to non-specialists, both from a global perspective as well as within a particular cultural context.

Competencies

The bachelor must:

- be able to independently and collaboratively dissect complex situations in order to identify the domain knowledge necessary for robust data processing and interpretation of outcomes, and to plan data collection on which robust statistical conclusions can be made.
- be able to independently and collaboratively develop domain-specific data cleaning, statistical models, and select algorithms and analysis and machine learning, based on input from domain experts.
- be able to articulate the importance of data origins, data collection context and impacts of inferential processes as part of reporting of findings.
- be able to independently and collaboratively consider the relevant context of decision-making given available data, and critically assess implications of data analytics and of the inferential process in a domain, including privacy and ethical concerns.
- be able to maintain and develop professional competencies for global state-of-the-art technologies, and adapt to new domains of application.
- be able to develop data-driven analyses and solutions from the bottom up, both independently and collaboratively.

Hvilke aftagere/aftagerorganisationer har været inddraget i behovsundersøgelsen

Navn	Stilling	Firma/organisation	Område	Kontakt
Kaare Brandt Petersen	Nordic Director, Education & Academic	SAS Institute	Business intelligence	Deltaget i aftagermøde ITU
Janus Sandsgaard	Fagchef for IT og Digitalisering	Dansk Erhverv	Association	Deltaget i aftagermøde ITU
Mikkel Holm Sørensen	Direktør / Adfærdsstrategi	KL7	Consulting	Deltaget i aftagermøde ITU i ITU aftagermøde
Søren Ilsøe Overgaard	Underdirektør, Head of Technology, Data & Security	SKAT	Public sector	Deltaget i aftagermøde ITU Respondent i markedsundersøgelse
Daniel Thunbo	Afdelingsdirektør, Head of Analytics & Omnichannel Communications	TDC	Tele-communications	Deltaget i aftagermøde ITU
Camilla Kerlauge	VP, Head of Analytics at Nordea	Nordea	Finance	Deltaget i aftagermøde ITU
Søren Holbech	Data Scientist, Ph.D. & Developer	LiveIntent	PR / marketing	Deltaget i aftagermøde ITU
Adam Lebech	Direktør	Dansk Industri	Association	Respondent i markedsundersøgelse
Birgitte Hass	Direktør	IT-branchen	Association	Respondent i markedsundersøgelse
Jens Jakob Svanholt	Senior Vice President of PD and Decision Modeling	Danske Bank	Finance	Respondent i markedsundersøgelse
Karsten Lund	Creative Director	Lego	Digital entertainement	Respondent i markedsundersøgelse
Erik Verdoner	Business Intelligence Specialist	Copenhagen Airports	Business intelligence	Respondent i markedsundersøgelse
Henrik Lauritzen	CTO	Audience Project	Market research	Respondent i markedsundersøgelse
Johan Kjeldgaard Pedersen	Direktør	Data2impact	Analysis and consulting	Respondent i markedsundersøgelse
Lisbeth Nielsen	Direktør	Sundhedsdatastyrelsen	Public sector	Respondent i markedsundersøgelse
Anders Kofod Petersen	Head of Data Science and Engineering Lab	Alexandra Instituttet	Research and IT	Respondent i markedsundersøgelse

Students Affairs and Programmes
Direkte tlf.: 7218 5087
E-mail: meho@itu.dk
Journalnr.: 2016-791

27 June 2016

Minutes employers meeting in relation to new programme in Data Science at the IT University of Copenhagen, Monday May 30, 2016, 9:00-12:30 in room 3A20

Participants:

Employers:
Kaare Brandt Petersen (Chairman), SAS institute; Janus Sandsgaard, Confederation of Danish Enterprise; Mikkel Holm Sørensen, KL7; Søren Ilsøe Overgaard, SKAT; Daniel Thunbo, TDC; Camilla Kerlauge, Nordea; Søren Holbech, LiveIntent

From The IT University:
Rasmus Pagh, professor; Irina Shklovski, associate professor; Thore Husfeldt, associate professor; Phillip Bonnet, associate professor; Lene Pries-Heje, head of studies, associate professor; Mette Holm Smith, programme coordinator, minutes taker.

1. Round of introductions (focusing on participants' careers and their relation to data science)
Rasmus Pagh welcomed and briefly introduced the background for the meeting. ITU is working on a new BSc in data science and would like to have feedback and input from employers in order to meet the needs for graduates in the industry and in public organisations. A brief round of introductions followed this.

2. Employers give brief oral summaries of the data science challenges and opportunities in their organizations
Under this item the employers gave a number of inputs that have been gathered under a number of headlines below.

Difference between smaller and larger companies.
The panel discussed the different roles for a data scientist employed in smaller and larger companies.

Larger companies and organisations - e.g. SKAT, Nordea, and TDC have many data scientists employed and can have employees who are hard-core data scientists and other employees who have more understanding of business processes.

There are not many employees that have both sides. Instead, they put teams of people together with different angles to cover the all sides. ITU might have to choose whether you want to educate hard-core data sciences to large companies or the general graduate to smaller business. Maybe create a programme with two branches?

Many smaller companies do not have a data scientist employed but they might wish to optimize their business by analysing and using the data available.

If this is the case the data scientist needs to know how to work with a boss without a clue, who only knows that he needs a data scientist. The data scientist is a one-man army that works with civilians. He should be selling confidence and trust, and needs to be able to do self-management.

Small start-ups have limited resources, and working for them you need to be able to do it all (programming, statistics, ...).

The employers' experience with hiring graduates is that they often would prefer to work with people that have similar background.

Teach how to fail and work with real world problems.
During the programme it is important to introduce students to real problems from the real world. In the real world, you have projects with many different people and people who have no knowledge and understanding of the tech side.
It is also important to teach the students to fail. By failing you broaden your skills and knowledge. There was a discussion about how to qualify failing within the frames of an educational institution who have the goal of students not failing but passing their exams. In this context failing should be thought of as "trying something that does not work" rather than as "failing to achieve learning objectives".

ITU should consider how to simulate a business world where projects involve different people, in order to teach the data scientist students where they have their core strengths.

Give students an understanding of what is an enterprise and present them to the flow in different types of industries.

Students should maybe also have possibility for some kind of internships in relevant companies.

How to fill the gap between the tech and the business side and make value?
A definition of a data scientist can be very broad within the area of tech and business.
It is necessary for ITU to develop one or more definitions in relation to the programme. The panel gave input to more data scientist roles but many of the comments focused on a role that can work and understand both the tech and business side:

- There is a need for someone who can understand and link between the business side and the data analytics side.
- In order to create value for the company a data scientist needs to be able to couple up with the business side. The business analyst role is often missing in a company.
- Currently many working data scientists are lacking the interest and understanding of creating value for the company. They are driven by math and tech and are not interested in business processes. They need to understand that it is not only about tech. They must be able to find a way of balancing, and learn to master and synthesise the "right" problems, and need to have an interest in understanding what they are helping the organisation to achieve.

It is important to align the value set for graduates and industries. Also, it is important to attract the kind of applicants to the programme who are open for both. They need to have a toolbox and mindset that are adaptable. And they need to be able to structure problems.

It was mentioned by an employer that it would probably be difficult to attract students to a programme with interest in both tech and business. So it might be wise to focus on the tech part in the programme and learn the rest in the company.

Another suggestion was to team up with other universities from the business side, e.g. CBS.

Algorithms and machine learning
More employers mentioned that on the algorithm side things are developing rapidly. It is possible to use machine learning to solve many new problems in e.g. image analysis. Will this cut out middle layer in the decision process and how will this influence the data scientist role? To what extent will improved tools and algorithms replace the need for a human data analyst?

Quality of data
IT Systems have a long life and contains data input by humans. In companies, ideas will often not be tried out because handling data is very complicated and resource intensive. E.g. why are data a not following the model? This calls for better methods and processes to understand data.

3. Round of feedback on and discussion of the curriculum draft (strengths, weaknesses, fit with perceived need, uncertainties)
The employers provided feedback to the current programme design. The following are comments made by individual employers and do not necessarily reflect the opinion of the panel as a whole:

- SQL seems to be missing. (ITU answered that it **is** part of the content)
- How much math do you want them to know? One third of the programme should be math and statistics. (ITU answered that a significant amount of applied math is part of the program.)
- Machine learning - it is the right book – but also requires strong math skills.
- Applied statistics - important not to focus on aspects of traditional statistics that are not important in data science.
- Network analysis: Cool but looks big – what about text analysis?
- Connection to deployment: Should there be something on bridging the gap to deploying a system for end-users? (ITU answered that this is addressed by the software development MSc.)
- Data Profiling – How to approach received data that may be messy and hard to interpret. E.g. receiving a bunch of database tables. Need to teach an initial approach when receiving a data set.

About data profiling and quality of data: ITU mentioned that it is part of many courses as ILOs. From employer's side it was stated that it is an important focus – it is a discipline in itself. You need to show them the messy data. It is a challenge to let the students gain experience with messy data within the time available. Perhaps let them scan for their own data set or give them a little messy data set.

- Job profiles – why are they at the last page? If you can design a BSc programme that produces graduates that are skilled enough we will hire. But it is also a matter of colleagues approach to working together with a BSc. If they have the skills we will hire them – but if they could study their MSc while they work it would be a perfect match also.
- Look into behavioral analytics as a potential job profile.
- What is a data scientist? It would clarify the intentions to give a definition
- Basic probability theory needs to be part of the statistics course.

- What are you going to learn in the project work – failing – business value? What do they achieve from working in projects?
- Group size. Not too big groups. Three members is a good number – when you are four members paople can start hiding.
- Seek feedback from working data scientists – would this programme have helped me in my current job?

4. How teachers can work with external partners (especially in the 2nd year project and bachelor projects, but possibly also by providing real-world data for courses)
There are three major projects during the programme, which all three calls for domain knowledge.

- First year project. An internal ITU project with interdisciplinary input
- Second year project. A project with external partners
- Third year. Bachelor project. Individual projects in smaller groups preferable in collaboration with industry

All employers would like to participate in a setup with collaboration. The following was discussed:

- Find and define projects with a suitable data set
- Group size – around five students (for 1st and 2nd year projects)
- Data security. For some data sets it might not be possible to take them outside the organisation. Students will have to work on-site or on synthetic data
- Sharepoint for projects. How to communicate projects from companies to students/supervisors? Some kind of platform/site where companies can offer projects to students and supervisors – maybe most relevant for in relation to bachelor projects?
- A contract for the project work needs to be defined
- ITU should make more detailed descriptions of the projects

5. Are external partners interested in helping making students interested in data science (e.g. small video introductions of "what you can do with data science").
Brainstorm ideas for exciting data science career stories (in DK), especially with focus on profiles that do not fit the "tech nerd" stereotype.

There was not enough time to touch on this item but the employers were in general very positive towards the new programme and interested in further collaboration.

6. Wrap-up: Conclusions from the panel
The employers stressed that they were positive towards the programme design, found it important that the programme is developed, and would like to hire the graduates.
They found that it was valuable that ITU invited employers to provide input to the new programme.
It was underlined that it is important to run the external projects and that the employers would like to participate in a future more obliged collaboration – both public sector and industry representatives.

TECHNICAL REPORT

An empirical analysis of the current Data Science Job Market in Denmark

From: Natalie Schluter and Rasmus Pagh

Date: September 21, 2016

Summary: In this report, we provide strong quantitative evidence of the demand for data scientists in Denmark. We also provide solid evidence that the specialisations additionally covered by ITU's proposed education combinations are aligned with this demand. **The numbers from this report provide concrete estimates for the current demand of data scientists in Denmark.** The report is intended to explain the technical details by which these figures were obtained and not in and of itself (above the figures it claims) to supplement the extensive qualitative reports that form part of ITU's Data Science programme application.

Objectives. In this study, we consider the current need for data scientists within Denmark from a quantitative standpoint, and using empirical means. Specifically, we aim to expose evidence of

(O1): the immediate open job market for data scientists in Denmark, and

(O2): the areas of specialisation for graduates pertinent for this job market.

The study is meant to provide concrete estimates for the current demand. This report provides the technical details by which the estimates were obtained. The estimates, provided in the next Section "Our findings", and not these technical details are meant to supplement the extensive qualitative reports provided with our application.

Our findings. Over a 6-month period, March to August 2016, ITU harvested job listings[1] relevant to Data Science in Denmark in order to (1) determine the immediate demand for data scientists, and (2) determine the sectors in which data scientists are required. We carried out this study from two different angles: "small data" quantitative and "big data" quantitative.

⇒ **"Small data" method findings.** We hand annotated the first 3 months (March, April, May) of job listings data for (i) openings that would be directly relevant to Bachelors/Graduates from our Data Science programme, in addition to being Data Science related positions, and (ii) the class of the openings in either (A) Business, (B) Software, or (C) Other.

We found 270 jobs where a BSc in Data Science would be relevant (but a BSc in, say, software development would not be a good enough background). About 65% of these are in Business Intelligence, 7.4% for Software Systems, and the rest in other data science roles.

⇒ **"Big data" method findings.** With very conservative search criteria (to ensure high precision), we extracted 1266 jobs in English and 829 jobs in Danish, for a total of 2195 new Data Science related job listings for the 6 month period alone. We note that some of the jobs may not be directly relevant for graduates of the proposed Data Science Bsc programme (i.e., too senior, or with too specialised knowledge, like pharmacology), however this conservative search covered all but 13 jobs from the "small data" method's hand-annotated dataset. Figure 1 shows the network of English job listings. We computed over the network the general topics of these job listings. We found that the most important data science jobs are in Business, with important contributions in Pharma, Software Systems, and Games.

For the Danish network (Figure 2), there is an even greater disparity. Well over 60% of data science job listings are in Business or the public sector (Kommuner, Skat, Ministries,...), with small minorities (2% or less each) in Finance and Health and other sectors.[2]

[1] From `http://dk.indeed.com/`.

[2] We note that one of the clusters is for Københavns kommune, Koncern-IT division. If there are a group of very similar postings for a single employer, it is natural for the algorithm to group them together.

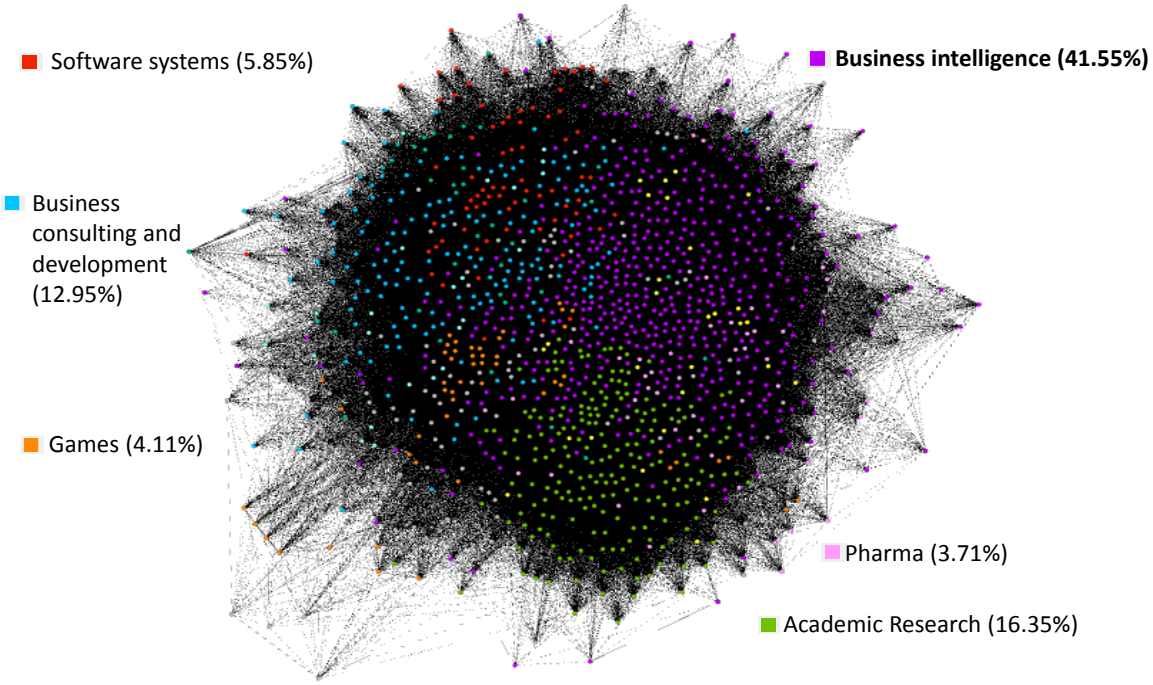


Figure 1: The network of Data Science job listings in English in Denmark (for the ALL_DS dataset), jobs linked by their similarity in content. The most important job types (in number) are given.

Methodology. We now give a brief outline of our "big data" approach. [3]

1. **Data collection.** We collected a dataset of job listings over 3-6 month periods, resulting in two separate datasets: one with a very restricted set of job listings to ensure relevance to Data Science, and a second smaller dataset covering a wider array of job listings but hand-annotated to ensure relevance to not only Data Science, but to Bachelors of Data Science as ITU would educate them in the proposed programme. We have made this dataset available.[4]

   The search words for the more restrictive dataset, covering the 6 months, are given in Table 1.

English			Danish	
data analyst	data analysis	predictive analysis	data analyse	dataanalyse
data scientist	data modeling	predictive modeling	chef analytiker	senior analytiker
data science	data warehouse	big data	maskinlæring	data modelering
data mining	senior analyst	business intelligence analyst	_bi	dataanalytiker
data engineer	chief analyst	knowledge discovery	data analytiker	
machine learning				

Table 1: Restrictive English and Danish keywords for data science job scraping. Note that most scientific terms for data science tend to appear in English, also in Danish texts. All keywords, from both languages, were used no matter the language of the job listing in question.

2. **Pre-processing.** We used an off-the-shelf language recogniser, `langdetect`[5] to separate these advertisements into either Danish or English (rejecting listings written in any other language). We carried out sentence and word tokenisation and calculated for each language's subset and each document of the subset the normalised tf-idf of the n-grams, where $n \in \{1, 2, 3\}$.

[3] The method used here is a simplified version of [2].
[4] link
[5] `https://pypi.python.org/pypi/langdetect?`

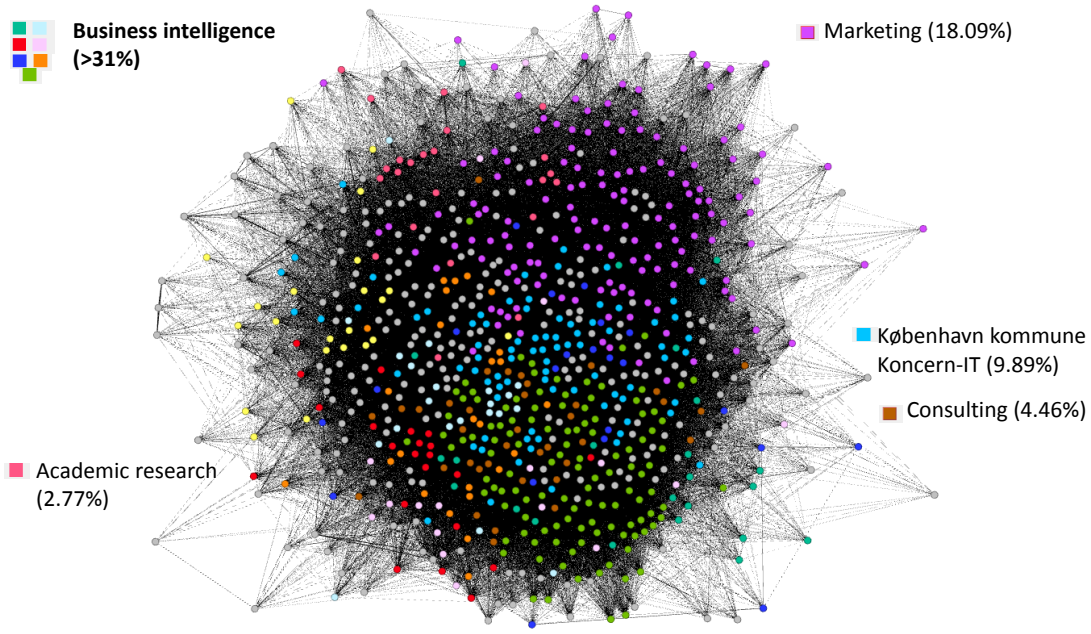


Figure 2: The network of Data Science job listings in Danish in Denmark (for the ALL_DS dataset), jobs linked by their similarity in content. The most important job types (in number) are given.

This resolves our first objective (O1): 1266 jobs in English and 829 jobs in Danish were extracted.

We now need to address objective (O2). We want to find out what kind of jobs are available and verify that they align with the offerings proposed by ITU in this application. We first try to represent job listings in a much more concise fashion, using keyword extraction. We then use these extracted keywords to compare job listings between each other and output natural groupings of job listings by topic area.

3. **Keyword extraction.** The normalised tf-idf score of a word given a document and that document's document set consists of two scores: (1) normalised term frequencey (tf), which is the frequency with which the word occurs in the document, normalised by the length of the document, and (2) inverse document frequency (idf), the inverse of the number of documents from the document set in which the term occurs. The idea is that words that are probably keywords for a document will have a higher tf-idf (See for example, [1]).

   For each document, any n-gram including a stopword was thereafter removed. The remaining n-grams form the set V, our vocabulary.

   For each job listing we discarded the lowest 85% scoring n-grams. We considered the remaining subset to be keywords for the job listing, and representative of it's unique contents. Each job listing can now be represented by a *score vector*, where element i is the score for the ith n-gram from V.

4. **Job listings network construction and clustering.** We now want to use these job listing representations to find the general types of job listings, according to topic matter.

   To do this, we constructed the job-listings network for the dataset, by taking each listing as a separate node and linking two nodes by an edge weighted by to the cosine similarity of their score vectors. The cosine similarity of the score vectors is meant to provide an indication on the similarity between the sets of keywords representing the two job listings.

	BUSINESS INTELLIGENCE	ACADEMIC RESEARCH	BUSINESS CONSULTING AND DEVELOPMENT	SOFTWARE SYSTEMS	GAMES	PHARMA
1	intelligence specialist	dtu	cognizant	cisco	moviestarplanet	clinical
2	business intelligence	big data analytics	digital	partner	game	cdm
3	nigel frank international	compute	client	adopt	data analyst	data
4	frank international limited	food supply chain	dep	psm	sales modeling	leo pharma
5	international limited københavn	phd	revenue	gcs	aps	biometrics
6	bi	project	enterprise content management	customer	global sales	pharma danmark
7	copenhagen nigel frank	university	business	expand	pricing	services department
8	specialist copenhagen nigel	chain data	portals	outcomes	friends	management deliverables
9	informatica	societal partnership	end	success	tweens	oversee
10	data warehouse	innovation	consulting services	emear	social network	trials

Table 2: Top 10 keywords per cluster for English 6 month un-annotated dataset.

We ran `infomap`[6], a state-of-the-art network clustering algorithm, over the resulting network to obtain the job listing clusters. Here a job listing cluster of a group of job listings that have similar descriptions (score vectors). The `infomap` attempted to construct the most simple description of the network in terms of groups of job listings (clusters) and found that the simplest description of the network was achieved by associating the job listings of each cluster that it output with each other.

5. **Examining the clusters.** For the largest clusters, we extracted the list of keywords for each cluster of documents, summing tf-idf scores across documents, and determined the cluster topic manually. As illustration, some cluster keywords for the six largest clusters are given in Table 2 for English. The keywords are automatically extracted to prevent possible overlap.

References

[1] *Unsupervised Approaches for Automatic Keyword Extraction Using Meeting Transcrips*, 2009.

[2] Natalie Schluter. An information-theoretic approach to supervised and unsupervised topic modelling. In preparation.

[6] http://www.mapequation.org/

Example job profiles

Profile 1: *Data Scientist*
Description: Generalist data scientist.
Example job: http://itu.dk/people/pagh/data-science/profile0.html

Profile 2: *Business Intelligence Specialist*
MSc degree: Digital Innovation and Management (ITU), Operations and Innovation Management (AAU-CPH)
Description: Works on defining and implementing business intelligence systems that bring value by giving insight into an organisation's data and enabling better decisions. Often existing frameworks are adapted to fit the organisation. Possibly in a global consulting company such as SAS Institute or IBM.
Example job: http://itu.dk/people/pagh/data-science/profile1.html

Profile 3: *Management Consultant*
MSc degree: Digital Innovation and Management (ITU), Operations and Innovation Management (AAU-CPH), Mathematical modeling and computation (DTU)
Description: This position requires some years of work experience. Work closely with decision-makers on modeling the effects of potential decisions to forecast what would yield the best outcome.
Example job: http://itu.dk/people/pagh/data-science/profile2.html

Profile 4: *Online Media Analyst*
MSc degree: Digital Design and Communication (ITU), IT & Cognition (KU), Operations and Innovation Management (AAU-CPH) Description: Help organisations understand what happens on social and online media related to their business, using Data Science tools to control the data avalanche.
Example job: http://itu.dk/people/pagh/data-science/profile3.html

Profile 5: *Game Data Analyst*
MSc degree: Games (ITU) or Software Development (ITU)
Description: Analyslyze data generated by computer games and work with game developers in order to e.g. improve gameplay, improve monetisation such as in-app purchases and ad-based revenue.
Example job: http://itu.dk/people/pagh/data-science/profile4.html

Profile 6: *Data Management Specialist*
MSc degree: Software Development (ITU) or computer science and engineering (DTU)
Description: Develops an organisation's data assets and the computational infrastructure needed to analyslyze them. Key employee in a larger organisation, or working as a consultant.
Example job: http://itu.dk/people/pagh/data-science/profile5.html

Profile 7: *Developer, Data-driven Software*
MSc degree: Software Development (ITU) or computer science and engineering (DTU)
Description: Develops new software products and platforms for Data Science, or more generally data-intensive systems.
Example job: http://itu.dk/people/pagh/data-science/profile6.html

Profile 8: *PhD in Data Science*
MSc degree: Programmes in disciplines where quantitative methods are used in research e.g. Digital Design and Communication (ITU), Digital Innovation and Management (ITU), Bioinformatics (KU),

Bioinformatics and systems biology (DTU), Digital Design and Communication (ITU), or Digital Innovation and Management (ITU), or alternatively a background in Software Development (ITU) and a PhD project on method development in Data Science.
Description: Use Data Science methods for scientific progress in a quantitative field, or a supporting field such as machine learning or algorithms.
Example job: http://itu.dk/people/pagh/data-science/profile7.html

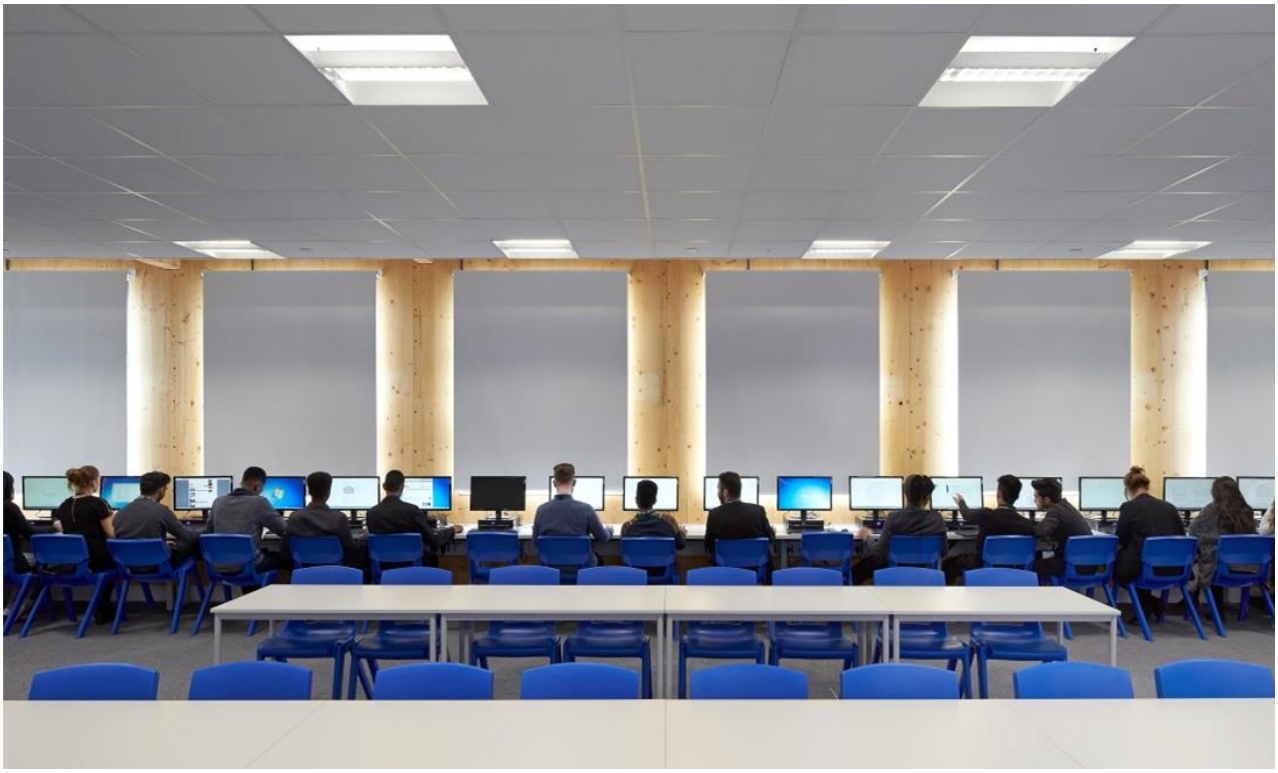
Profile 9: *Behavioral analyst*
MSc degree: Behavioral science; no suitable programme identified in DK, but internationally such programmes can be taken for example at Warwick or LSE.
Description: Use Data Science methods to study human behavior, in particular with the aim to make "behavioral design" that changes the behavior of groups of people (e.g. customers, citizens) in a desirable way.
Example job: Consultant at KL.7, see e.g. http://kl7.dk/cases.html

Til
IT-Universitetet

Dokumenttype
Afrapportering

Dato
Maj 2016

MARKEDSUNDERSØGELSE FOR DATA SCIENCE



RAMBØLL

INDHOLD

1.	**INDLEDNING**	**1**
1.1	Baggrund for undersøgelsen	1
1.2	Undersøgelsens tilrettelæggelse	1
1.3	Læsevejledning	3
2.	**SAMMENFATNING OG HOVEDKONKLUSIONER**	**4**
3.	**DELOPGAVE I: EKSISTERENDE UDDANNELSER INDEN FOR DATA SCIENCE**	**6**
3.1	Kortlægning af danske uddannelser	6
3.2	Et udpluk af anerkendte internationale uddannelser	10
4.	**DELOPGAVE II: EFTERSPØRGSLEN EFTER DATA SCIENCE**	**14**
4.1	Data science – del af en teknologisk udvikling	14
4.2	Data science rummer stort potentiale	16
4.3	En øget efterspørgsmål efter data scientists	17
4.4	De efterspurgte kompetencer	20

BILAG

Bilag 1 - De inkluderede studier

Bilag 2 - Detaljeret metodebeskrivelse

Bilag 3 – Oversigt over respondenter

Bilag 4 - Referenceliste

Rambøll
Hannemanns Allé 53
DK-2300 København S
T +45 5161 1000
F +45 5161 1001
www.ramboll.dk

1. INDLEDNING

I denne rapport præsenteres resultaterne af en markedsundersøgelse om uddannelser og arbejdsmarkedet for data science. Undersøgelsen er gennemført af Rambøll Management Consulting (herefter Rambøll) på opdrag af IT-Universitetet i perioden februar-maj 2016.

Formålet med undersøgelsen har været at kortlægge den eksisterende viden om og udbuddet for uddannelser inden for data science samt skabe et overblik over den forventede størrelse af arbejdsmarkedet og efterspørgslen på data scientists i Danmark. Markedsundersøgelsen skal samlet set kvalificere vidensgrundlaget for oprettelsen af en uddannelse indenfor data science på IT-Universitetet.

1.1 Baggrund for undersøgelsen

I takt med den teknologiske udvikling er flere og større datamængder – såkaldt Big Data – blevet tilgængelige. Den danske regering har efterspurgt en bedre udnyttelse af de muligheder, som Big Data giver, og her bliver data science yderst relevant.

Der findes mange forskellige definitioner af data science, men hovedparten definerer data science som et interdisciplinært felt, der handler om at udforske kompleks data – typisk Big Data – og udlede brugbar ny viden, der ellers ville ligge skjult. En data scientist beskrives typisk som en person, der formår at anvende sofistikerede kvantitative metoder til at forstå data og omsætte data i et forståeligt sprog til brugbar viden. Flere undersøgelser påpeger, at en data scientist skal besidde en række forskellige kompetencer, men i særlig grad have en profil med kompetencer inden for teknologi, matematik og forretningsforståelse.

Data science har et stort potentiale, og flere internationale såvel som nationale undersøgelser estimerer, at der vil være et større ønske om at arbejde med data på denne måde, hvorfor flere arbejdspladser inden for data science vil opstå i fremtiden.

De konkrete resultater, som flere undersøgelser mener, at et arbejde med Big Data vil kunne medføre, er bl.a. følgende:

- Virksomheder kan opnå øget kenskab til deres kunders behov og kommunikere direkte med kunden. Fx kan virksomheder, gennem analyser af Big Data, træffe beslutninger på baggrund af velfunderede forudsigelser om, hvad kunden vil få af behov.
- Virksomheder kan optimere sine produktions- eller tjenesteprocesser ved at trække på Big Data, hvilket kan lede til forbedringer af processerne og i sidste ende øget indtægt for virksomhederne.
- Analyser af Big Data kan give grobund for udviklingen af innovative og nyskabende forretningsmodeller til gavn for både virksomheder såvel som brugere/individer.

Der er således flere eksempler på, at data science allerede anvendes og i højere grad vil blive anvendt i fremtiden inden for flere brancher og sektorer – og både internationalt og nationalt. Dette er udgangspunktet for nærværende markedsundersøgelse og vil blive uddybet yderligere i rapportens to delopgaver.

1.2 Undersøgelsens tilrettelæggelse

Denne markedsundersøgelse har haft to fokusområder, der samlet set afdækker muligheden for at etablere en uddannelse inden for data science:

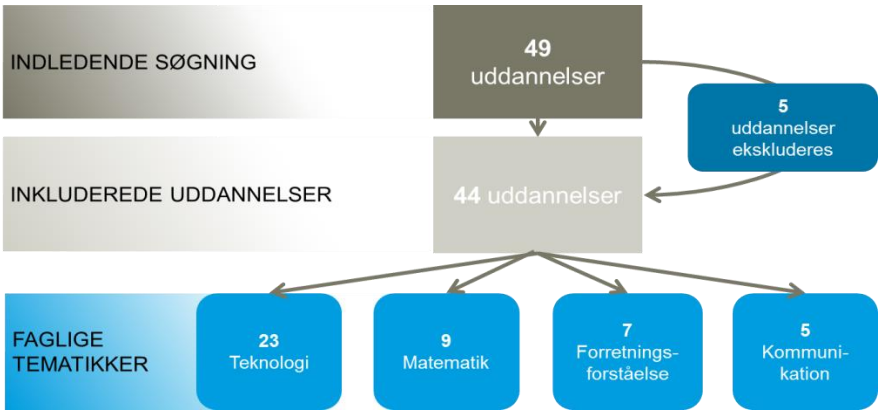
Fokusområde 1): En undersøgelse af udbuddet for uddannelser inden for data science

Fokusområde 2): En afdækning af den forventede størrelse af arbejdsmarkedet inden for data science med særlig vægt på en 5-10-årig horisont.

De to fokusområder i boksen ovenfor afdækkes hver især i henholdsvis delopgave I og delopgave II.

I *delopgave I* undersøger vi fokusområde 1, hvilket vil sige de eksisterende uddannelser i Danmark, der er centreret om eller har elementer af data science som del af deres curriculum. Til at gennemføre denne kortlægning har vi anvendt metoden *rapid evidence assessment,* hvilket har sikret en systematisk og effektiv kortlægning inden for en kort tidsramme (for en detaljeret metodebeskrivelse se bilag 2). Figuren nedenfor illustrerer processen fra antal uddannelser efter håndsøgningen, til inkluderede studier efter screeningen og klassificeringen af disse i de tematikker, der uddybes i rapporten.

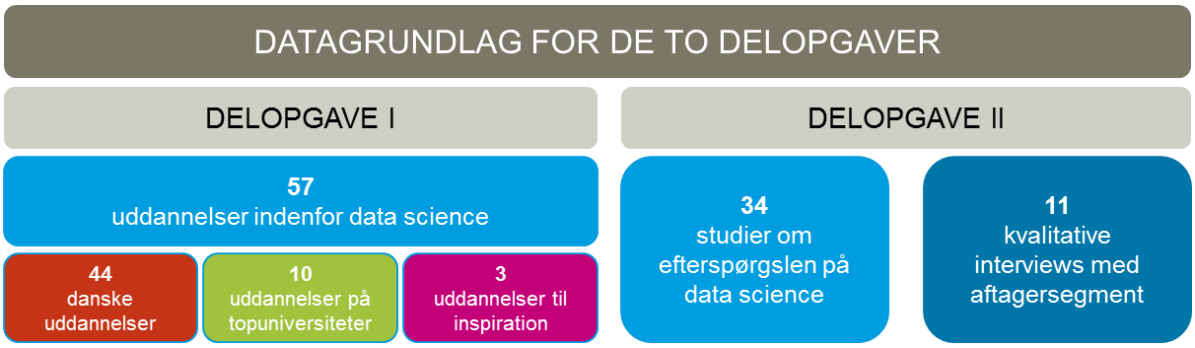
Figur 1: Overblik over screeningsproces og kategorisering



På baggrund af de inkluderede studier har vi foretaget en tværgående syntese, der identificerer tendenser og eventuelle 'huller' i de eksisterende danske uddannelser. Udover afdækning af de danske uddannelser har vi også foretaget en søgning på engelsksprogede internationale uddannelser inden for data science, hvor vi har valgt et fokus på de bedste internationale universiteters uddannelser inden for datalogi og informationsvidenskab. Derudover fremhæver vi tre specifikke internationale bacheloruddannelser i data science, som kan fungere som inspiration til en potentiel bacheloruddannelse på ITU.

I *delopgave II* afdækker vi fokusområde 2, som omhandler efterspørgslen efter data science. Til denne del af undersøgelsen har vi også foretaget en håndsøgning, der har resulteret i en række inkluderede studier. Da data science er et forholdsvis nyt felt, er mængden af relevante studier tilsvarende få, og de identificerede studier har primært karakter af undersøgelser og rapporter udarbejdet af konsulent – og analysevirksomheder. Kortlægningen af studier er dog blevet suppleret af 11 kvalitative interviews med repræsentanter fra forskellige brancher, sektorer og virksomheder i Danmark.

Figur 2: Datagrundlaget for markedsundersøgelsen



I figuren ovenfor ses en oversigt over markedsundersøgelsens samlede datagrundlag.

1.3 Læsevejledning

Ud over dette første kapitel består rapporten af tre øvrige kapitler samt to bilag.

Kapitel 2 – Sammenfatning og hovedkonklusioner: Indeholder en opsamling på tværs af begge rapportens delopgaver. Her uddybes de store linjer i rapporten med særligt fokus på de anvendelsesorienterede resultater og konklusioner.

Kapitel 3 – Delopgave I: Indeholder syntesen på tværs af de identificerede danske uddannelser, en syntese på tværs af uddannelserne på højt rangerende universiteter inden for informationsteknik og datalogi samt inspiration til indhold og struktur på en bacheloruddannelse i data science.

Kapitel 4 – Delopgave II: Indeholder den samlede analyse af efterspørgslen på data science. Kapitlet består af fire afsnit, der omhandler den nuværende situation på arbejdsmarkedet, potentialer ved data science, efterspørgslen efter data science og de konkrete kompetencer, der efterspørges.

2. SAMMENFATNING OG HOVEDKONKLUSIONER

Rambøll Management Consulting har gennemført en markedsundersøgelse for data science på opdrag af IT-Universitetet. Markedsundersøgelsen har haft til formål at afdække de eksisterende uddannelser inden for data science såvel som efterspørgslen på data scientists på arbejdsmarkedet i Danmark.

Få danske uddannelser understøtter den rette kombination af kompetencer
Kortlægningen af de eksisterende danske uddannelser inden for data science er udmundet i en gruppering af uddannelserne inden for følgende fire tematikker: Teknologi, matematik, forretningsforståelse og kommunikation. Kortlægningen viser, at få af de identificerede bacheloruddannelser har et curriculum, der direkte understøtter en uddannelse til data scientist. En af bacheloruddannelserne (Matematik og Teknologi på DTU) består af et curriculum, der giver den studerende en dybdegående tværfaglig forståelse, hvilket kan understøtte data science, og desuden kan den enkelte studerende tilpasse flere af bacheloruddannelserne, især inden for teknologi, så vedkommende får de rette kompetencer til at praktisere data science.

På nuværende tidspunkt er data science en specialisering på kandidatniveau
Kortlægningen af de eksisterende danske uddannelser viser, at data science i højere grad er en specialisering, som foretages på kandidatniveau, fremfor en vej som den enkelte vælger på bachelorniveau. De mest relevante danske uddannelser inden for data science har typisk en bacheloruddannelse, der giver de studerende generelle kompetencer inden for fx teknologi eller matematik, hvorefter den enkelte kan vælge at specialisere sig inden for data science på kandidatuddannelse.

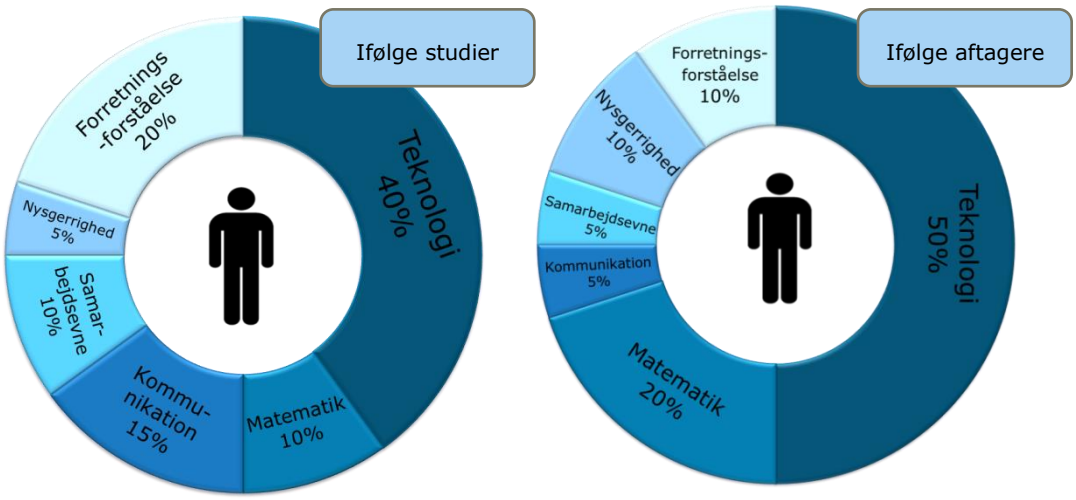
På tværs af de internationale uddannelser ses mange af de samme tendenser som ved de danske uddannelser: Data science er i højere grad en specialisering, fremfor udgangspunktet for en bacheloruddannelse. Dog er flere universiteter netop ved at udvikle deres eksisterende uddannelseskatalog yderligere ved bl.a. også at etablere bacheloruddannelser inden for data science for at imødekomme behovet for data scientists samt interessen fra de studerende.

Big Data og data science har uvist potentiale
På tværs af de inkluderede studier såvel som de kvalitative interviews med forskellige aftagersegmenter, er oplevelsen, at data science og Big Data er en del af en teknologisk udvikling, der er kommet for at blive. Udover at Big Data og data science er på den teknologiske dagsorden i indtil flere lande (primært USA og i Europa), understreger flere undersøgelser også, at det har stort potentiale for både at øge virksomheders produktivitet samt at give stor økonomisk værdi. Aftagersegmentet deler denne oplevelse, og mener, at Big Data og data science kan understøtte et øget kendskab til individets behov og adfærd samt medvirke til udviklingen af nye innovative forretningsmodeller. Flere af respondenterne udtrykker dog samtidig en vis usikkerhed over, *hvordan* dette potentiale udnyttes bedst muligt, og her håber mange, at data scientists kan hjælpe.

Efterspørgslen på data scientists er stor og stigende
På tværs af både studier såvel som interviews er det tydeligt, at der er en stor efterspørgsel på data scientists på nuværende tidspunkt både nationalt og internationalt, og alt tyder på, at denne efterspørgsel vil stige i fremtiden. Data scientists ansættes typisk inden for teknologibranchen, men det tyder på, at der vil komme en stigende efterspørgsel efter data science inden for den finansielle sektor, marketingbranchen samt inden for den offentlige sektor. Efterspørgslen efter data scientists i Danmark er i øjeblikket så høj, at flere virksomheder må opgive at besætte stillinger eller søge efter medarbejdere i udlandet (primært England og USA). Alle respondenter er enige om, at data scientists er yderst efterspurgte i Danmark på nuværende tidspunkt og 5-10 år frem i tiden. Hertil understreger flere, at det *"kun kan gå for langsomt"* at uddanne data scientists i Danmark.

En data scientist skal være nysgerrig på data og kunne formidle den
På baggrund af de inkluderede studier om efterspørgslen på data science samt de kvalitative interviews bliver det muligt at fremhæve de konkrete kompetencer, som en data scientist bør besidde.

Figur 3: Sammenligning af de efterspurgte kompetencer



Som figuren ovenfor viser, tegner der sig et billede af en data scientist, der har kompetencer inden for især teknologi, matematik og forretningsforståelse. Dog viser nærværende undersøgelse også, at studierne (der primært har amerikansk ophav) vægter de teknologiske kompetencer lavere end de danske respondenter, som omvendt vægter nysgerrighed og innovative kompetencer højere, end de står beskrevet i studierne. Man fristes derfor til at påpege, at de nuværende danske efterspurgte kompetencer kan ændre sig i takt med den teknologiske udvikling, hvorfor en uddannelse inden for data science bør være dynamisk og sikre, at de studerende klædes på inden for alle ovennævnte fagområder.

3. DELOPGAVE I: EKSISTERENDE UDDANNELSER INDEN FOR DATA SCIENCE

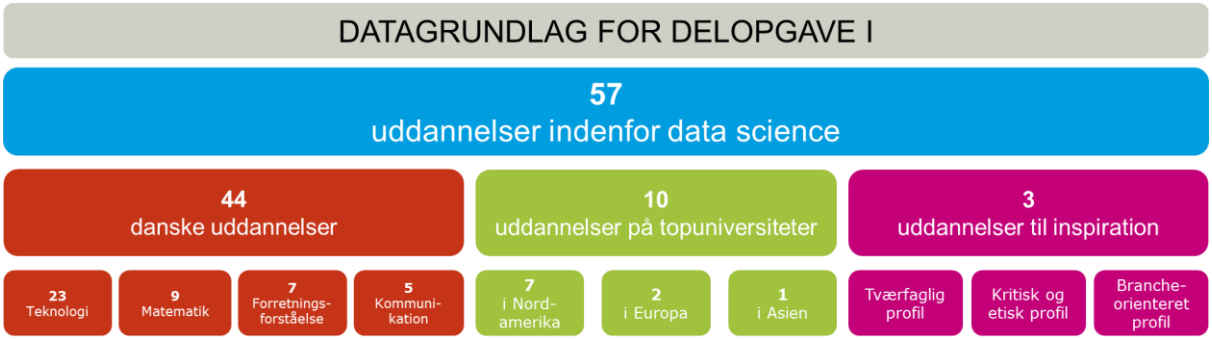
Formålet med denne del af rapporten er at kortlægge de eksisterende uddannelser inden for data science, og således arbejde med fokusområde 1: *En undersøgelse af udbuddet for uddannelser inden for data science*.

Kapitlet er inddelt i følgende tre afsnit:

- I *første afsnit* fremlægger vi syntesen af de identificerede danske uddannelser, der rummer centrale elementer af undervisning i data science.

- I *andet afsnit* præsenterer vi udvalgte internationale uddannelser inden for data science fra toprangerende universiteter samt giver tre konkrete eksempler på bacheloruddannelser, der kan være relevante at lade sig inspirere af i oprettelsen af en uddannelse inden for data science i Danmark.

Figur 4: Datagrundlag for delopgave I



I ovenstående figur ses datagrundlaget for delopgave I, der består af 57 uddannelser inden for data science, som fordeles på henholdsvis 44 danske uddannelser, 10 udvalgte uddannelser fra internationale topuniversiteter samt tre uddannelser, der er udvalgt som inspiration til oprettelsen af en bacheloruddannelse inden for data science.

3.1 Kortlægning af danske uddannelser

Rambøll har systematisk gennemgået danske uddannelser, der indeholder centrale aspekter af data science. Der er taget udgangspunkt i både bachelor- og kandidatuddannelser samt diplom-ingeniøruddannelser. Derudover er der medtaget enkelte relevante diplom-, efter- og videreuddannelser.

Vi har valgt at systematisere de identificerede uddannelser ud fra det indholdsmæssige fokus, de har. Således har vi grupperet uddannelserne i følgende faglige tematikker:

- Teknologi
- Matematik
- Forretningsforståelse
- Kommunikation.

Grupperingen er foretaget med udgangspunkt i de kompetencer, der er identificeret i delopgave II, som er relevante for arbejdet med data science. De inkluderede uddannelser indeholder alle elementer af teknologiske fag, som er påkrævede i arbejdet med Big Data[1], og grupperingen er

[1] Alle inkluderede uddannelser består bl.a. af teknologiske fag, da dette var et inklusionskriterie i kortlægningens søgestrategi.

således foretaget på baggrund af uddannelsens øvrige fokus. Uddannelser, der eksempelvis er placeret under *forretningsforståelse,* indeholder dermed teknologiske fag men har en hovedvægt af fag, der omhandler forretningsforståelse. Grupperingen i faglige tematikker bevirker, at vi kan påpege tværgående tendenser inden for hvert uddannelsesfelt, hvilket gør det muligt at identificere faglige 'huller', som en potentiel bacheloruddannelse på ITU vil kunne udfylde.

Uddannelser inden for teknologi

Det karakteriserende ved uddannelser inden for teknologi er, at hovedfokus i uddannelsen - hovedvægten af ECTS-point – lægges inden for fag som fx softwareudvikling, it, programmering mv. Uddannelser, der kun indeholder et eller to valgfag inden for teknologi, er således placeret under en af de andre faglige tematikker.

De inkluderede uddannelser:
4 bacheloruddannelser
8 bachelor – og kandidatuddannelser
4 kandidatuddannelser
7 diplom-, professions – eller efteruddannelser.

På tværs af de identificerede relevante bacheloruddannelser inden for teknologi ses en overvægt af bacheloruddannelser knyttet til uddannelsen til civilingeniør. For at blive civilingeniør på DTU kan den studerende eksempelvis vælge *Softwareteknologi* eller *Netværksteknologi og It*, fag, hvor de studerende bl.a. lærer at anvende avanceret datalogi og moderne matematik til at udvikle nye it-løsninger.

De inkluderede bacheloruddannelser indeholder få fag direkte målrettet arbejdet med data science, Big Data og/eller maskinlæring. Uddannelserne sigter primært mod at give de studerende grundlæggende evner inden for datalogi og software, og flere af dem indeholder en række matematiske og naturvidenskabelige fag, hvor de fx undervises i algoritmer, fysik eller matematisk modellering. Mængden af valgfag og muligheder for specialisering er begrænset på bacheloruddannelserne, og der er forholdsvis dårlig mulighed for at vælge et forretningsorienteret fag, der kan supplere de teknologiske fag. De nævnte bacheloruddannelser giver bl.a. adgang til kandidatuddannelsen i Datalogi og It eller Data Engineering på AU, Informationsteknologi på DTU eller Digitale Netværksstrategier på DTU. Både bachelor- såvel som kandidatuddannelserne på DTU kræver solide evner inden for teknologi. Som del af de fleste af de inkluderede kandidatuddannelser fra DTU kan den studerende specialisere sig i data science og Big Data. Her optjener den kandidatstuderende 45 ECTS-point gennem fag, der har fokus på både indsamling, behandling, analyse og forretningsprocesser.

En anden type bacheloruddannelse, der er værd at fremhæve, er bachelor- såvel som kandidatuddannelsen i *Datalogi* på en række af de danske universiteter (RUC, KU, SDU, AU og AAU). Uddannelsen i datalogi indeholder bl.a. fag, der lærer de studerende at analysere og løse komplekse problemer i forbindelse med it, samt fag i både programmering, softwarearkitektur, algoritmer, datastrukturer og systemudvikling. Indholdet i bacheloruddannelsen i datalogi fremstår forholdsvis ens på tværs af de forskellige uddannelsesinstitutioner, mens de efterfølgende kandidatuddannelser giver mulighed for specialisering inden for data science, enten i form af relevante valgfag på uddannelsen, eller som en formaliseret retning, hvilket er tilfældet på KU og AAU. Det kendetegner derfor kandidatuddannelserne i datalogi, at der er stor frihed til egen specialisering, og flere kandidater indeholder specifikke fag inden for data science.

Udover kandidatuddannelserne har vi også inkluderet en række relevante diplom-, professions- og efteruddannelser. Disse er karakteriseret ved et højere fokus på praksis ved programmering og softwareudvikling. Uddannelser sigter mod at lære de studerende specifikke kompetencer inden for databehandling, men giver ikke en mere generel indføring i datalogiske og teknologiske processer, som universitetsuddannelserne gør det.

På tværs af uddannelser inden for teknologi tegner der sig et billede af et specialiseret område, hvor de studerende får konkrete teknologiske værktøjer og viden, der gør dem i stand til at bearbejde større datamængder samt udvikle software og større it-systemer. Specifikke fag om data science, Big Data eller maskinlæring er ikke fremtrædende på bacheloruddannelserne. Bacheloruddannelserne giver dermed i højere grad generelle kompetencer inden for det teknologiske og datalogiske område, mens kandidatuddannelserne gør det muligt at specialisere sig inden for data science. Dog bør det fremhæves, at det især er de teknologiske fag, der dominerer, og der er begrænset mulighed for at tage fag inden for forretningsforståelse og tilegne sig andre ikketeknologiske kompetencer, der kan være relevante for en data scientist. Data science får derfor karakter af en teknologisk specialiseringsvej fremfor en hovedvej, der også indbefatter tilegnelsen af et bredere sæt kompetencer for de studerende.

Uddannelser inden for matematik

De inkluderede uddannelser:
1 bacheloruddannelse
6 bachelor – og kandidatuddannelser
2 kandidatuddannelser.

De inkluderede uddannelser inden for matematik har en overvægt af matematiske fag som fx differentialligninger, statistiske metoder, algebra og algoritmer mv.[2] Uddannelserne inden for matematik kan opdeles i tre grupper: De såkaldte klassiske matematiske uddannelser, uddannelser, der kombinerer matematik og teknologi, og uddannelser, der omhandler statistik.

På de klassiske uddannelser i matematik er hovedfokus på matematikkens kerneområder, fx algebra, geometri og differentialligninger. På bacheloruddannelserne har den studerende mulighed for at specialisere sig inden for datalogi, mens specialiseringerne på kandidatuddannelserne er knap så teknologiske. For at uddanne sig til at arbejde med Big Data kan det således være relevant for den studerende at tage en bacheloruddannelse i matematik, og kombinere denne med en kandidatuddannelse inden for teknologi, for dermed også at kunne arbejde med databaser og softwareudvikling.

To af de inkluderede uddannelser inden for matematik kombinerer matematik og teknologi i høj grad. På uddannelserne undervises de studerende i grundlæggende matematiske fag, men også i hvordan man benytter computerbaserede løsninger samt statistiske metoder. Således er der potentiale for, at de studerende tilegner sig flere af de kompetencer, som en data scientist skal besidde. Særligt en bacheloruddannelse på DTU indeholder en relevant specialiseringsmulighed, fx inden for computerintensive dataanalyser. På DTU vil det også være en relevant mulighed at læse en af deres øvrige kandidatuddannelser, som i endnu højere grad omhandler de teknologiske aspekter af data science (se fx uddannelser inkluderet i forrige afsnit).

I kortlægningen har vi også inkluderet to kandidatuddannelser inden for statistik, der indeholder fag inden for sandsynlighedsregning, statistiske teorier og praktisk datahåndtering. Uddannelserne opfylder flere af de matematiske krav, der typisk påkræves af en data scientist, men de er dog ikke nær så teknologisk tunge. Det er derfor begrænset, hvor store kompetencer uddannelserne giver til fx at udvikle software til behandling og håndtering af store datamængder. De naturlige bacheloruddannelser, der er adgangsgivende til kandidatuddannelserne i statistik, har ligeledes en overvægt af matematiske fremfor teknologisk fag.

På tværs af uddannelser inden for matematik tegner sig et billede af uddannelser, der opfylder de grundlæggende krav til en data scientists' kompetencer inden for avanceret matematisk forståelse. Dog er uddannelserne så specialiserede, at den studerende får matematiske kompetencer, der er mindre relevante for en data scientist. På flere af uddannelserne er der mulighed for at supplere matematikken med datalogiske og teknologiske fag, hvorved den studerende bed-

[2] Andre naturvidenskabelige uddannelser, som fx fysik, indeholder også en række matematiske fag, som benyttes inden for data science. Disse uddannelser er dog ekskluderet under screeningsprocessen grundet manglende relevans i uddannelsernes curriculum. Det skal dog bemærkes, at der blandt de nuværende data scientists er en mindre andel, som har en baggrund inden for fx fysik.

re opnår en profil som data scientist. Uddannelserne giver dog kun begrænsede kompetencer til at arbejde med ustruktureret data, hvilket er en af de centrale arbejdsopgaver inden for data science, og derudover er kompetencer inden for forretningsforståelse samt formidling af data næsten ikkeeksisterende på uddannelserne.

Uddannelser inden for forretningsforståelse

De inkluderede studier:
1 bacheloruddannelse
2 bachelor – og kandidatuddannelser
2 kandidatuddannelser.

De inkluderede uddannelser inden for forretningsforståelse er kendetegnet ved at have en overvægt af fag med fokus på forretningsforståelse. Disse fag kan fx handle om at undervise de studerende i ledelse og organisering, forandringer i organisationer, forretningsadministration og bredere samfundsvidenskabelige fag som fx politik. De teknologiske fag på uddannelserne kan fx omhandle programmering, design af databasesystemer og data mining, men flere af uddannelserne indeholder også fag, der omhandler forretningsintelligens og analyse, hvorved den studerende får viden om, hvordan data skaber værdig for virksomheden.

Bacheloruddannelsen Strategisk Analyse og Systemdesign på DTU er den mest teknologiske uddannelse inkluderet under denne tematik, da den fx introducerer til maskinlæring, og har derudover også matematiske fag om modellering. Uddannelsens forretningsfokus omhandler i høj grad forretningsintelligens (business intelligence), og det er en oplagt mulighed at kombinere denne uddannelse med en matematisk eller teknologisk kandidatuddannelse for at undgå en for 'overfladisk' profil, der risikerer ikke at leve op til kravene som data scientist.

Tilsvarende er både bachelor- og kandidatuddannelsen på CBS i Erhvervsøkonomi og It blandt de mindre teknologiske uddannelser. De teknologiske fag, der undervises i på uddannelsen, handler i højere grad om dataanalyse fremfor databehandling, hvorfor håndtering af ustruktureret data bl.a. ikke tillæres. Uddannelsen giver dog en bred erhvervsøkonomisk indsigt og indeholder et fag på kandidatdelen i Big Data management.

Kandidatuddannelserne har ikke særligt fokus på det teknologiske eller matematiske, hvorfor det er afgørende, at en kandidatstuderende på en uddannelse inden for forretningsforståelse har taget en bachelor i en matematisk eller teknologisk uddannelse for at kunne kalde sig data scientist. Flere af kandidatuddannelserne er dog netop målrettet bachelorer, der kommer fra andre fagområder end det forretningsmæssige, hvorfor der er fin mulighed herfor.

På tværs af uddannelser inden for forretningsforståelse tegner sig et billede af en række uddannelser, der ikke opfylder de teknologiske kompetencekrav, som er nødvendige for de fleste jobs inden for data science. Dette gælder især kompetencer inden for arbejdet med ustruktureret data og udvikling af softwaresystemer til datahåndtering. Omvendt giver uddannelserne en solid forretningsforståelse, der fremhæves som en vigtig kompetence i arbejdet med data science inden for forskellige brancher og sektorer.

Uddannelser inden for kommunikation

De inkluderede uddannelser:
3 bacheloruddannelser
1 bachelor – og kandidatuddannelse
3 kandidatuddannelser.

Uddannelserne inden for kommunikation indeholder typisk fag som kommunikations- og/eller organisationsanalyse samt design af it-systemer til kommunikation og forretning, udover også at indeholde forskellige teknologiske fag.

Med undtagelse af uddannelsen i Kommunikationsteknologi på AAU og DTU, er der begrænset vægt på, at de studerende tilegner sig teknologiske kompetencer på uddannelser inden for kommunikation. Dette selvom der på hovedparten af uddannelserne undervises i programmering.

Således kan man med fordel kombinere en kandidatuddannelse inden for kommunikation med en mere teknologisk og/eller matematisk bacheloruddannelse.

På kandidatuddannelserne inden for kommunikation er der flere muligheder for specialisering, fx indenfor it-ledelse, en mere kommunikationsteknologisk retning eller forretningsstrategier. Kandidatuddannelserne giver dermed også mulighed for at tilegne sig en vis forretningsforståelse.

Bacheloruddannelsen på AAU og kandidatuddannelsen på DTU i Informationsteknologi falder begge en smule uden for de andre uddannelser inden for kommunikation, da de er noget mere teknologiske. På kandidatuddannelsen på DTU er der blandt andet mulighed for at tage den tidligere nævnte specialisering inden for data science og Big Data, mens bacheloruddannelsen på AAU giver tilsvarende mulighed for at gennemføre projekter inden for en relevant specialisering. Eksempelvis skal de studerende gennemføre et projekt om database – og softwareudvikling på 4. semester af uddannelsen.

På tværs af uddannelser inden for kommunikation tegner sig dermed et billede af en række uddannelser, der ikke opfylder de teknologiske kompetencekrav, som en data scientist i høj grad bør besidde. I modsætning til uddannelserne inden for forretningsforståelse indeholder uddannelserne inden for kommunikation dog en del fag i softwareudvikling. Uddannelserne indeholder således fag, der understøtter relevante teknologiske kompetencer, dog i højere grad målrettet kommunikationssystemer frem for databaser. Flere af uddannelserne underviser ligeledes i ledelse og organisering, hvorfor de også giver de studerende kompetencer i en vis forretningsforståelse. Med undtagelse af uddannelsen i Informationsteknologi, er det dog oplagt, at den studerende kombinerer en uddannelse inden for kommunikation med en teknologisk eller matematisk uddannelse for at kunne arbejde som data scientist.

3.2 Et udpluk af anerkendte internationale uddannelser

Som del af kortlægningen af eksisterende uddannelser inden for data science, har Rambøll ligeledes undersøgt internationale, engelskprogede uddannelser inden for data science. I afdækningen af internationale uddannelser har det været nødvendigt at afgrænse søgningen markant grundet det omfangsrige antal af relevante uddannelser.

Vi har derfor valgt et fokus på de ti bedste internationale universiteters uddannelser inden for datalogi og informationsvidenskab, som er udvalgt på baggrund af *QS World University Rankings*[3] (se bilag 1 for yderligere metodebeskrivelse). I det følgende afsnit opridser vi de overordnede tendenser på tværs af uddannelserne på de ti internationale universiteter.

Tabel 1: Oversigt over ti topuniversiteters uddannelser inden for data science e.l.

Rangering	Universitet	Geografisk placering	Uddannelsestitel
1	Massachusetts Institute of Technology (MIT)	USA	Hovedfag i *Business Analytics* Hovedfag i *Computer Science and Engineering* Hovedfag i *Mathematics and Computer Science* Sidefag i *Statistics and Data Science*
2	Stanford University	USA	Hovedfag i *Computer Science* Hovedfag i *Management and Computer Science* Kandidat i *Statistics: Data Science*
3	Oxford University	England	Bachelor i *Mathematics and Computer Science*
4	Harvard University	USA	Bachelor i *Computer Science*
5	Carnegie Mellon University	USA	Hovedfag i *Information System* Hovedfag i *Statistics and Machine Learning*
6	University of Cambridge	England	Bachelor og engelsk kandidat i *Engineering* Bachelor i *Computer Science*
7	Berkeley – University of California	USA	Bachelor i *Computer Science and Engineering* *Bachelor* i *Data Science*
8	National University of Singapore	Singapore	Engelsk bachelor i *Computer Engineering*
9	Princeton University	USA	Bachelor i *Computer Science* (ingeniøruddannelse)

[3] Ranglisten udgives af det britiske firma Quacquarelli Symonds, der er specialiseret i uddannelse, sammen med det hollandske akademiske forlag Elsevier, og indeholder mere end 800 universiteter.

10	University of Toronto	Canada	Bachelor i *Computer Science*

Som tabellen ovenfor viser, foregår hovedparten af uddannelserne i USA, hvilket må antages at hænge sammen med den teknologiske udvikling inden for både data science og Big Data, der er længere fremme i USA end fx i Europa (se afsnit 4.1).

En begrænset andel af bacheloruddannelserne på de bedste, engelsksprogede universiteter inden for datalogi og informationssystemer indeholder et særskilt fokus på data science og Big Data. Dette kan skyldes, at de medtagne universiteter har stærke faglige traditioner og discipliner, og en ny faglighed, som fx data science, først skal modnes, før universiteterne etablerer uddannelser inden for området. På Berkeley Universitet har man dog netop (pilothold i efterår 2015, første undervisningshold i forår 2016) etableret en bacheloruddannelse med specifikt fokus på data science som en tværfaglig uddannelse. Uddannelsen er udviklet af en task force på tværs af fakulteter på Berkeley, og den består af både undervisning i teknologi, statistik, dataetik, visualisering af data samt fag til at understøtte samfunds – såvel som forretningsforståelse. Hovedparten af de øvrige medtagne uddannelser er inden for datalogi, matematik eller statistik samt teknologiske ingeniøruddannelser. På de fleste af disse uddannelser er der mulighed for at specialisere sig inden for data science eller tage relevante valgfag, som fx maskinlæring, data mining eller valgfag i data science og Big Data. Fælles for hovedparten af de internationale uddannelser er, at de er både teknologisk og matematisk krævende. To uddannelser, der skiller sig ud fra den hårde teknologiske profil, er 1) uddannelsen i *Information System* på Carnegie Mellon University, som er organiseret blandt de humanistiske uddannelser samt giver et bredere sæt af kompetencer, samt 2) uddannelsen i *Business Analytics* på Massachusetts Institute of Technology, der udover klassiske fag inden for forretningsanalyse, supplerer med fag i programmering, statistik, maskinlæring og data mining for netop at sikre en bredere faglig profil, der således også kan det teknologiske og matematiske.

Kortlægningen af internationale uddannelser har haft fokus på uddannelser på bachelorniveau. Dog forekommer der flere relevante kandidatuddannelser inden for data science på flere af universiteterne. På hovedfaget i *Statistics and Machine Learning* på Carnegie Mellon University er fokus på klassiske fag i matematik og statistik, der suppleres af fag, der specifikt omhandler data science og maskinlæring. På samme vis har Stanford University en kandidatuddannelse i *Statistics: Data Science*, som primært indeholder fag i matematik, statistik og computervidenskab, men også giver mulighed for at tage en lang række valgfag, der sikrer en klar specialisering.

Det bør desuden nævnes, at flere af de udvalgte universiteter, ligesom Berkeley Universitet har gjort det, er i færd med at udvikle flere relevante uddannelser og øvrige initiativer inden for data science i fremtiden. Massachusetts Institute of Technology har fx fornyligt oprettet valgfag (sidefag) i statistik og data science, mens National University of Singapore har annonceret, at de opretter en bacheloruddannelse i Data Science og Analytics til august 2016.

På tværs af de internationale uddannelser ses mange af de samme tendenser, som angivet i kortlægningen af de danske uddannelser. Data science er i højere grad en specialisering fremfor udgangspunktet for en bacheloruddannelse. Dog viser kortlægningen af uddannelser også, at flere topuniversiteter er ved at – eller netop har – etableret uddannelser med fokus på data science, og flere af universiteterne tilbyder allerede adskillige online-uddannelser og -kurser i data science. Det er således tydeligt, at der etableres flere og flere fag – såvel som deciderede uddannelser – inden for data science, hvilket må antages som del af et forsøg på at efterkomme en øget efterspørgsel.

Eksempler til inspiration
Med henblik på at understøtte tilrettelæggelsen af en bacheloruddannelse i Data Science har Rambøll identificereret tre konkrete internationale bacheloruddannelser, som kan fungere som inspiration. De tre eksempler er udvalgt, fordi de hver især har en særlig data science-profil, og eksemplerne kan således fungere som inspiration til, hvordan en bacheloruddannelse i Data Science kan tilrettelægges på forskellige måder.

Figur 5: Første eksempel til inspiration

Bachelor i *Data Science and Analytics* på University of Essex, England

University of Essex udbyder en treårig bachelor i *Data Science and Analytics*. Det interessante ved denne uddannelse er, at data science beskrives som en tværfaglig disciplin, der handler om at give de studerende specialiserede evner inden for flere forskellige discipliner. Hovedparten af fagene ligger inden for afdelingen for computervidenskab, men uddannelsen består også af fag på afdelingen for matematik samt afdelingen for forretning (Essex Business School), hvilket dermed skal sikre disse tværfaglige kompetencer.

På uddannelsens første år bliver de studerende introduceret for fag i programmering, databaser og statistik, samt web-udvikling og diskret matematik, mens de på andet år også får fag i informationssøgning, kunstig intelligens, maskinlæring samt lærer, hvordan man løser moderne ledelsesproblemer gennem simuleringsmodeller (*Quantitative Management*). På både første og andet år kan de studerende desuden vælge moduler i fx matematik ift. data science, anvendt statistik, softwareudvikling samt data strukturer og algoritmer. På det sidste år er der ingen faste moduler, men derimod god mulighed for at sammensætte den specialisering inden for data science, som den studerende selv ønsker.

Udover ovenstående fag har uddannelsen fokus på projektledelse samt det at arbejde i en forretningsmæssig og organisatorisk kontekst. Under uddannelsen indgår de studerende således også i flere omgange af gruppearbejde, hvor de arbejder med konkrete projekter, der understøtter kompetencer i både samarbejde samt praksisorienteret data science.

Efter en afsluttet bachelor er der mulighed for at fortsætte på en kandidatuddannelse i Data Science på University of Essex (Link: Bachelor i Data Science and Analytics).

Figur 6: Andet eksempel til inspiration

Bachelor i Statistik og Dataanalyse på Linköping Universitet, Sverige

Linköping Universitet i Sverige udbyder en bacheloruddannelse i Statistik og Dataanalyse. Det interessante ved denne uddannelse, udover at den foregår i Skandinavien, er, at den har specifikt fokus på kritisk og selvstændig tænkning i forhold til data science.

Over halvdelen af uddannelsen består af fag i statistik, hvilket gør den yderst matematisk. På uddannelsens første år introduceres de studerende til den statistiske metodik og har fag i både programmering samt visuel dataanalyse. Førsteåret indeholder desuden undervisning i videnskabsteori og dataetik, hvor de studerende lærer, hvordan viden skabes samt præsenteres på en etisk måde. På andet år undervises i fag som matematisk analyse, grundlæggende SAS-programmering og design samt programmering af databaser, mens der på det tredje år er fag i statistisk teori, multivariate metoder samt statistisk analyse af komplekse data og data mining.

Bacheloruddannelsens naturlige overbygning er en engelsksproget kandidatuddannelse i statistik og data mining, der blandt andet omhandler maskinlæring, data mining, kunstig intelligens og visualisering (Link: Bachelor i statistik og dataanalyse).

Figur 7: Tredje eksempel til inspiration

Bachelor i *Data Science* fra University of Rochester

University of Rochester har investeret mere end $50 millioner i at udvide deres arbejde med Big Data og har forpligtet sig til at investere yderligere $50 millioner i fremtiden. Disse investeringer har blandt andet resulteret i oprettelsen af to bachelorer i Data Science, der begge varer fire år. Udover universitetets massive fokus på Big Data, er det interessante ved uddannelserne, at der er mulighed for at specialisere sig inden for forskellige anvendelsesområder. Uddannelserne baserer sig på datalogi og statistik, men giver mulighed for, at den studerende kan tage avancerede datalogiske fag inden for et anvendelsesområde, som fx forretning, biologi, miljø eller statskundskab.

På første år er det obligatorisk at følge fag i programmering, datastrukturer og regning, mens der på andet år undervises i lineær algebra, datalogisk statistik, datalogiske metoder, samt gives en introduktion til kunstig intelligens. På tredje år undervises i data mining og databasesystemer, mens det fjerde år indeholder fag i design og analyse af algoritmer samt et afsluttende fag i data science.

Ved siden af de obligatoriske fag tilføjes fag, der ligger inden for den specialisering, som den studerende har valgt. Disse fag følges på alle fire år. Inden for specialiseringen i økonomi og forretning er blandt andet fag i mikro- og makroøkonomi, spilteori samt finansiering og marketing, mens der på specialiseringen i datalogi, statistik og matematik blandt andet er fag inden for maskinlæring, programmeringssprog og matematisk statistik.

Rochester udbyder ligeledes en kandidatuddannelse i Data Science (Link: BS/BA i Data Science).

4. DELOPGAVE II: EFTERSPØRGSLEN EFTER DATA SCIENCE

Formålet med delopgave II er at skabe et overblik over den nuværende og kommende efterspørgsmål på data science – data scientists – på arbejdsmarkedet. Herunder hvilke konkrete kompetencer inden for data science, der efterspørges i Danmark. Delopgave II er afdækket gennem en desk research med specifikt fokus på efterspørgslen på data science internationalt set samt kvalitative interviews med udvalgte repræsentanter fra det danske erhvervsliv.

Da data science er et forholdsvis nyt marked og deraf også undersøgelsesfelt, findes der begrænsede mængder af international og national litteratur af høj kvalitet på området. Desk researchen indbefatter således primært internationale undersøgelser, typisk udformet af konsulenthuse i USA, og ganske få danske og nordiske undersøgelser. De kvalitative interviews anvendes løbende til at aktualisere og forbinde undersøgelserne til en dansk kontekst, og de uddyber samtidigt den efterspørgsel på data scientists samt de specifikke behov for kompetencer, som ses i det danske erhvervsliv.

Figur 8: Datagrundlag for delopgave II



* Andet refererer til undersøgelser fra henholdsvis World Economic Forum og OECD.

I ovenstående figur ses datagrundlaget for delopgave II, der består af 43 internationale og nationale studier om efterspørgslen på data science, samt 11 kvalitative interviews med et aftagersegment af respondenter fra forskellige brancher, sektorer og virksomheder.

4.1 Data science – del af en teknologisk udvikling

På tværs af de inkluderede studier fremhæves det, at data science er opstået som del af en større teknologisk udvikling inden for de sidste 5-10 år. Nærmere bestemt, har digitaliseringen og den øgede computerkraft medvirket til, at virksomheder skaber **flere data end nogensinde før**. Allerede i 2012 meddelte IBM, at der på to år var produceret 90 pct. af al data i verden, og at datamængden ville forsætte med at vokse eksponentielt(IBM, 2012). Denne vækst estimerer McKinsey Global Institute til at være 40 pct. per år(Mckinsey, 2011). Væksten i data har ligeledes fået OECD og World Economic Forum til at anbefale lande at satse på Big Data fremadrettet (OECD, 2015; World Economic Forum, 2012).

Big data og data science har også fået **stort politisk fokus**. I 2014 blev der på foranledning af den amerikanske præsident Barack Obama udarbejdet en rapport om data science, der beskriver Big Data som en social, økonomisk og teknologisk revolution (White House, 2014). I 2014 lavede EU ligeledes en ny strategi for Big Data for at støtte og accelerere transformationen mod datadreven økonomi i Europa. Strategien omhandler initiativer som fx at udvide infrastrukturen, offentlige indkøb og en øget markedsgørelse af datateknologier. Årsagen til denne strategi var bl.a.

at sikre Europas position inden for udviklingen af arbejdet med Big Data samt for at understøtte, at flere europæiske virksomheder påbegynder arbejdet hermed.[4]
Den danske regering er opmærksom på potentialet for vækst gennem Big Data. Erhvervsstyrelsen har i slutningen af 2013 offentliggjort en rapport, der beskriver, hvordan man kan fremme potentialet for Big Data i Danmark (Erhvervsstyrelsen, 2013). Analysen er baseret på interviews og casestudier blandt "best practice-virksomheder". Virksomhederne efterspørger mere uddannelse af medarbejderne, både grunduddannelser og efteruddannelser, lettere adgang til offentlige data, og desuden vurderes offentlig regulering og kampagneaktiviteter at have betydning for udbredelsen og udnyttelsen af Big Data.

Den amerikanske markedsrådgivningsvirksomhed International Data Corporation(IDC) nedsatte i 2011 en tværfaglig interessegruppe af personer inden for it-branchen til at varetage både forskning samt gennemføre drøftelser med analytikere, leverandører og købere inden for området (IDC, 2011). På den baggrund forudsagde IDC, at Big Data-teknologi og dertilhørende service vil vokse globalt med ca. 40 pct. årligt i perioden fra 2010 til 2015. Dette er syv gange mere end det generelle marked for informations- og kommunikationsteknologi vokser, hvorfor udviklingen må siges at være markant. Denne antagelse har IDC opdateret til ca. 23 pct. årlig global vækst inden for Big Data i perioden 2014-2019 (IDC, 2015), hvilket dog stadigvæk er seks gange større end det generelle marked for informations – og kommunikationsteknologi.

"Udviklingen er ikke overraskende men velkommen"
I aftagersegmentet spænder *kendskabet* til data science og Big Data vidt; fra respondenter, der arbejder med Big Data til daglig, og har rig erfaring med det, til respondenter, der har hørt om Big Data, men er usikre på, hvad Big Data er og kan. På tværs af alle respondenter er dog en klar tendens til, at Big Data og arbejdet hermed er en ny udvikling, der ses indenfor både det offentlige og private såvel som i både store og små virksomheder.

Flere af respondenterne beskriver arbejdet med Big Data som en udvikling til at arbejde med ustruktureret fremfor struktureret data i forlængelse af den stigende mængde af tilgængelige data. Udviklingen inden for Big Data og data science opleves ikke hverken overraskende eller banebrydende, men ses tværtimod som en velkommen udvikling, der ligger i forlængelse af de eksisterende tendenser inden for datadrevet innovation, kunstig intelligens og et stigende fokus på forbrugeren (mere herom afsnit 4.2).

Aftagersegmentet er vidt forskellige steder i forhold til at *arbejde* med Big Data. I den ene ende af spektret ses virksomheder, der lever af at levere analyser af Big Data til andre virksomheder, og det er således virksomheder, der har rig erfaring med og kendskab til data science. I midten af spektret ses den del af virksomhederne, der lige har – eller er ved at – etablere en afdeling, der skal specialisere sig i Big Data, da de kan se et potentiale heri. Citatet i boksen til højre er fra en repræsentant i en større servicevirksomhed, og indikerer den modsatte ende af spektret: virksomheder, der kender til Big Data-analyser og data science, men endnu mangler indblik i potentialet ved at arbejde med det eller ikke finder det aktuelt på nuværende tidspunkt.

> *"Jeg følger med i, hvad der sker på sidelinjen. Det er ikke relevant for os endnu, men jeg er meget spændt på Microsofts' implementering af R-programmering, da det er et skridt i den rigtige retning, synes jeg."*

Gennemgående for alle i aftagersegmentet er således, at de oplever en udvikling i retning af både at ville og kunne arbejde med Big Data og data science i Danmark. Derfor har hovedparten af aftagersegmentets virksomheder enten iværksat initiativer til at følge udviklingens gang, som fx

[4] Ifølge Europa-Kommissionen er kun 2 ud af 20 virksomheder, der tjener penge på Big Data, europæiske (Europa-Kommissionen, 2016; SIRA 2014).

at etablere en afdeling med dette formål, eller også følger de udviklingen fra sidelinjen for at holde sig opdateret.

4.2 Data science rummer stort potentiale

Arbejdet med Big Data og data science har en række potentialer. Big Data-analyser har blandt andet potentiale for at sikre øget kendskab til kundeadfærd og behov, understøtte en optimering af produktions – eller tjenesteprocesser samt danne grobund for udvikling af nye forretningsmodeller.

På tværs af studierne ses det, at efterspørgslen efter data science er drevet af **potentialet for øget produktivitet**. I en række analyser foretaget af Andrew McAffee og Erik Bynjolfsson fra MIT konkluderes det, at de nordamerikanske virksomheder, der anvender datadrevne beslutninger samt ligger i den bedste tredjedel inden for deres industri, i gennemsnit er 5 pct. mere produktive samt 6 pct. mere profitable end deres konkurrenter(Harvard Business Review, 2012).

Flere af de inkluderede undersøgelser fremhæver især **det økonomiske potentiale** i at arbejde med Big Data og data science. I en rapport fra McKinsey Global Institute forudsiges det, at arbejdet med Big Data kan give en potentiel årlig værdi på 250 milliarder dollars til den europæiske offentlige sektor. En forudsigelse, der bekræftes i en rapport i 2013 (McKinsey, 2011; McKinsey 2013).

Det store økonomiske potentiale ved at arbejde med Big Data styrkes yderligere af en undersøgelse fra Warsaw Institute for Economic Studies (2014). I undersøgelsen er foretaget en analyse af de makroøkonomiske effekter af Big Datas indtog i Europa, baseret på økonomiske nøglemål (WISE, 2014). Her sammenlignes den makroøkonomiske effekt af Big Data på 28 EU-landes økonomi. Analysen viser, at den europæiske økonomi i gennemsnit vil stige med ekstra 1,9 pct. i 2020 som resultat af arbejdet med Big Data. Ifølge analysen vil især Irland, Storbritannien, Luxembourg og Sverige opleve denne vækst. Danmark placerer sig i den øverste top ti af de lande, der vil opleve en positiv vækst i økonomien som følge af et stigende fokus på og arbejde med Big Data. Dette må forventes at influere efterspørgslen på medarbejdere, der kan arbejde med Big Data i Europa.

Der findes ikke et præcist estimat over antallet af danske arbejdspladser inden for data science. Estimatet over den makroøkonomiske gevinst ved Big Data i Danmark, fra studiet af Warsaw Institute for Economic Studies, er det bedste bud på den fremtidige danske situation ift. data science. Dog bør det fremhæves, at It- og Telestyrelsen i 2011, med hjælp fra rådgivningsfirmaet Zangenberg & Company, foretog interviews med repræsentanter for seks forskellige brancher i Danmark om Big Data (Zangenberg & Company, 2011). Resultatet heraf var, at der opleves et stort potentiale i at arbejde med Big Data i Danmark, hvilket bl.a. skyldes de store mængder af offentlig data. I rapporten opregnes et potentiale på op til 25 milliarder kroner, og alene 20 milliarder i energi- og byggesektoren, ved udlevering af offentlige data til erhvervslivet.[5]

I en spørgeskemaundersøgelse blandt virksomheder i Europa fra 2013, foretaget af det internationale datacenter Interxion (Interxion, 2013), konkluderes dog, at danske virksomheder ikke har været gode nok til at udnytte potentialet for Big Data. Ifølge analysen skyldes dette, at Danmark er et lille marked, hvor virksomheder er mindre tilbøjelige til at have en strategisk plan for området. Ligeledes viser analysen, at færre danske virksomheder betragter Big Data som en prioritet sammenlignet med virksomheder i større europæiske lande (Interxion, 2013). De danske virksomheder har således **ikke formået at udnytte potentialet** for Big Data, hvilket bl.a. kan skyldes uvidenhed om mulighederne ved analyser af Big Data – også på et mindre marked.

[5] Her bør dog påpeges, som det gøres i rapporten, at dette tal ikke er et fyldestgørende og endeligt billede af potentialet, da potentialet først vil materialisere sig, når virksomhederne for alvor formår at anvende data (Zangenberg & Company, 2011).

Det fremgår af Regeringens redegørelse om Danmarks digitale vækst for 2015, at dansk erhvervsliv kan styrke deres produktivitet og konkurrenceevne gennem større anvendelse af Big Data (Regeringen, 2015). Derfor iværksatte den tidligere regering, bl.a. ved aftalen om vækstplan for digitalisering i Danmark i 2015 (Aftaletekst, 2015), en række initiativer for at fremme brugen af Big Data.

"Big Data og data science har kæmpe potentiale"
På tværs af alle respondenter i aftagersegmentet er der bred enighed om, at Big Data og data science har kæmpe forretningsmæssigt potentiale. Flere af respondenterne fremhæver både Google og Amazon som forbilledlige frontløbere, der har formået at optimere deres forretningsmodel og kundebase, fordi de har arbejdet med Big Data og data science.

Især to typer af potentiale ved Big Data-analyser og data science fremhæves af respondenterne:

1. En øget indsigt i forbrugeren/kunden/klienten – og de behov, denne måtte have.
2. En optimering og innovation af forretningen.

I henhold til det førstnævnte potentiale fortæller flere af respondenterne, at udviklingen inden for arbejdet med Big Data og data science i Danmark også hænger sammen med et *øget fokus på individet*. Citatet i boksen, fra en respondent i en større servicevirksomhed, indikerer, at Big Data og data science giver mulighed for at få indblik i målgruppens adfærd, hvilket understøtter udviklingen af et produkt, der rammer behov i højere grad og dermed bliver af højere værdi for kunden. Ved at udvikle bedre målrettede produkter og services til forbrugeren opnår virksomheden også en højere indtjening, som flere af respondenterne forklarer.

"Vi bruger i højere grad data til beslutningsstøtte, og Big Data kan hjælpe os endnu tættere på forbrugerens behov, da vi kan se nærmere på deres opførsel."

Udover et potentiale for at komme tættere på forbrugerens behov, fortæller flere af respondenterne, at Big Data-analyser og data science også understøtter en *optimering og innovation af forretningen*. Citatet nedenfor er fra en respondent fra den finansielle sektor. Vedkommende mener, som flere af de øvrige respondenter, at data science kan understøtte en højere grad af kreativ og innovativ tænkning, der kan bidrage til udvikling af nye og bedre forretningsmodeller. Flere og flere større virksomheder har store mængder tilgængeligt data, men udnytter det kun i begrænset omfang, hvorfor der er stort potentiale i at ansætte medarbejdere, der kan anvende data og ændre den eksisterende forretningsmodel.

"Ved at arbejde mere eksplorativt og udforske data på en anden måde, kan nye idéer jo opstå. Idéer, som kan være med til at udvikle vores forretning."

På tværs af alle respondenter er det således tydeligt, at alle oplever et stort potentiale i arbejdet med Big Data og data science. Hovedparten af respondenterne har kendskab til Big Data og data science, men det tyder på, at der blandt nogle er stor tvivl om, *hvordan* potentialet udnyttes bedst muligt.

4.3 En øget efterspørgsmål efter data scientists

På baggrund af de inkluderede studier samt gennemførte interviews er det tydeligt, at der opleves et stærkt behov for ansatte, der formår at omsætte data til værdi for virksomheder både internationalt og nationalt set.

Det mest citerede konkrete bud på den potentielle efterspørgsel efter personer med evner til at analysere Big Data er en rapport fra McKinsey Global Institute fra 2011, hvor de har studeret Big Data i fem sektorer: Sundhedssektoren i USA, den offentlige sektor i Europa, detailhandlen i USA, samt fremstillingsvirksomheder og personlig lokaliseringsdata globalt set (McKinsey, 2011).

Rapporten forudsiger, at der i 2018 er potentiale for jobs i USA for mellem 140.000 til 190.000 flere personer med analytiske evner inden for Big Data, og 1,5 millioner flere managere, der formår at analysere de indsamlede data(McKinsey, 2013).

Dette bekræftes af flere andre undersøgelser, bl.a. i en artikel af Thomas H. Davenport, en af de mest anerkendte analytikere og forskere inden for business og teknologi, der sammen med data scientist D.J. Patil beskriver, at der er **stor mangel på arbejdskraft inden for data science**, og at dette kun vil fortsætte i fremtiden (Harvard Business Review, 2012). I deres årlige ledelsesanalyse for 2015, der er baseret på interviews med 430 virksomhedsledere fra 10 lande samt 30 industrier, konkluderer det globale konsulentfirma A. T. Kearney ligeledes, at 43 pct. af virksomhederne oplever, at mindst 10 pct. af deres stillinger inden for det digitale og analytiske arbejde ikke bliver besat, mens 33 pct. af virksomhederne forventer en stigning i efterspørgslen efter netop denne type arbejde inden for de næste 5 år (A. T. Kearney, 2015). Det tyder således på at være en global tendens, at der er efterspørgsel på data scientists.

Da USA er længere i udviklingen inden for Big Data-analyse og data science, ses hovedparten af efterspørgslen på data scientists især i USA. Præcise estimater for, hvor stor den europæiske eller danske efterspørgsel er, er ikke mulige at fremlægge. Det mest konkrete, samt mest citerede, europæiske estimat omhandler Storbritannien, hvor SAS og Tech Partnership i 2014 har foretaget en spørgeskemaundersøgelse af Big Data-stillinger i Storbritannien og de ansatte heri (SAS, 2014). Undersøgelsen beskriver og forudsiger udviklingen fra 2013 til 2020, hvor der ses en voldsom stigning i efterspørgslen inden for Big Data-ansatte i Storbritannien, som fra 2008-2013 er steget med 41 pct. årligt (fra 1.800 ansatte i 2008 til 21.400 ansatte i 2013). Til sammenligning er efterspørgslen efter it og data warehouse/business intelligence faldet med hhv. 9 pct. og 6. pct. i samme periode. Udviklingen fra 2013 til 2020 forventes at fortsætte med en samlet stigning på 160 pct. jobmuligheder mod en forventning om en generel beskæftigelsesfremgang på 6 pct. i samme periode. Det giver 56.000 jobmuligheder per år i 2020 og en årlig stigning på 23 pct. Udviklingen i Storbritannien minder dermed meget om det tidligere nævnte skøn fra IDC for den globale vækst i Big Data (se afsnit 4.1), som var 40 pct. fra 2010 til 2015 og en vækst på 23 pct. fra 2014-2019 (IDC, 2011). Det tyder således på **en lignende efterspørgsmål i Storbritannien**, hvor 77 pct. af arbejdsgiverne med ansatte inden for Big Data-analyse har oplevet det vanskeligt at udfylde stillingerne inden for Big Data (SAS, 2014).

En væsentlig pointe i forbindelse med efterspørgslen på data scientists, især i USA, er, at udbuddet af data scientists ligeledes er steget. Dette påpeges i en international undersøgelse foretaget af den amerikanske softwarevirksomhed RJMetrics i 2015. Med udgangspunkt i den erhvervsorienterede netværksside LinkedIn, finder RJMetrics 11.400 personer på LinkedIn, der specifikt præsenterer sig selv som data scientist. Dette er en fordobling inden for de sidste fire år. RJMetrics understreger, at estimatet er konservativt, da de yderligere har fundet 60.200 personer med registreret, professionel erfaring inden for data science, 27.700 registreringer af uddannelse inden for området og 254.600 registreringer af evner inden for data science (RJMetrics, 2015). Der er således allerede et solidt udbud af data scientists i en amerikansk kontekst. Her bør det tilføjes, at udviklingen i USA er markant længere inden for Big Data, end den er i en dansk kontekst, hvorfor udbuddet af data scientists i Danmark må formodes at være proportionelt markant lavere.

Det er ikke muligt at lokalisere konkrete tal for efterspørgslen efter data scientists i Danmark. Dog opstiller ITEK[6] under Dansk Industri en prognose, som viser **et udækket behov i Danmark** i 2020 på ca. 3.000 kandidater generelt i ITEK-branchen. ITEK konkluderer, at der skal tillægges en generel vækst inden for teknologibranchen, bl.a. i form af fremdrift i Big Data, hvilket skønnes at udgøre op til yderligere 3.000 kandidater (ITEK, 2015). I den tidligere nævnte undersøgelse af

[6] ITEK har sidenhen skiftet navn til DI Digital. Det er et branchefællesskab i Dansk Industri for virksomheder inden for it, tele, elektronik og kommunikation. DI Digital repræsenterer på nuværende tidspunkt ca. 300 teknologivirksomheder af forskellige størrelser.

IDC tilkendegiver 13 pct. danske virksomheder, at de har planer om at ansætte en specialist inden for Big Data, mens 21 pct. udtrykker et behov for kompetencerne, men mangel på ressourcer (IDC, 2011). Ligeledes peger regeringens redegørelse for den digitale vækstplan på, at på trods af en stigning i antallet af it-uddannede, er det fortsat en udfordring at få dækket virksomhedernes behov inden for it, især i forhold til adgangen til specialiserede it-kompetencer, som fx it-udviklere og dataanalytikere (Regeringen, 2015).

Flere af undersøgelserne i nærværende kortlægning har også undersøgt, **hvilke brancher** data scientists typisk ansættes i. På trods af at data scientists anvendes inden for flere og flere felter, fremgår det tydeligt, at den største andel til stadighed ansættes inden for teknologibranchen. Ifølge en amerikansk ledelsesundersøgelse såvel som en europæisk spørgeskemaundersøgelse, forventes den største vækst i antallet af ansatte data scientists at være inden for kommunikation, medier og teknologi og dernæst den finansielle sektor samt inden for marketingbranchen. Andre interessante sektorer, der forventer at vækste i antallet af ansatte data scientists, er inden for telekommunikation, forbrugerprodukter og detailsektoren samt energisektoren og sundhedssektoren (A.T. Kearney, 2015; Interxion, 2013). It-sektoren er dermed stadigvæk den sektor, der aftager flest kandidater inden for data science, men efterspørgslen på data scientists inden for især den finansielle sektor og marketing kunne fordre et stigende behov for, at data scientist også besidder solide kompetencer inden for forretningsforståelse.

Som del af nærværende kortlægning har Rambøll også valgt at undersøge **dimittendledigheden** blandt de eksisterende uddannelser inden for/med relation til data science. De seneste tal fra Uddannelses- og Forskningsministeriet er fra 2. kvartal 2013 viser en samlet dimittendledighed for alle kandidatuddannelser på 26,4 pct.. Af tallene fremgår det, at der generelt er større efterspørgsel efter kandidater fra det naturvidenskabelige og teknologiske område, som har en ledighed på henholdsvis 25,7 pct. og 24,4 pct. Inden for de teknologiske uddannelser, som er medtaget i denne kortlægning, er der dog en endnu lavere dimittendledighed, nemlig på 17 pct., mens de medtagne matematiske uddannelser har en dimittendledighed på kun 11,2 pct. Det tyder således på, at de kandidater, der besidder de kompetencer, som efterspørges hos data scientists, allerede i 2013 var yderst efterspurgte på arbejdsmarkedet.

"Jeg aftager gerne fem data scientists med det samme"
Der er bred enighed blandt respondenterne om, at der er og bliver stor efterspørgsel på data scientists.

Især respondenter fra de større danske (især finansielle) virksomheder, der netop har – eller er i færd med – at etablere en afdeling med fokus på Big Data og data science, står i akut mangel på medarbejdere, som har den rette faglighed til at arbejde med sådanne analyser.

Dette understreges endvidere af respondenter fra brancheforeningerne, der fortæller, at flere og flere virksomheder må opgive at besætte stillinger, da de ikke kan finde de relevante kompetente medarbejdere. En respondent fra en virksomhed understreger netop dette problem, og fortæller, at de må takke nej til kunder, fordi de ikke har nok medarbejdere, der kan levere analyser af Big Data og som forstår sig på data science.

"I den nuværende situation tager virksomheder, hvad de kan få af medarbejdere, der har grundlæggende færdigheder inden for området. Men kun fordi der ikke er den profil, som de faktisk har brug for."

Flere af respondenterne ansætter både datalogi-, matematik-, og CBS-uddannede til at foretage databehandling og analyser, men de oplever ikke, at disse profiler stemmer nøjagtig overens med det, de har behov for. Citatet i boksen er fra en respondent i den finansielle sektor, og som det indikerer, vurderer de enkelte virksomheder på nuværende tidspunkt selv, hvilke profiler kan udfylde rollen som data scientist. Enkelte virksomheder er også begyndt at

kigge til udlandet efter potentielle kandidater, fordi de ikke oplever at kunne finde dem i Danmark.

Det tyder således på, at der er stor efterspørgsmål efter data scientists i Danmark. Dog er der stor forskel på, hvad respondenterne forstår ved en data scientist og de kompetencer, de efterspørger. Dette uddybes i næste afsnit.

4.4 De efterspurgte kompetencer

Som de forrige afsnit har opridset, tyder det på, at efterspørgslen på data scientists vil stige i fremtiden – både internationalt og nationalt. Men hvilke konkrete kompetencer efterspørges?
I flere af de inkluderede undersøgelser af nuværende **data scientists' uddannelsesbaggrunde** konkluderes, at hovedparten af data scientists er uddannede i datalogi, matematik og statistik eller business management (RJMetrics 2015; Burtch Works, 2014; EMC, 2011). Som påpeget indledningsvist, defineres data science og data scientists dog vidt forskelligt på tværs af de identificerede studier, hvilket også influerer, hvordan studierne opgør de kompetencer og krav hertil, som en data scientist har. I en amerikansk undersøgelse af uddannelsesbaggrunde hos personer, der definerer dem selv som data sicentists på LinkedIn, har lige dele enten baggrund inden for datalogi (13 pct.) eller business management (12,5 pct.) (RJMetrics, 2015). Dog viser en anden amerikansk undersøgelse, foretaget af en amerikansk rekrutteringsvirksomhed med speciale i kvantitative professionelle, at 32 pct. data scientists har baggrund inden for matematik og statistik, 19 pct. inden for datalogi, mens kun 4 pct. har baggrund indenfor business management (Burtch Works, 2014).

Forskelle undersøgelserne imellem kan bl.a. tilskrives, at kriteriet for at indgå som interviewperson i den sidstnævnte undersøgelse er, at man skal kunne arbejde med ustruktureret data. En data scientist er således defineret forskelligt i de to undersøgelser; ud fra en stillingsbetegnelse som data scientist i den ene, og ud fra konkrete teknologiske kompetencer i den anden. I den forbindelse må estimatet for andelen med baggrund inden for business management, i den sidstnævnte undersøgelse, vurderes som det mest realistiske. Begge undersøgelser finder derudover en stor andel data scientists med en ph.d.-grad i de føromtalte fag, og understreger, at arbejdet med data science kræver en høj grad af specialisering samt avancerede kompetencer inden for det teknologiske og matematiske felt (RJMetrics, 2015; Burtch Works, 2014).

Hovedparten af de inkluderede studier om data science fremhæver især **teknologiske kompetencer** som de mest efterspurgte. Flere af undersøgelserne konkluderer, at arbejdet med Big Data-analyser og data science kræver, at den professionelle har solide teknologiske evner, fx i forhold til arbejdet med avancerede algoritmer, dataanalyse, data mining og maskinlæring, som man typisk får gennem studiet i datalogi. Data scientists bør have kendskab til en lang række avancerede værktøjer og have evnen til at arbejde med både SQL, avancerede statistiske pakker, NoSQL databaser såvel som de nyere værktøjer Hadoop og visualiseringsredskaber som fx Tableau. Derudover bør data scientists kunne benytte en række open source-værktøjer som R, Python og Perl (EMC, 2011; RJMetrics, 2015). Samme tendenser fremgår af en online survey blandt 600 professionelle fra forskellige industrier fra hele verden (dog 2/3 fra USA), hvor de mest benyttede redskaber er SQL, R og Python, og indenfor det sidste år også SPARK og Scala Software (O'Reilly, 2015). På samme vis konkluderer en undersøgelse blandt 3.400 jobopslag for data scientists på LinkedIn fra 2016, at de mest efterspurgte værktøjer er SQL og Hadoop, og dernæst følger Python, Java og R (CrowdFlower, 2016). Især manglende kompetencer inden for Hadoop beskriver en af de inkluderede undersøgelser er en stor barriere for rekruttering af medarbejdere til arbejdet med data science (Gartner, 2015).[7]

[7] For en uddybelse af konkrete kompetencer henvises til survey-undersøgelsen af O'Reilley (2012, s.30), der indeholder en liste med 22 konkrete efterspurgte kompetencer hos en data scientist, samt top 20-listerne i henholdsvis RJMetrics (2016) og CrowdFlower (2016).

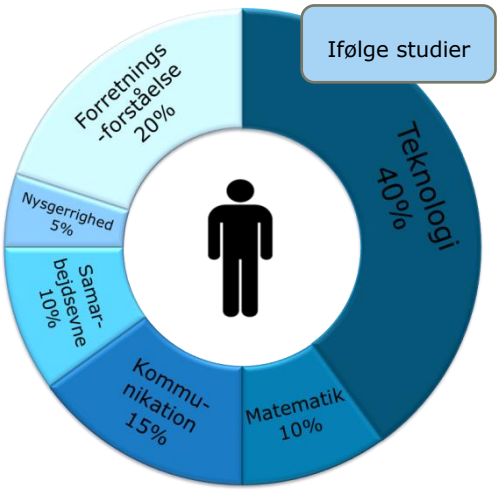
Det fremgår tydeligt af denne kortlægning, at de teknologiske kompetencer er de mest fremtrædende i efterspørgslen efter data scientists. Udover at besidde solide teknologiske kompetencer, fremhæver flere af studierne dog også, at der er behov for et **bredere sæt af andre kompetencer.** Her er især tre kompetencer helt centrale:

1. At kunne kommunikere data til kolleger og udenforstående, således at data giver mening, og potentielt kan skabe værdi.
2. At kunne arbejde med kompleks vidensindsamling og bearbejdning i samarbejde med andre, samt være kreativ og innovativ i dette arbejde. Som data scientist kan/skal man sjældent løse opgaver alene, da kompleksiteten ofte kræver evner til at tænke i nye og tværfaglige løsninger.
3. At have en vis forretningsforståelse. Flere virksomheder efterspørger data scientists, der kan udarbejde business cases, samt forstår, hvordan datalogiske problemer kan opsættes til en forretning, der kan optimere virksomheden.

I arbejdet med analyser af Big Data kræver det typisk evnen til at samarbejde på tværs af professionelle med teknologiske kompetencer, med kompetencer inden for forretningsforståelse såvel som kommunikationskompetencer. En vigtig kompetence hos en data scientist er således evnen til at kunne *kommunikere* outputtet af sit arbejde til andre i virksomheden, for at sikre at det skaber en værdi. Netop evnen til at kommunikere sine resultater og sætte dem ind i den rette forretningsmæssige kontekst er kompetencer, der understreges i adskillige af de inkluderede undersøgelser. Den tidligere nævnte analytiker og forsker inden for forretning og teknologi Thomas Davenport forklarer, at der er et stigende behov for, at en data scientist formår at *formidle* data (Davenport, 2015). Det er altafgørende, at en data scientist formår at indgå i tværfaglige samarbejder og kan kommunikere data på en forståelig måde overfor kolleger, der ikke besidder samme teknologiske kompetencer (EMC, 2011).

Det fremhæves også som essentielt, at en data scientist har evnen til at skelne mellem, hvad der er vigtigt, hvad der er interessant og hvad der ikke er, hvilket kræver en vis *forretningsforståelse* (Burtch Works, 2014). Hertil beskriver den amerikanske data scientist D.J. Patil, at data scientists udover at have en forretningsforståelse, også skal besidde en stor grad af *nysgerrighed*, der gør dem i stand til at se problemer fra forskellige kreative vinkler (Patil, 2011). Som data scientist skal man besidde evnen til at udvikle business cases og have en kreativ samt innovativ tilgang til både analysen samt præsentation af data (A.T. Kearney, 2015; EMC, 2011).

Figur 9: De efterspurgte kompetencer



I en af de inkluderede undersøgelser fremgår det, at næsten 60 pct. af de adspurgte virksomheder oplever, at nyuddannede universitetskandidater inden for data science eller lignende er utilstrækkeligt forberedte på, hvad der forventes af dem i erhvervslivet. International Institution for Analytics forudsiger et stigende behov for en øget forretningsforståelse hos de nyuddannede kandidater i 2016, men beskriver også, hvordan en uddannelse i data science kommer tættere og tættere på en uddannelse i 'Predictive Analytics'[8] i fremtiden, og med denne bevægelse risikeres det, at man går på kompromis med det påkrævede teknologiske fundament af kompetencer (IIA, 2016). Netop denne risiko fremhæves også i interviewene med aftagersegmentet, som uddybes nedenfor.

[8] *Predictive Analytics* handler om at anvende statistiske algoritmer og maskinlæringsteknikker på historisk data til at identificere sandsynligheden for forskellige fremtidige udfald, tendenser og trends (http://www.sas.com/en_us/insights/analytics/predictive-analytics.html).

I figur 9 ovenfor illustreres den fordeling af efterspurgte kompetencer hos en data scientist, som studierne giver anledning til. Procentangivelserne er udtryk for Rambølls vurdering af de inkluderede studier, og er således ikke direkte angivet i studierne. Det, der i særlig grad kendetegner de inkluderede studier, er, at de understreger vigtigheden i, at en data scientist kan *mere* end det teknologiske (og matematiske), og især besidder en vis forretningsforståelse for at kunne indgå i flere af de nyere sektorer, der begynder at arbejde med Big Data (fx den finansielle sektor og marketingbranchen).

"Det handler ikke om at finde svar, men om at stille de rette spørgsmål"
Der tegner sig et lidt andet billede blandt aftagersegmentet end på tværs af undersøgelserne. De efterspurgte kompetencer er fuldstændig ens, men størstedelen af respondenterne fremhæver de teknologiske kompetencer som afgørende for en data scientist.

Størstedelen af respondenterne udtrykker, at det er altafgørende, at en data scientist besidder stærke kompetencer inden for datalogi, statistik og it. Det er vigtigt, at den ansatte data scientist har solide programmeringskompetencer, og ikke *'blot kendskab til programmeringssprog"*. På samme vis er det vigtigt, at vedkommende forstår sig på data; ikke kun på at aflæse data, men på at analysere data statistisk og afdække relevante spørgsmål. Flere af respondenterne fremhæver et behov for, at kommende data scientists har kendskab til en række af de anvendte it-værktøjer som fx Hadoop og Python, samt formår at arbejde med SQL og R-programmering. Hertil understreger enkelte respondenter, at det ikke nødvendigvis handler om hvilket værktøj, men i højere grad handler om at have prøvet værktøjerne af, hvilket vil gøre det nemmere at anvende lignende værktøjer under ansættelse i den pågældende virksomhed.

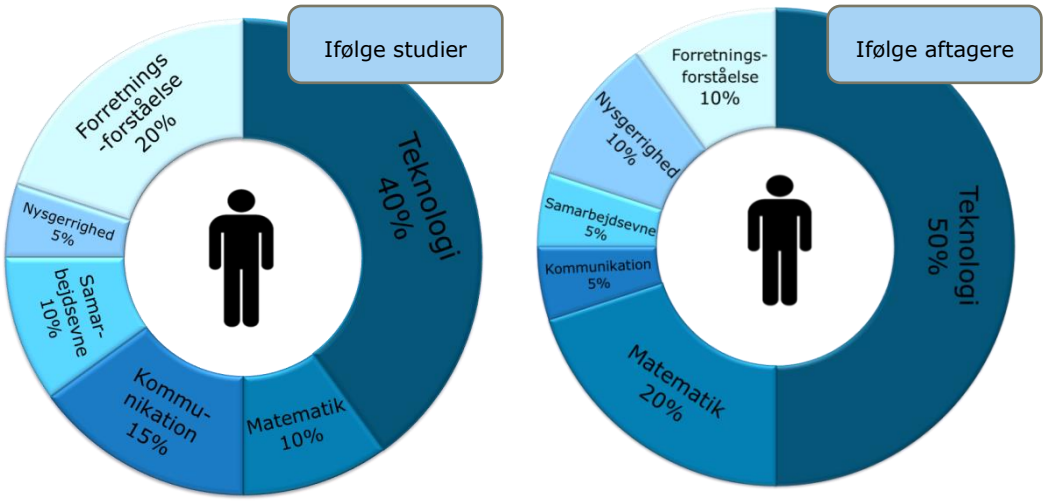
Respondenterne efterspørger således en data scientist med stærke teknologiske kompetencer, der gerne suppleres af kompetencer inden for matematik og forståelse for algoritmer. Dog fremhæver respondenterne også behov for de tre kompetencer, som undersøgelserne fremhæver; det at kunne kommunikere data, være kreativ i arbejdet med data samt have en vis forretningsforståelse. Især kreative kompetencer og såkaldt 'entreprenør-ånd' fremhæves af flere respondenter som afgørende.

På tværs af respondenter med både rig og begrænset erfaring, er der ønske om en medarbejder, der formår at *"stille de spændende spørgsmål".* En af respondenterne fra den finansielle sektor understreger, hvor vigtigt det er at have 'entreprenør-ånd' og 'tænke ud af boksen' i arbejdet med data. Især inden for den finansielle verden har man haft en 0-fejls kultur, men denne kultur fungerer ikke, når man arbejder med data science. Tværtimod skal en data scientist kunne forstå og lære af sine fejl, og videreudvikle analysen efterfølgende.

Også forretningsforståelse fremhæves af enkelte respondenter som vigtigt at lære i samspil med det teknologiske. Respondenterne oplever, at de nuværende medarbejdere, der arbejder med data, til tider mangler den rette forståelse for, hvad der kan skabe værdi for virksomheden, og udfører 'bestillingsarbejde' fremfor 'udviklingsarbejde', som de egentlig er ansat til. Hovedparten af respondenterne mener dog, at forretningsforståelsen kan tillæres, når vedkommende starter i en virksomhed, og derfor ikke behøver en særlig opmærksomhed under uddannelsen. Bekymringen, der illustreres i boksen til højre, blev udtrykt af respondenter fra både den finansielle sektor, brancheorganisationerne og marketingvirksomhederne. Som citatet illustrerer, er flere af respondenterne bekymrede for, om hvorvidt de studerende får de rette dybdegående teknologiske kompetencer, hvis de også skal lære om forretningsprocesser, organisationsteori og andre lignende fag. De vigtigste kompetencer for respondenterne er således især de *teknologiske* såvel som de mere *innovative og kreative* kompetencer.

> *"… Men hvis de skal kunne lidt af hvert, kan de så det teknologiske godt nok? Det kunne godt være min bekymring."*

Figur 10: Sammenligning af de efterspurgte kompetencer



Som figur 10 illustrerer, er der stor lighed mellem de efterspurgte kompetencer, som studierne beskriver, og de kompetencer, som respondenterne efterspørger. Den største forskel på de to 'cirkler' ovenfor er, at respondenterne har et større ønske om teknologiske kompetencer og nysgerrighed, mens studierne i højere grad lægger vægt på vigtigheden ved forretningsforståelse og kommunikation.

En mulig forklaring på denne forskel kan være, at data science er langt mere udviklet og anvendt i USA (hvorfra de fleste studier har ophav) end i Danmark (respondenternes udgangspunkt). På nuværende tidspunkt er kravene til de teknologiske kompetencer forholdsvis høj i Danmark, men som en af de mere erfarne respondenter inden for data science fra den finansielle sektor siger: *"Kravene til databehandling var langt højere før i tiden. Nu har vi jo adskillige værktøjer, der udfører arbejdet for os. Vi skal blot trykke på knapper og så selvfølgelig forstå resultaterne".* Som citatet understreger, er der den senere tid sket en udvikling inden for teknologien (jf. afsnit 4.1), der har medvirket til, at flere og flere data kan behandles og ikke kræver omfattende databearbejdning. Denne udvikling, mener enkelte af respondenterne, kan også medvirke til, at der i højere grad bliver behov for data scientists, der kan *forstå* data og indsætte dem i en forretningsmæssig kontekst, fremfor at have de hårde teknologiske kompetencer, som de fleste efterspørger på nuværende tidspunkt. Citatet i boksen ovenfor er fra en respondent fra en offentlig institution. Man fristes derfor til at påpege, at indholdet – og især vægtningen af fokus på teknologiske kompetencer – i en uddannelse inden for data science bør genbesøges i takt med den teknologiske udvikling for at sikre, at udbuddet matcher efterspørgslen.

> *"Vi vil gerne have de tunge profiler, men jeg kan se, at vi faktisk også får behov for 'light-profiler' i fremtiden."*

"En uddannelse i data science skal kunne..."
Hovedparten af respondenterne finder både timing og belæg for oprettelsen af en bacheloruddannelse inden for data science relevant. Som den store efterspørgsel indikerer, kan det dårligt *"gå hurtigt nok"* at få uddannet flere data scientists.

Uddannelsen bør, ifølge respondenterne, indeholde både teknologiske fag som fx programmering, maskinlæring og data mining, samt matematik og statistik. Derudover bør de studerende blive introduceret for relevante it-værktøjer og enkelte tematikker ift. forretningsforståelse. Et forslag, som flere af respondenterne understreger, er *praktikforløb*, hvor de studerende oplever, hvordan en virksomhed arbejder med data i praksis. Dette vil, udover at give den studerende konkret

praktisk erfaring med data science, også være til gavn for virksomhederne, der kan få (gratis) hjælp til at løse flere af deres opgaver inden for dataanalyse.

En anden pointe, som flere respondenter fremhæver, er, at ITU også bør tænke over, hvordan de vil 'brande' uddannelsen. *"Man vil vel gerne have at uddannelsen appellerer til de teknologistærke drenge, men der er også en pointe i at få 12-talspigerne og en bredere gruppe studerende med"*, fortæller en respondent fra en af interesseorganisationerne. Det kan således være værd at overveje, hvem man ønsker at tiltrække med uddannelsen, og forsøge at nå bredere ud til et større segment end segmentet, der typisk læser datalogi, fx ved at understrege, at data science giver en bredere sammensætning af kompetencer end de eksisterende uddannelser indenfor især teknologi og matematik gør.

Flere af respondenterne påpeger desuden, at en bacheloruddannelse ikke nødvendigvis er den rette vej at gå. En bacheloruddannelse kan have tendens til at blive *"overfladisk",* som enkelte respondenter fortæller, og en data scientist skal netop kunne gå i dybden med dataanalyse. Flere af respondenterne foreslår derfor, at der etableres en kandidatuddannelse også, eller i stedet for. Dog mener endnu andre respondenter, at en bacheloruddannelse kan være det helt rigtige, da de grundlæggende kompetencer inden for især teknik, analyse og forretningsforståelse kan tilegnes her, hvorefter den enkelte studerende kan vælge en retning på sin kandidatuddannelse, fx inden for et af de tre spor eller inden for en specifik sektor.

BILAG 1 - DE INKLUDEREDE STUDIER

De identificerede uddannelser præsenteres i tabelform nedenfor. Først følger de danske uddannelser, der er grupperet under hver faglige tematik, og dernæst følger de udvalgte internationale uddannelser.

De inkluderede danske uddannelser

UDDANNELSER INDENFOR TEKNOLOGI

Uddannelsens titel	**IT**
Niveau og uddannelsesinstitution	Bacheloruddannelse på AU
Link	Link til AU-præsentationsside
Grundlæggende indhold	Uddannelsen har fokus på it-produktudvikling og der undervises i både designmæssige og teknologiske aspekter. De designmæssige aspekter foregår i tæt samarbejde med Arkitektskolen Aarhus. Uddannelsen har en overvægt af teknologiske/ datalogiske fag som programmering, web-teknologi, informationssystemer, datastrukturer og systemudvikling. Uddannelsen er derudover projektorienteret, og gennem uddannelsen har man flere store projektforløb, hvor man arbejder med problemstillinger inden for IT-produktudvikling.
Uddannelsens profil	I forhold til denne kortlægning er uddannelsen blandt de mere teknologiske og indeholder en del datalogiske fag. På uddannelsens andet år kan man specialisere sig inden for tre retninger: 1) Produktdesign, hvor fokus er på de designmæssige aspekter, 2) softwareudvikling, der har fokus på udvikling og arkitektur af softwaresystemer og webservices og 3) forretningsforståelse, der handler om forretningsmodeller for IT-produkter. Af relevante kandidatmuligheder i forhold til data science er uddannelsen adgangsgivende til datalogi og informationsteknologi på AU. Disse to kandidatuddannelser beskrives senere.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015):139 Optag (inkl. standby 2015): 68 Dimittendledighed (2013 k2): -
Uddannelsens titel	**Softwareteknologi**
Niveau	Bacheloruddannelse (en del af en civilingeniøruddannelse) på DTU
Link	Link til DTU-præsentationsside
Grundlæggende indhold	På uddannelsen undervises i, hvordan man bruger matematik og avanceret datalogi til at udvikle it-løsninger. Der undervises i naturvidenskabelige grundfag, men også teknologiske fag som computerarkitektur, programmering og operative systemer.
Uddannelsens profil	Uddannelsen er i høj grad rettet mod den teknologiske del af denne kortlægning. Uddannelsen har en del relevante teknologiske fag f.eks. som algoritme og datastrukturer (5 ECTS). Der er forholdsvis mange valgfag på uddannelsen (45 ECTS), og man kan fx vælge at specialisere sig inden for datasikkerhed, softwareudvikling eller databasesystemer. En relevant kandidatuddannelse kunne være informationsteknologi på DTU, der også beskrives i denne kortlægning.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 264 Optag (inkl. standby 2015): 77 Dimittendledighed (2013 k2): -
Uddannelsens titel	**General Engineering**
Niveau	Bacheloruddannelse (en del af en civilingeniør uddannelse) på DTU
Link	Link til DTU-præsentationsside
Grundlæggende indhold	Uddannelsen er en ny, generel og tværfaglig ingeniøruddannelse. Uddannelsen indeholder en del matematiske fag om avanceret matematik og statistik. Derudover indeholder uddannelsen også fag i programmering.
Uddannelsens profil	Uddannelsen er teknologisk, men dog også forholdsvis generel. På uddannelsens 4. semester er der mulighed for specialisering. En relevant kandidatuddannelse på DTU kunne være informationsteknologi, der også er beskrevet i kortlægningen.

Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - (uddannelse fra 2016) Optag (inkl. standby 2015): - (uddannelse fra 2016) Dimittendledighed (2013 k2): - (uddannelse fra 2016)
Uddannelsens titel	**Netværksteknologi og IT**
Niveau	Bacheloruddannelse (en del af en civilingeniør uddannelse) på DTU
Link	Link til DTU-præsentationsside
Grundlæggende indhold	Uddannelsen omhandler samspillet mellem hardware og software. Der undervises i fag som matematik, fysik og kemi, som kombineres med teknologiske fag som softwareudvikling og programmering.
Uddannelsens profil	Uddannelsen er blandt de mere teknologiske uddannelser i denne kortlægning, hvor hoved-områderne er matematik, kommunikationsteknologi, programmering samt udvikling af software/hardware. Uddannelsen giver forholdsvis mange muligheder for valgfag (45 ECTS), som f.eks. kunne bruges på flere teknologiske fag eller suppleres af fag i økonomi, som DTU selv foreslår.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 143 Optag (inkl. standby 2015): 44 Dimittendledighed (2013 k2): -
Uddannelsens titel	**Software**
Niveau	Bachelor- og kandidatuddannelse (civilingeniør) på AAU
Link	Link til AAU-præsentationsside
Grundlæggende indhold	På uddannelsen undervises i brug af programmeringssprog, netværksteknologi og database-teknologi. Man får overblik over arbejdsprocessen i udvikling af software. Der undervises derudover i algoritmer og datastrukturer.
Uddannelsens profil	Uddannelsen er blandt de mere teknologiske uddannelser i kortlægningen med fokus på data- og softwaresystemer. På kandidatuddannelsen er der flere relevante specialiseringsmuligheder. Relevante at fremhæve er databaseteknologi, hvor der undervises i multidimensionelle databaser, som data warehousing og data mining, samt hvordan man håndterer kompleks data i databaser. Derudover er specialiseringsmuligheden 'Maskinintelligens' også særlig relevant, da den omhandler tilpasning af intelligente computersystemer til forskellige situationer, som f.eks. hvordan et intelligent informationssystem giver information som brugeren har behov for.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 180 Optag (inkl. standby 2015): 106 Dimittendledighed (2013 k2): -
Uddannelsens titel	**Software Engineering**
Niveau	Bachelor- og kandidatuddannelse (civilingeniør uddannelse) på SDU
Link	Link til SDU-præsentationsside(bachelor)
Grundlæggende indhold	Uddannelsen omhandler udvikling af software fra bunden og udvikling af eksisterende soft-ware. Indeholder design, programmering og testning af softwaren. Derudover undervisning i samfunds- og virksomhedsrelaterede problemstillinger.
Uddannelsens profil	Uddannelsen er i relation til kortlægningen forholdsvis teknologisk, med fokus på opbygning af software og videreudvikling. Uddannelsen har dog også fokus på informationsteknologi og samspillet mellem softwaresystemer og organisationer. På kandidatuddannelsen er data science/datavidenskab et af fem obligatoriske fagområder. Her er fokuseret på indsamling, filtrering, behandling, dannelse og distribution af data. De andre områder er software-livscyklus, intelligent software, software økosystemer, innovative softwareløsninger.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 157 (bachelor) Optag (inkl. standby 2015): 67 (bachelor) Dimittendledighed (2013 k2): 19,7 % (for Civilingeniør generelt, uddannet i Odense)
Uddannelsens titel	**Softwareudvikling**
Niveau	Bachelor- og kandidatuddannelse på ITU
Link	Link til ITU-præsentationsside(bachelor)
Grundlæggende indhold	Uddannelsen handler om at designe og programmere software, herunder udnyttelse af programmeringssprog, operative systemer, distribuerede systemer, databasesystemer og modellering. Indeholder klassiske datalogiske kurser i algoritmer, datastrukturer og databa-ser, som suppleres af kurser i organisationsteori, forretningsprocesser og ledelse.
Uddannelsens profil	Uddannelsen er blandt de teknologiske uddannelser i forhold til kortlægningens formål. På trods af dette er der som nævnt fag i forretningsprocesser, kommunikation og organisation i forbindelse med projektarbejde. På kandidaten kan man specialisere sig enten i et design-spor, med særligt fokus på design og strukturering af data og databaser eller i et avanceret computing-spor, med fokus på programmering og softwareudvikling. Især design-sporet er målrettet studerende med ikke it-faglig bachelor. Det kunne f.eks. være en studerende med indsigt i forretning og marke-ting, der ønsker teknologiske kompetencer. På uddannelsen er der mulighed for en specialisering i fx databaser og forretningsanalyse – og i begge specialiseringer er der fag i 'Big Data Management'.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 332 (bachelor) Optag (inkl. standby 2015): 100 (bachelor) Dimittendledighed (2013 k2): -
Uddannelsens titel	**Datalogi**
Niveau	Bachelor-og kandidatuddannelse på AU, AAU, KU, SDU, samt RUC som kombinationsuddan-nelse
Link	Link til UddannelsesGuiden - datalogi (bachelor) Link til datalogi på RUC

	Link til datalogi på KU (kandidat) Link til datalogi på SDU (kandidat) Link til datalogi på AU (kandidat) Link til datalogi på AAU (kandidat)
Grundlæggende indhold	Uddannelsen handler om at analysere og løse komplekse problemer i forbindelse med it. Indeholder programmering, softwarearkitektur, databaser, algoritmer, datastrukturer og systemudvikling.
Uddannelsens profil	Uddannelsen er forholdsvis teknologisk i forhold til denne kortlægning. På trods af enkelte specialiseringsmuligheder minder de fleste af bacheloruddannelser meget om hinanden. Dog med undtagelse af RUC, hvor bacheloruddannelsen tilbydes som en kombinationsuddannelse. Her udbydes datalogi / computer science til 30 ECTS, og denne skal kombineres med et humanistisk-teknologisk eller naturvidenskabeligt fag. På kandidatuddannelserne er der en række relevante specialiseringsmuligheder: På KU kan man specialisere sig inden for Data Science, der handler om at uddanne specialister i data-baseret beslutningsstøtte. Der undervises i informationssøgning, maskinlæring og praktisk anvendelse og analyse af Big Data. En anden specialiseringsmulighed er inden for algoritmer og datastrukturer. Her undervises i algoritmiske paradigmer og teknikker samt de underliggende datastrukturer. På SDU er der mulighed for at tage flere fag, der handler om datasystemer. Mest relevant er fagene; Big Data (10 ECTS) og Principper for databasesystemer (10 ECTS). På AU er kandidatuddannelsen på engelsk og indeholder 30 valgfri ECTS, hvor relevante fag kan være maskinlæring (10 ECTS), avanceret algoritmik, datastrukturer(10 ECTS) og distributive systemer (5 ECTS). På AAU kan man på andet år specialisere sig i databaseteknologi, der omhandler avancerede databaser, som multidimensionelle databaser, spatiotemporale databaser og kompleks data i databaser som f.eks. XML. På RUC er der mulighed for at kombinere datalogi /computervidenskab med f.eks. 'Informatics' eller kommunikation. Der er mulighed for at lære teknikker til analyse, design og konstruktion af softwaresystemer, softwareudvikling, computer- og systemarkitektur, algoritmer og datarepræsentationer.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 863 (bachelor, egen sammentælling, minus RUC) Optag (inkl. standby 2015): 346 (bachelor, egen sammentælling, minus RUC) Dimittendledighed (2013 k2): 14,5 % (egen sammentælling)
Uddannelsens titel	**Data Engineering**
Niveau	Kandidatuddannelse på AAU
Link	Link til AAU-præsentationsside
Grundlæggende indhold	Uddannelsen udbydes på engelsk og omhandler håndtering af store mængder data, bl.a. tids-, rum- og multidimensionalt data. Uddannelsen giver viden om dataintensive systemer, Data Warehousing, Business Intelligence, Big Data og mobilt internet.
Uddannelsens profil	Uddannelsen er blandt de mere teknologiske uddannelser i kortlægningen, særligt i forhold til datahåndtering. Derudover er der også fokus på entreprenørskab på uddannelsen. Direkte adgangsgivende bacheloruddannelser på DTU og Software og Datalogi, der også beskrives i denne kortlægning.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): - Dimittendledighed (2013 k2): -
Uddannelsens titel	**Matematisk Modellering og Computing**
Niveau	Kandidatuddannelse (civilingeniør) på DTU
Link	Link til DTU-præsentationsside Link til Big Data- specialisering
Grundlæggende indhold	Uddannelsen foregår på engelsk og giver viden om anvendt matematik og matematisk modellering. Der undervises desuden i moderne computerudstyr og håndtering af store datamængder.
Uddannelsens profil	Uddannelsen er i forhold til denne kortlægning blandt de mere teknologiske uddannelser. Uddannelsen giver mulighed for at vælge fokusområder inden for flere forskellige teknologiske områder. Blandt de mest relevante fokusområder er maskinel læring og signalbehandling, industriel og anvendt statistik. Her er bl.a. fag i dataanalyse, mens fokusområdet 'finansiel ingeniørvidenskab' kunne være relevant, hvis man ønsker en mere forretningsorienteret profil. Uddannelsen er også adgangsgivende til DTU's særlige profil inden for data science, med specialisering i Big Data (45 ECTS), der er en særlig profil med fokus på både indsamling, behandling, analyse og forretningsprocesser i forhold til Big Data. Det er derudover et krav, at der skrives speciale inden for data science.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): - Dimittendledighed (2013 k2): -
Uddannelsens titel	**Digitale medieteknologier**
Niveau	Kandidatuddannelse (civilingeniør) på DTU
Link	Link til DTU-præsentationsside Link til big data specialisering
Grundlæggende indhold	Uddannelsen er engelsksproget og handler om at udvikle services og løsninger inden for informations- og kommunikationsteknologi.

Uddannelsens profil	Uddannelsen er i forhold til denne kortlægning blandt de mere teknologiske, dog også med undervisning, der giver en vis forretningsforståelse. Man kan specialisere sig inden for et af fem områder. Her er området data science det mest relevante. Her er der blandt flere relevante fag i computerværktøjer til Big Data, it-mæssig dataanalyse og introduktion til maskinlæring og data mining (alle 5 ECTS) Uddannelsen er også adgangsgivende til DTU's særlige profil inden for data science, med specialisering i Big Data (45 ECTS), der er en særlig profil med fokus på både indsamling, behandling, analyse og forretningsprocesser i forhold til Big Data. Det er derudover et krav at der skrives speciale inden for data science.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): - Dimittendledighed (2013 k2): -
Uddannelsens titel	**Machine Intelligence**
Niveau	Kandidatuddannelse på AAU
Link	Link til AAU-præsentationsside
Grundlæggende indhold	Uddannelsen er engelsksproget og omhandler metoder og teknologier til intelligente webbaseret systemer. Man bliver undervist i Data mining, kunstig intelligens, Agentteknologi, Web Intelligence, og får en bred indføring i maskinlæring.
Uddannelsens profil	Uddannelsen er blandt de mere teknologiske i denne kortlægning, særligt i forhold til maskinlæring. En bachelor i software eller datalogi, som også er beskrevet i denne kortlægning, er adgangsgivende til uddannelsen.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): - Dimittendledighed (2013 k2): -
Uddannelsens titel	**IT**
Niveau	Diplomingeniør på DTU, SDU, AU og VIA University College (under forskellige navne)
Link	Link til UddannelsesGuiden - Diplomingeniør IT
Grundlæggende indhold	Uddannelsen omhandler udvikling af informationssystemer. Man lærer om programmering af hardware og software og om datanetværk.
Uddannelsens profil	Uddannelsen er forholdsvis teknologisk omkring programmering, software og systemudvikling. På de forskellige uddannelsessteder er dog rig mulighed for specialisering. På Via University College kan man f.eks. specialisere sig i retning af Business Information Systems, SDU giver blandt andet mulighed for at specialisere sig mod det mere organisationsorienterede, AU giver bl.a. mulighed for at specialisere sig mod kommunikationsteknologi eller indlejrede softwaresystemer, mens uddannelsen på DTU hovedsageligt omhandler datalogi og softwareudvikling. Alle uddannelser giver mulighed for at læse videre til civilingeniør. Det kan f.eks. være computerteknologi på AU, hvor en specialisering i distributive systemer eller softwareteknologi kan være relevant.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 715 (egen sammentælling) Optag (inkl. standby 2015): 260 (egen sammentælling) Dimittendledighed (2013 k2): 20,1 (egen sammentælling)
Uddannelsens titel	**Datamatiker og softwareudvikling**
Niveau	Professionsuddannelse (to dele) (flere undervisningssteder landet over)
Link	Link til UddannelsesGuiden - Datamatikker Link til UddannelsesGuiden - Softwareudvikling
Grundlæggende indhold	'Datamatiker' omhandler databaser og deres opbygning, programmering samt udvikling og vedligeholdelse af it-systemer. Uddannelsen tager 2 ½ år. Softwareudvikling er en overbygning til datamatiker på 1 ½ år. Uddannelsen omhandler brugen af forskellige databasetyper, samt udvikling af og videreudvikling af systemer.
Uddannelsens profil	I forhold til kortlægningen har uddannelsen fokus på de teknologiske aspekter af data science, særligt i forhold til programmering og systemudvikling. På særligt softwareudvikling undervises der også i hvordan man konstruerer store datatunge distribuerede it-systemer.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 3288 (Datamatiker - egen sammentælling) Optag (inkl. standby 2015): 845 (Datamatiker - egen sammentælling) Dimittendledighed (2013 k2): 14,6 (datamatiker, egen sammentælling) 19 % (softwareudvikling, egen sammentælling)
Uddannelsens titel	**Data Science & Big Data**
Niveau	Efter- og videreuddannelse på AAU
Link	Link til AAU-præsentationsside
Grundlæggende indhold	Faget er en del af linjen Softwarekonstruktion under Master i IT. Faget giver viden og praktisk erfaring med data science og Big Data gennem begreber, statistisk dataanalyse og Big Data-Skalering
Henvendt til:	It-udviklere Forretningsanalytikere

UDDANNELSER INDENFOR MATEMATIK

Uddannelsens titel	**Matematik og teknologi**
Niveau	Bacheloruddannelse (DTU), bachelor- og kandidatuddannelse på AAU
Link	Link til DTU-præsentationsside Link til AAU-præsentationsside (bachelor)
Grundlæggende indhold	Uddannelsen omhandler algoritmer til store beregninger, og computerbaserede løsninger.

	Man får et kendskab til forskellige teorier og metoder til løsning af differentialligninger, samt statistiske metoder til at analysere informationer.
Uddannelsens profil	Uddannelsen er forholdsvis matematisk men kombineret med en teknologisk forståelse for, hvordan man bruger computerbaserede løsninger til matematiske og statistiske problemstillinger. Særligt uddannelsen på DTU giver mulighed for egen specialisering. Den mest relevante specialisering er "computerintensiv dataanalyse", der primært omhandler statistisk analyse og matematisk modellering. Den naturlige kandidatoverbygning på DTU, i forhold til kortlægningen, er enten matematisk modellering og computing eller digitale medieteknologier, der begge er beskrevet i denne kortlægning. Kandidatuddannelsen på AAU er en videreudvikling af bacheloruddannelse.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 286 (bachelor, egen sammentælling) Optag (inkl. standby 2015): 100 (egen sammentælling) Dimittendledighed (2013 k2): -
Uddannelsens titel	**Matematik**
Niveau	Bachelor-og kandidatuddannelse på AU, AAU, KU, SDU, samt RUC som kombinationsuddannelse
Link	Link til UddannelsesGuiden - matematik (bachelor) Link til anvendt matematik - SDU (bachelor)
Grundlæggende indhold	På uddannelsen undervises i de matematiske kerneområder såsom algebra, geometri og differentialligninger.
Uddannelsens profil	Uddannelsen er i sagens natur matematisk. Der er på bacheloruddannelsen forskellige specialiseringsmuligheder, hvor det mest relevante på alle uddannelser er at kombinere med et datalogisk spor. På kandidatuddannelser er der ligeledes mulighed for specialisering indenfor især det matematiske spor. På SDU kan det også være relevant at læse anvendt matematik på både bachelor- og kandidatniveau. Det er, til forskel fra matematik, en mere tværfaglig uddannelse med fokus på at løse konkrete problemstillinger gennem matematisk modellering. Uddannelsen benytter ligeledes, til forskel fra de øvrige indenfor matematik, oftere computerteknologier.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 760 (bachelor, egen sammentælling, minus RUC) Optag (inkl. standby 2015): 302 (bachelor, egen sammentælling, minus RUC) Dimittendledighed (2013 k2): 8,9 % (egen sammentælling)
Uddannelsens titel	**Statistik**
Niveau	Kandidatuddannelse på KU og AU
Link	Link til KU-præsentationsside Link til AU-præsentationsside
Grundlæggende indhold	Uddannelsen er engelsksproget og omhandler komplicerede statistiske problemstillinger. Man lærer om sandsynlighedsregning, statistiske teorier og praktisk datahåndtering. Der er mulighed for at specialisere sig i anvendt statistik.
Uddannelsens profil	Uddannelsen er blandt de mere matematiske uddannelser i denne kortlægning. På både KU og AU er der visse specialiseringsmuligheder, dog inden for rammen af statistik. En bacheloruddannelse i matematik med specialisering i statistik på KU eller matematisk modellering på AU er adgangsgivende.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): 37 Dimittendledighed (2013 k2): 8,9 % (egen sammentælling)

UDDANNELSER INDENFOR FORRETNINGSFORSTÅELSE

Uddannelsens titel	**Strategisk analyse og systemdesign**
Niveau	Bacheloruddannelse på DTU
Link	Link til DTU-præsentationsside
Grundlæggende indhold	Uddannelsen kombinerer en teknologisk forståelse med samfundsmæssig og økonomisk indsigt, hvor der læres om matematisk modellering, beslutningsstøtte, økonomi samt organisation og ledelse.
Uddannelsens profil	Uddannelsen kombinerer teknologiske kompetencer med økonomisk indsigt. Der er en række meget relevante teknologiske linjefag på uddannelsen som databasesystemer, forretningsanalyse og introduktion til maskinlæring og datamodellering. Uddannelsen giver mulighed for en del valgfag (45 ECTS) blandt DTU's øvrige fag eller på et andet universitet. En oplagt adgangsgivende kandidatuddannelse på DTU vil være Matematisk Modellering og Computing, der også er beskrevet i denne kortlægning.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 118 Optag (inkl. standby 2015): 33 Dimittendledighed (2013 k2): -
Uddannelsens titel	**Erhvervsøkonomi**
Niveau	Bacheloruddannelse på AU
Link	Link til AU-præsentationsside (HA) Link til Economics and Business Administration Link til Erhvervsøkonomi med tilvalg
Grundlæggende indhold	Uddannelsen er centreret om økonomi og ledelse, hvor man arbejder med virksomheders økonomiske problemstillinger og udvikler erhvervsøkonomiske strategier.

Uddannelsens profil	Uddannelsen er ikke blandt de mest teknologiske i denne kortlægning. Til gengæld giver uddannelsen indsigt i de erhvervsmæssige sammenhænge som data science kan fungere i. Der er blandt andet et fag i Business Data Collection (10 ECTS). Grunden til at uddannelsen er medtaget i denne kortlægning er, at hele 5. semester er tillagt en specialisering. Blandt flere specialiseringer er en specialisering inden for informationssystemer relevant, hvor man lærer, hvordan it-teknologi bedst indføres i virksomheder, eller en specialisering inden for Business Analytics, hvor man lærer at håndtere og analysere data, så beslutningstagere kan opnå en bedre forståelse. Der er også en engelsksproget udgave af uddannelsen, der hedder Economics and Business Administration, hvor der er mere fokus på internationale forhold, men ligeledes mulighed for specialisering i Business Analytics. En tredje relevant erhvervsøkonomisk mulighed på AU er at vælge erhvervsøkonomi med tilvalg, hvor man kan læse fag i f.eks. datalogi ved siden af den erhvervsøkonomiske uddannelse. En relevant adgangsgivende kandidatuddannelse på AU kan være IT, kommunikation og organisation eller Erhvervsøkonomi – Business Intelligence, der også er beskrevet i denne kortlægning.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 1419 (Generelt for HA i Aarhus) Optag (inkl. standby 2015): 617 (Generelt for HA i Aarhus) Dimittendledighed (2013 k2): -
Uddannelsens titel	**Erhvervsøkonomi og it**
Niveau	Bachelor- og kandidatuddannelse på CBS
Link	Link til CBS-præsentationsside (bachelor)
Grundlæggende indhold	Uddannelsen handler om at løse erhvervsøkonomiske problemstillinger ved hjælp af it og informationssystemer. Uddannelsen kombinerer finansiering, regnskab og international økonomi med udvikling af it-systemer og programmering af små systemer og databaser.
Uddannelsens profil	I forhold til kortlægningen er uddannelsen blandt de mindre teknologiske, på trods af fag i blandt andet programmering. Til gengæld har uddannelsen et stort fokus på at de studerende opnår kompetencer i forhold til forretningsforståelse. Bacheloruddannelsen giver mulighed for 2 valgfag til i alt 15 ECTS point, hvor man, inden for erhvervsøkonomi og IT, kan vælge mellem fag på alle danske universiteter. De naturlige kandidatuddannelser er E-business og Business Administration og Information System, hvor særligt sidstnævnte synes relevant for kortlægningen. Denne kandidatuddannelse er engelsksproget og handler om information og informationsteknologi som essentielle for aktiviteter og beslutninger i virksomheder. Uddannelsen integrerer erhvervsøkonomi og informationsteknologi. Fra andet semester kan man specialisere sig mod enten Information Management, der har fokus på kommunikation, strategi og forretningsmetoder i forhold til information eller IT Management og Business Economics, der fremstår mere fokuseret på de teknologiske aspekter og på hvordan informationsteknologier kan forbedre eller skabe nye forretningsmodeller. Uanset specialisering er der på første semester et fag i Big Data Management på 7, 5 ECTS point.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 432 (bachelor) Optag (inkl. standby 2015): 110 (bachelor) Dimittendledighed (2013 k2): **-**
Uddannelsens titel	**Erhvervsøkonomi – Business Intelligence**
Niveau	Kandidatuddannelse på AU
Link	Link til AU-præsentationsside
Grundlæggende indhold	Uddannelsen er engelsksproget og understøtter kompetencer ift. at forbedre beslutningstagen i virksomheder gennem dataanalyse.
Uddannelsens profil	Uddannelsen er ikke blandt de mest teknologiske i denne kortlægning, til gengæld giver den indsigt i, hvordan man opsætter data til at være relevant beslutningsstøtte i virksomheder. På uddannelsen er der fag i f.eks. Data Warehousing (5 ECTS) og Data Mining for Business Decisions (10 ECTS). Derudover består uddannelsens tredje semester af valgfag (30 ECTS) på AU eller andre universiteter, som man f.eks. kunne benytte til mere teknologiske fag.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): - Dimittendledighed (2013 k2): -
Uddannelsens titel	**IT- ledelse**
Niveau	Kandidatuddannelse på AAU
Link	Link til AAU-præsentationsside
Grundlæggende indhold	Uddannelsen giver indsigt i it og informationssystemer. Der undervises i at udvikle og implementere it for virksomheder og organisationer. Derudover er der på uddannelsen fokus på hvordan ledelse, forandring og udvikling af organisationer kan gennemføres via IT-løsninger.
Uddannelsens profil	Uddannelsen er ikke blandt de mest teknologiske uddannelser i denne kortlægning. I stedet er der fokus på styring og ledelse af komplekse it-projekter. Uddannelsen indeholder dog et fag i programmering og databasesystemer (5 ECTS). En bred vifte af bacheloruddannelser er adgangsgivende til uddannelsen som f.eks. kan være informationsteknologi, politik og administration eller informatik. Sidstnævnte kombinerer kommunikation og datalogiske grundprincipper, som kommunikationsprocesser, avanceret systemudvikling og design af brugergrænseflader.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): - Dimittendledighed (2013 k2): 38,7 %
Uddannelsens titel	**It, Design an Application**
Niveau	Kandidatuddannelse på AAU
Link	Link til AAU-præsentationsside

Grundlæggende indhold	Uddannelsen er designet til de bachelorer, der ikke har en teknologisk baggrund. Uddannelsen lægger op til at studerende med baggrund indenfor f.eks. samfundsvidenskab, økonomi eller kommunikation kan bruge denne viden til at udvikle it-systemer. Uddannelsen har fokus på design af IT-systemer, menneske-maskine-interaktion, systemudvikling og programmering
Uddannelsens profil	Uddannelsen er teknologisk i et vist omfang, tilpasset den målgruppe, der ikke har teknologisk baggrund i forvejen. Det betyder dog samtidig, at de teknologiske fag ikke er ligeså videregående som på andre teknologiske kandidatuddannelser i denne kortlægning. Uddannelsen indeholder blandt andet et fag i databaseudvikling. En lang række bacheloruddannelser indenfor humaniora, samfundsvidenskab og design er adgangsgivende til uddannelse. En mere forretningsorienteret bacheloruddannelse kunne f.eks. være en relevant kombination i forhold til kortlægningens formål.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): - Dimittendledighed (2013 k2): -
Uddannelsens titel	**Information management**
Niveau	Bacheloruddannelse (CBS)
Link	Link til CBS-præsentationsside
Grundlæggende indhold	Uddannelsen er engelsksproget og omhandler informations- og videnshåndtering. Deraf bl.a. at lære teknologier til strukturering af information.
Uddannelsens profil	Uddannelsen er ikke blandt mest teknologiske uddannelser, der er med i denne kortlægning. Til gengæld giver uddannelsen kompetencer omkring kommunikation og forretningsadministration, der kombineres med indsigt omkring informationssystemer. En oplagt mulighed er at læse kandidat i erhvervsøkonomi og IT, som uddannelsen er adgangsgivende til.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 371 Optag (inkl. standby 2015): 60 Dimittendledighed (2013 k2): -

UDDANNELSER INDENFOR KOMMUNIKATION

Uddannelsens titel	**Informationsvidenskab**
Niveau	Bachelor- og kandidatuddannelse på AU og AAU - en specialisering af Kommunikation og digitale medier
Link	Link til AU-præsentationsside (bachelor)
Grundlæggende indhold	Uddannelsen omhandler programmering, kommunikationsteori og organisationsanalyse. Uddannelsen kombinerer humanistiske, samfundsvidenskabelige og naturvidenskabelige fag. Faglige emner er blandt andet programmering, systemudvikling og databaser.
Uddannelsens profil	Uddannelsen er blandt de mindre teknologiske uddannelser. Til gengæld er uddannelsen forholdsvis bred, og giver indsigt i forhold til forretning, organisationsprocesser og kommunikation. På AU er den direkte kandidat ligeledes Informationsvidenskab, hvor der er forskellige specialiseringsmuligheder. Her er det mest oplagte valg at tage fag i it-projektledelse og vælge analyse- og konsulentprofilen på tredje semester, for at opnå kompetencer indenfor data science. På AAU er kandidatuddannelse fokuseret på systemudvikling, IKT, og menneske-maskine-organisation.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 273 (AU) og 471(AAU) Optag (inkl. standby 2015): 82 (AU) og 69 (AAU) Dimittendledighed (2013 k2): -
Uddannelsens titel	**Informationsteknologi**
Niveau	Bacheloruddannelse (AAU), kandidatuddannelse (DTU)
Link	Link til AAU-præsentationsside (bachelor) Link til DTU-præsentationsside (kandidat)
Grundlæggende indhold	Bacheloruddannelsen (AAU) omhandler design af IT-systemer, systemudvikling og programmering, IT-systemer til kommunikation og forretningsforståelse. Kandidatuddannelsen (DTU) er engelsksproget og omhandler design og brug af beregningskomponenter til software- og hardwaresystemer. Algoritmer, systemintegration, programmelkonstruktion er blandt uddannelsens nøgleområder.
Uddannelsens profil	Bacheloruddannelsen på AAU kombinerer kommunikation og teknologisk it-indsigt, mens kandidatuddannelsen på DTU primært er teknologisk. På AAU skal der på 4. semester udarbejdes et større projekt, der kombinerer indsigt fra fagene i Databaseudvikling og Software Engineering. Her opnås viden om udvikling af databasesystemer til en specifik brugssammenhæng. På DTU kan man specialisere sig i indlejrede og distributive systemer. Derudover kan man tage en særlig profil inden for data science, med specialisering i Big Data (45 ECTS), der er en særlig profil med fokus på både indsamling, behandling, analyse og forretningsprocesser i forhold til Big Data. Det er derudover et krav, at der skrives speciale inden for data science. En oplagt kandidatuddannelse på AAU kunne være IT-ledelse eller informationsvidenskab, mens en relevant bacheloruddannelse på DTU er softwareteknologi. De er alle beskrevet i kortlægningen.

Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): 115 (AAU) Optag (inkl. standby 2015): 50 (AAU) Dimittendledighed (2013 k2): 13,1 % (DTU)
Uddannelsens titel	**It, kommunikation og organisation**
Niveau	Kandidatuddannelse på AU
Link	Link til AU-præsentationsside
Grundlæggende indhold	Uddannelsen handler om, hvordan kommunikationsteori og programmering kan benyttes til at udvikle it i en virksomhed eller organisation. Man lærer at løse krævende teknologiske, udviklings- og ledelsesmæssige opgaver inden for kommunikations- og informationsteknologi. Uddannelsen har desuden fokus på projektarbejde.
Uddannelsens profil	Uddannelsen er ikke blandt de mest teknologiske i denne kortlægning, selvom der også læres programmering. Til gengæld giver uddannelsen indsigt i kommunikation og organisering. Man kunne med fordel kombinere uddannelsen med en mere teknologisk bachelor. Uddannelsen har tre spor som alle gennemgår: organisation (med fokus på kultur, udvikling, forretningsstrategi, ledelse mm.), kommunikation (med fokus på teknologi, interne/eksterne processer og strategier) og informationsteknologi (med fokus på udvikling af it-baserede informations- og kommunikationssystemer). Hvis den studerende har en bachelor med fokus på IT, indeholder kandidaten bl.a. fag med fokus på produktudvikling og virksomhedsarkitektur- og design, mens bachelorer, der f.eks. har en mere økonomisk baggrund, i stedet vil have flere fag med fokus på programmering, systemudvikling og databaser. I forhold til denne kortlægning vil en adgangsgivende baggrund som bachelor i IT eller Informationsvidenskab, samt en professionsuddannelse som datamatiker og softwareudvikling være nærliggende. Alle disse uddannelser er beskrevet i denne kortlægning. Man kunne tilsvarende også forestille sig en mere økonomisk bacheloruddannelse inden for økonomistyring, organisation eller marketing, da kandidatuddannelsen giver mulighed for introduktion til de mere teknologiske aspekter.
Oplysninger relateret til efterspørgsel	Antal ansøgere(2015): - Optag (inkl. standby 2015): 90 (36 % fremgang 2010-2015) Dimittendledighed (2013 k2): -

De internationale uddannelser

Universitet	**Massachusetts Institute of Technology (MIT)**
Uddannelsesnavn og niveau	Hovedfag i Business Analytics Hovedfag i Computer Science and Engineering Hovedfag i Mathematics and Computer Science Sidefag i Statistics and Data Science
Link	Link til hovedfag i Business Analytics Link til hovedfag i Computer Science og Engineering Link til hovedfag i Mathematics og Computer Science Link til sidefag i Statistics og Data Science
QA-Ranking:	1
Indhold	*Hovedfag i Business Analytics* Business Analytics omhandler Operations Research og deraf statistik og dataanalyse. På uddannelsen lærer de studerende, hvordan man kan bruge data til at træffe bedre beslutninger i en virksomhed. MIT beskriver bl.a., at en uddannet i Business Analytics kan arbejde med Big Data. Uddannelsen bevirker således, at de studerende får bedre matematiske evner især inden for dataanalyse. Uddannelsen indeholder bl.a. fag i programmering, statistik, maskinlæring og data mining. Uddannelsen er både teknologisk og forretningsorienteret, hvor den studerende lærer at benytte data gennem matematik, statistik og programmering til at støtte beslutningstagere i virksomheder. *Hovedfag i Computer Science and Engineering* Uddannelsen kombinerer klassiske ingeniørfag med computervidenskabelige fag. Af relevante fag på uddannelsen er blandt andet 'Introduction to Computational Thinking and Data Science', introduktion til maskinlæring, samt flere matematiske/statistiske fag. Uddannelsen er blandt de mere teknologiske i kortlægningen, hvor særligt klassiske ingeniørfag spiller en større rolle. *Hovedfag i Mathematics and Computer Science* Da udfordringer indenfor computervidenskab ofte løses via den matematiske metode, er de to fag nært forbundet. Uddannelsen indeholder en række fag i algebra og algoritmer samt softwareudvikling. Uddannelsen er blandt de mere teknologiske i kortlægningen. *Sidefag i Statistics and Data Science* Uddannelsen er helt ny, og er tilgængelig for studerende fra alle tre førnævnte hovedfagsuddannelser. Studerende skal vælge kurser inden for minimum fire af seks emner indenfor grundlæggende matematiske og datalogiske fag, inden for statistiske fag (både grundlæggende og i forhold til data analyse) og inden for dataanalyse, hvor der vælges to fag som f.eks. maskinlæring eller forretningsanalyse. Derudover udarbejdes en opgave inden for

	faget 'Statistics, Computation and Applications'.
Universitet	**Stanford University**
Uddannelsesnavn og niveau	Hovedfag i Computer Science Hovedfag i Management Science and Computer Science Kandidat i Statistics: Data Science
Link	Link til hovedfag i Computer Science Link til hovedfag i Management Science and Computer Science Link til kandidat i Statistics: Data Science
QA-Ranking:	2
Indhold	*Hovedfag i Computer science* Uddannelsen starter med en forholdsvis klassisk introduktion til datalogi, hvorefter man kan specialisere sig. Af relevante specialiseringer er der f.eks. informationslinjen, der indeholder en del matematiske fag, programmering og et fag, der introducerer til databaser. Uddannelsen er forholdsvis datalogisk, hvor data science spiller en mindre rolle. *Hovedfag i Management Science and Computer science* Uddannelsen indeholder en del matematiske fag i f.eks. stokastisk modellering, anvendt statistik og probabilitv analyse. Derudover er der fag i programmering, som kombineres med økonomiske fag som 'Economic Analyses', finansiering og organisationsanalyse. Uddannelsen kombinerer dermed en teknologisk indsigt med mere økonomiske erkendelser. Kandidat *i Statistics: Data Science* Studerende på uddannelsen får stærke evner indenfor matematik, statistik og programmering for dermed at blive stærk i data science. Uddannelsen indeholder fag i maskinlæring, stor-skala softwareudvikling, avanceret multicore systemer, 'Mining Massive Data Sets' og i 'Business Intelligence for Big Data'. Uddannelsen er blandt de teknologiske og matematiske i denne kortlægning og giver stor indsigt i statistik og softwaresystemer.
Universitet	**Oxford University**
Uddannelsesnavn og niveau	Bachelor i Mathematics and Computer Science
Link	Link til bachelor i Mathematics and Computer Science
QA-Ranking:	3
Indhold	Uddannelsen kombinerer matematisk argumentation med en forståelse for computere, da både matematik og datalogi i mange sammenhænge er dybt integrerede. Uddannelsen er opbygget på den måde, at det første år består af fag i matematik, bl.a. med fokus på komplekse numre, lineær algebra og sandsynlighedsregning, og fag i datalogi, med fokus på bl.a. programmering, design og analyse af algoritmer. På andet år af uddannelsen bygges videre på disse fag, men med mulighed for en valgfrihed i de matematiske fag. På det tredje år har de studerende endnu større valgfrihed ift. fag, og kan bl.a. vælge fag i databaseforståelse indenfor datalogi. Der er også mulighed for at tage uddannelsen i fire år og deraf få en "Master of Mathematics and Computer Science". I dette tilfælde har de studerende mulighed fag i maskinlæring, databasesystemers implementering og/eller algoritmer og data strukturer. Uddannelsen er forholdsvis teknologisk, men dog uden specifikke fag vedrørende data science. Man kan desuden også på Oxford tage en bachelor i datalogi, som giver mulighed for samme datalogiske fag som ovenfor beskrevet.
Universitet	**Harvard University**
Uddannelsesnavn og niveau	Bachelor i Computer Science
Link	Link til bachelor i Computer Science
QA-Ranking:	4
Indhold	Uddannelsen kan tages på flere forskellige måder ud fra de forudsætninger, den studerende i forvejen har inden for datalogi. Derudover giver uddannelsen mulighed for forskellige specialiseringer ud fra interesser. Er den studerende fx interesseret i finansiering, kan vedkommende tage fag i data science, og ved interesse for matematik, kan vedkommende følge fag om datastrukturer og algoritmer.
Universitet	**Carnegie Mellon University**
Uddannelsesnavn og niveau	Hovedfag i Information System Hovedfag i Statistic and Machine Learning
Link	Link til hovedfag i Information System Link til hovedfag i Statistic and Machine learning Link til hovedfag i Computer Science Link til kandidat i Business Intelligence og Data Analysis, Link til kandidat i Machine Learning, Link til kandidat i Computational Data Science
QA-Ranking:	5
Indhold	*Hovedfag i Information System* Uddannelsen er organiseret under de humanistiske og socialvidenskabelige uddannelser, og derfor ikke under de teknologiske eller matematiske, som de fleste andre uddannelser i denne kortlægning. På uddannelsen undervises i, hvordan man designer og implementerer løsninger, der møder organisations-, samfunds- og ledelsesmæssige behøv for information og beslutningsstøtte. På uddannelsen undervises i, hvordan man skaber, distribuerer og benytter information via teknologi. Det første år af uddannelsen består af en del fag i matematik og statistik, såsom regning og statistisk ræsonnement og praksis, samt fag i datalogi, som fx programmering og introduktion til datastrukturer. Herefter bliver uddannelsen mere fokuseret med mulighed for en del valgfag, hvor man blandt f.eks. kan vælge en forret-

	ningsorienteret vinkel. Blandt flere relevante fag er der fag i databasedesign og -udvikling, databasedesign og -implementering, Big Data og Analyse, 'Modern Data Management'. Uddannelsen giver dermed en grundlæggende introduktion til matematik og statistik, samt datalogi på akademisk niveau, hvorefter uddannelsen bliver mere fokuseret mod informationssystemer, med mulighed for f.eks. at tage en forretningsorienteret vinkel på data science. *Hovedfag i Statistics and Machine Learning* Carnegie beskriver selv uddannelsen som ideel for studerende med interesse for statistisk beregning, data science og Big Data-problematikker. Uddannelsen indeholder klassiske statistiske fag, som lineær / matrix algebra, sandsynlighedsregning og statistisk inferens. Derefter bliver uddannelsen mere målrettet data science og maskinlæring med fag i store datasæt, avancerede metoder til dataanalyse, begyndende og videregående dataanalyse samt algoritmer og avancerede datastrukturer. Derudover indeholder uddannelsen også fag i programmering samt giver mulighed for valgfag. Der er naturligvis også mulighed for at læse hovedfag i datalogi på Carnegie. Her er der også mulighed for forskellig specialisering, dog ikke med samme relevans for data science som de to ovennævnte uddannelser. Carnegie har også forskellige relevante kandidatuddannelser f.eks. en master i Business Intelligence & Data Analytics, Master i Computational Data Science eller en master i maskinlæring. Se links til uddannelserne.
Universitet	**University of Cambridge**
Uddannelsesnavn og niveau	Bachelor og kandidat i Engineering Bachelor i Computer Science
Link	Link til bachelor og kandidat i Engineering Link til bachelor i Computer science
QA-Ranking:	6
Indhold	*Bachelor og kandidat i Engineering* Uddannelsen begynder med en grundlæggende, generel introduktion til ingeniørfaget og dertil fag om design, fremstillingsvirksomhed og ledelse, samt matematiske metoder. Herefter skifter studiet til at være mere emneorienteret, hvor der blandt andet er mulighed for en forretningsorientering. På uddannelsens tredje år er der mulighed for en decideret specialisering, som f.eks. kan være omkring 'Information and Computer Engineering'. *Bachelor i Computer Science* Uddannelsen kombinerer programmering, matematik og ingeniørfag med naturvidenskabelige-, psykologiske- og sproglige fag. Uddannelsen giver mulighed for forskellige specialiseringer, der dog ikke har udpræget fokus på data science, selvom der er mulighed for at tage fag i maskinlæring og algoritmer ved data mining.
Universitet	**Berkeley – University of California**
Uddannelsesnavn og niveau	Bachelor i Computer Science and Engineering Bachelor i Data Science
Link	Link til bachelor i Computer Science and Engineering Link til bachelor i Data Science Link til master i Information og Data Science
QA-Ranking:	7
Indhold	*Bachelor i Computer Science and Engineering* Uddannelsen kombinerer datalogi og ingeniør-faget med en række klassike fag inden for de to discipliner, som heriblandt programmering og software engineering. Uddannelsen er ikke særligt målrettet mod data science, der er dog mulighed for længere fremme på uddannelsen at tage et introducerende fag i databasessystemer, samt et mere videregående fag i databaser. *Bachelor i Data Science* Bacheloruddannelsen er netop etableret. Efter et pilothold i efteråret 2015 blev de første undervisningshold startet op i foråret 2016. Uddannelsen er etableret i samarbejde med en task force bestående af repræsentanter fra flere af universitetets fakulteter. Uddannelsen består bl.a. af et grundlæggende kursus i rationalerne bag data science, fag om dataetik, dataanalyse og – anvendelse ift. udfordringer indenfor byudvikling, samfundsmæssige emner samt individets sundhed. Desuden indeholder uddannelsen fag i statistik, it og visualisering af data. Berkeley Universitet har en forhåbning om, at denne uddannelse vil være udgangspunktet for en fremtidig udvikling indenfor arbejdet med og uddannelse indenfor data science. Der er mulighed for at læse en kandidat (master) i Information og Data Science, som er en online uddannelse. Der er linket til uddannelsen.
Universitet	**National University of Singapore**
Uddannelsesnavn og niveau	Bachelor i Computer Engineering
Link	Link til bachelor i Computer Engineering Link til kandidat i Business Analytics
QA-Ranking:	8
Indhold	*Bachelor i Computer Engineering* I starten af uddannelsen vil den studerende modtage en forholdsvis bred introduktion til matematik, informationsvidenskab og datalogi. Herefter vil der fokuseres på elektronisk og computermæssig udvikling, hvorefter der er mulighed for en specialisering. Det kunne f.eks. være inden for intelligente systemer eller stor-skala it-processer. Her er der fag i datastrukturer og algoritmer, databasesystemer, databasesystemers implementering, database-

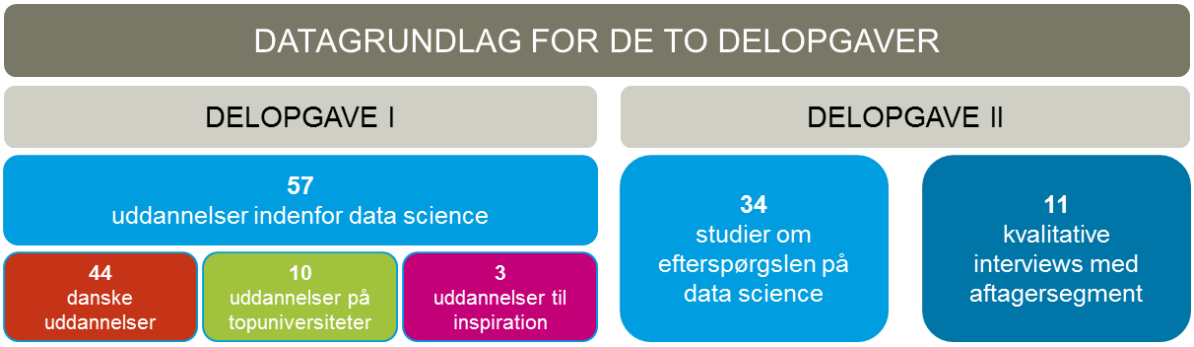
	applikationers design og optimering og maskinlæring. På universitetet er der mulighed for at læse en kandidat i Business Analytics, hvilket er den mest relevante mulighed på universitetet. Der er linket til uddannelsen. Derudover har universitetet også annonceret, at der kommer en bacheloruddannelse i Data Science og Analytics til august 2016.
Universitet	**Princeton University**
Uddannelsesnavn og niveau	Bachelor i Computer Science (ingeniøruddannelse)
Link	Link til bachelor i Computer Science Link til certifikat i Statistics and Machine Learning
QA-Ranking:	9
Indhold	I uddannelsen er der krav om at tre større fag skal gennemføres: et generelt fag i datalogi, et fag, der introducerer til programmeringssystemer, samt et fag om algoritmer og datastrukturer. Udover disse fag er der krav om fag inden for matematik, fysisk og kemi, hvilket er gældende for ingeniøruddannelser på Princeton. Herefter er der mulighed for at vælge flere forskellige fag til, fx faget 'Database og Information Management System' samt indenfor fundamental maskinlæring og informationssikkerhed. Som en del af uddannelsen kan der tages et certifikat i statistik og maskinlæring, som er et tillæg til uddannelsen. Dette indeholder blandt andet fag i analyse af Big Data og flere forskellige kurser i maskinlæring.
Universitet	**University of Toronto**
Uddannelsesnavn og niveau	Bachelor i Computer Science
Link	Link til bachelorforløb i Computer science
QA-Ranking:	10
Indhold	Uddannelsen er længere end en typisk bachelor og giver mulighed for at specialisere sig inden for informationssystemer. Uddannelsen giver grundlæggende kendskab til datalogi i form af moderne computer software og hardware, teoretiske aspekter og relevante områder inden for matematik og statistik. Herefter er der mulighed for at specialisere sig. Specialiseringen inden for informationssystemer har stort fokus på ingeniørdiscipliner, deriblandt computersystemer. Derudover undervises der i aspekter af forretningsforståelse og -styring. Uddannelsen er forholdsvis teknologisk, og indeholder fag om databaser, maskinlæring og data mining og databasesystem teknologi. Derudover indeholder specialiseringen også en del forretningsorienterede fag med fokus på især styring og ledelse.

BILAG 2 - DETALJERET METODEBESKRIVELSE

I dette kapitel beskriver vi det metodiske grundlag for markedsundersøgelsen nærmere. Kapitlet er inddelt i følgende to afsnit:

- I det første afsnit opridses den metodiske tilgang til *delopgave I*
- I det andet afsnit gennemgår vi den metodiske tilgang til *delopgave II.*

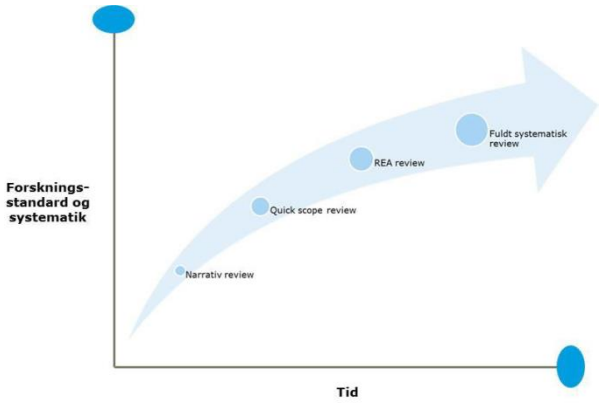
Figur 11: Markedsundersøgelsens samlede datagrundlag



Metodisk tilgang til *delopgave I*

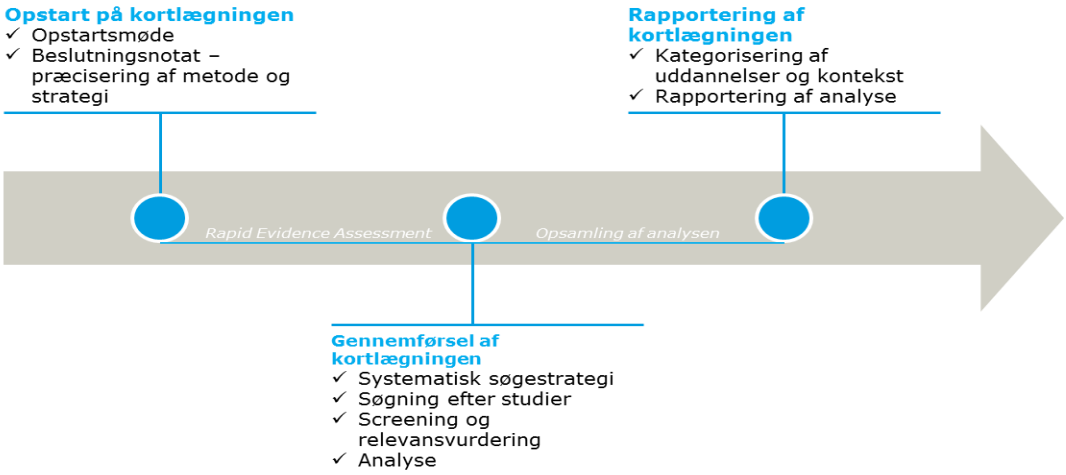
Delopgave I har haft til hensigt at kortlægge eksisterende uddannelser indenfor data science i Danmark, suppleret af enkelte relevante internationale uddannelser.

Kortlægningen af eksisterende danske uddannelser, har taget udgangspunkt i metoden Rapid Evidence Assessment (REA). Metoden sikrer en hurtig, systematisk og transparent tilgang til indsamling af viden, der er baseret på forskningsmæssige standarder. REA er inspireret af systematiske forskningsreviews, men udviklet med henblik på at kortlægge og indsamle eksisterende viden og evidens om konkrete sektorområder inden for meget kort tid. Metoden har en mere snæver afgrænsning end systematiske forskningsreviews, eksempelvis afsøges færre databaser, men metoden er kendetegnet ved den samme systematiske tilgang inden for den valgte afgrænsning.



Dette fremgår også af figuren til højre, der viser REA-metodens placering i et diagram – sammenlignet med andre review-metoder, hvor Y-aksen viser metodens forskningsmæssige standard og systematik, mens X-aksen viser tidsdimensionen i forhold til metodens indsamling af viden, screening og kvalitets- og relevansvurdering. Figuren viser, at REA-metodens kvalitet tidsmæssigt gør den anvendelig i nærværende opgaveløsning, der har en meget afgrænset tidsramme.

REA er baseret på en række operationelle trin. Trinene fremgår af figuren nedenfor – i figuren ses også, hvordan de metodiske tilgange til kortlægningen understøtter faserne og aktiviteterne.



Ovenstående trin har sikret en systematisk tilgang til opgavebesvarelsen, samt en løbende inddragelse af ITU, hvilket har understøttet en fælles forståelse og strategi for rapporten, samt givet mulighed for input fra ITU til fx. søgestrategien.

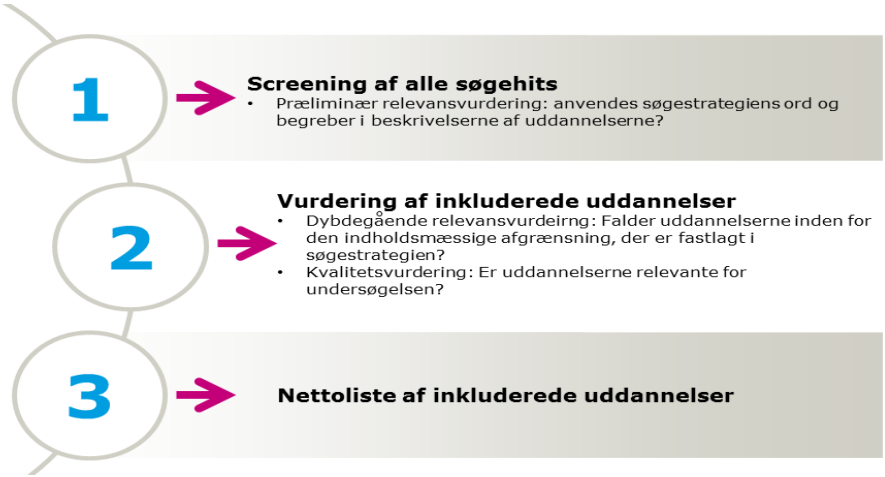
Søgestrategien har været essentiel for at sikre en dækkende kortlægningen af området. Søgestrategien har haft til formål at afgrænse genstandsfeltet via relevante inklusions- og eksklusionskriterier. Rambøll og ITU har løbende været i dialog omkring relevante søgetermer. Der er søgt på søgetermer på både dansk og engelsk. Nedenfor ses den anvendte søgestrategi.

Kategorier i søgestrategi	Indhold
Undersøgelsesspørgsmål	Hvilke danske uddannelser er der inden for data science og hvordan ser internationale, engelsksprogede, tilsvarende uddannelser ud?
Søgetermer	Data science, Big Data, statistik, maskinlæring, dataarkitektur, programmering, algoritmer, data management, data ingeniør, data mining, analytiker, data visualisering, business intelligence.
Inklusionskriterier	*Typer af uddannelser:* Bachelor-, kandidat, diplomingeniør-, civilingeniør-, professions- og efter- og videreuddannelse *Sproglig afgrænsning af uddannelser:* Dansk og engelsk
Kilder	Uddannelsesinstitutionernes hjemmesider: Danske universiteter: http://www.au.dk/ http://www.aau.dk/ http://www.cbs.dk/ http://www.dtu.dk/ http://www.itu.dk/ http://www.ku.dk/ http://www.ruc.dk/ http://www.sdu.dk/ Internationale universiteter: http://web.mit.edu/ https://www.stanford.edu/ http://www.ox.ac.uk/ http://www.harvard.edu/ http://www.cmu.edu/ http://www.cam.ac.uk/ http://www.berkeley.edu/ http://www.nus.edu.sg/ https://www.princeton.edu/main/ https://www.utoronto.ca/ QS World University Ranking: http://www.topuniversities.com/ Uddannelsesguiden: https://www.ug.dk/

Den korte tidsramme for kortlægningen har stillet store krav til balancen mellem alsidighed og præcision. Vores erfaringer fra tidligere kortlægninger viser, at når alsidighed øges, mindskes præcision med risiko for, at indsamlet litteratur falder uden for søgekriterierne. Det er derfor

afgørende, at søgningen er systematisk og fokuseret fra start. Undersøgelsen af danske uddannelser har derfor fra start haft fokus på de danske universiteters uddannelsesmuligheder. På denne baggrund er der udarbejdet en bruttoliste af danske uddannelser. Næste skridt var at gennemføre en systematisk og kritisk vurdering af de indsamlede uddannelser.

Figuren nedenfor illustrerer de forskellige trin i udvælgelse og vurdering af uddannelser



Udgangspunktet for søgningen efter uddannelser har i den første fase været uddannelsesguiden.dk[9]. På uddannelsesguidens præsentationsside til de enkelte uddannelser præsenteres også relaterede uddannelser, som også er blevet undersøgt.

1

Som det fremgår af figuren har alle uddannelser undergået en præliminær relevansvurdering og metodisk screening. Ved den præliminære relevansvurdering har vi vurderet beskrivelsen af uddannelsen ift. søgestrategiens ordpar eller relaterede begreber. Kun de uddannelser, som inkluderes på baggrund af første trin, underkastets en dybdegående relevansvurdering og kvalitetsvurdering i andet trin.

2

Den dybdegående relevansvurdering i andet trin omfatter en endelig vurdering af, om uddannelserne falder inden for kortlægningens indholdsmæssige afgrænsning. Her er der kigget nærmere uddannelsesinstitutionernes eget materiale om de enkelte uddannelser, og hvorvidt dette harmonerer med en uddannelse i data science. Her har vi især skelet til opbygningen af uddannelserne, deriblandt obligatoriske fag på uddannelserne og muligheder for valgfag, samt hvilke kombinationsmuligheder uddannelsen giver. På baggrund af dette har Rambøll foretaget en kvalitetsvurdering af, hvorvidt de enkelte uddannelser er relevante i forhold til oprettelsen af en uddannelse i data science. Ved de medtagne uddannelser er der derefter undersøgt antal ansøgere og optag på uddannelse, samt dimittendledighed, for også at kunne give en vurdering af, hvor efterspurgt uddannelserne er.

3

På baggrund af relevans- og kvalitetsvurderingen er der udarbejdet en nettoliste over de inkluderede 44 danske uddannelser inden for data science, som fremgår af bilag 1.

INTERNATIONALE UDDANNELSER

Tilgangen til de internationale, engelsksprogede uddannelser har baseret sig på samme søgestrategi som de danske uddannelser for at sikre en sammenhæng og til dels sammenlignelighed.
Da en systematisk søgning indenfor internationale uddannelser er udenfor opdragets ramme, har Rambøll valgt at foretage en afgrænsning i udvælgesen af de internationale uddannelser ud fra en 'best practice-tilgang'. Rambøll har ved brug af den meget anvendte internationale universitets rangliste "QS World University Rankings" udgivet af Quacquarelli Symonds, taget udgangspunkt i de 10 bedste universiteter, der har engelsksproget relevant uddannelse indenfor datalogi

[9] Uddannelsesguiden udgives af Ministeriet for Børn, Undervisning og Ligestilling, og indeholder alle anderkendte danske uddannelser.

og informationsvidenskab. Formålet med dette har været at afdække, hvilke muligheder der er for at studere data science på nogle af verdens bedste universiteter, samt hvordan de konkret tilrettelægger uddannelserne.

Uddannelserne på de ti udvalgte universiteter er, ligesom de danske uddannelser, undergået en dybdegående relevans- og kvalitetsvurdering, som er baseret på de samme fokuspunkter som ved de danske uddannelser.

Ranglisten "QS World University Ranking" er udvalgt af tre grunde:

4. Den er blandt de mest anvendte og anerkendte internationale ranglister for sammenligning af universiteter.[10]
5. Den underopdeler universiteter på deres evner inden for forskellige område, hvilket giver mulighed for at sikre, at de udvalgte universiteter er blandt de bedste i verden inden for det overliggende område, som data science opererer indenfor, nemlig datalogi og informationssystemer.
6. Den inddrager en indikator på, hvor arbejdsgivere mener, at de bedste kandidater uddannes. Ifølge QS er der ikke andre internationale raglister for universiteter, der medtager denne indikator. Det er relevant i forhold til denne kortlægning, da det kan give en klar indikation af efterspørgslen.

Ranglisten udgives af det britiske firma Quacquarelli Symonds, der er specialiseret i uddannelse, sammen med det hollandske akademiske forlag Elsevier, og indeholder mere end 800 universiteter. Som alle forsøg på at lave ranglister der sammenligner universiteter, har dette forsøg også en række svagheder og forbehold.[11] Rambøll har ikke belæg for at sige at QS' rangliste er hverken bedre eller dårligere end de 2-3 andre store globale universitets ranglister, men mener at denne liste kan give en fin indikation på, hvordan de bedste universiteter, inden for datalogi og informationssystemer, arbejder med data science og Big Data.

Metodisk tilgang til *delopgave II*
Delopgave II har haft til hensigt at kortlægge viden om og efterspørgslen på data science på det danske arbejdsmarked såvel som internationalt set.

I denne delopgave har vi foretaget en kortlægning af internationale samt nationale studier, der kunne belyse efterspørgslen på data science nationalt såvel som internationalt. Udover at formålet var at skabe et overblik over det nuværende arbejdsmarked for en uddannelse i data science, samt det potentielle arbejdsmarkedet primært i en 5-10 årig horisont, er det blevet afdækket hvilke brancher og jobfunktioner, som efterspørger de kompetencer, der typisk tillæres på en uddannelse i data science.

Data science er et forholdsvis nyt felt, hvorfor der findes flere forskellige definitioner på, hvad en data scientist er og gør. For at sikre en sammenhæng med delopgave I har vi valgt at tage udgangpunkt i mange af de samme søgetermer, som anvendes i denne, men har tilføjet ord og termer der relaterer sig til arbejdsmarked i højere grad end uddannelse (fx ord som *employment, job opportunities, arbejdsmarked, beskæftigelse* mv.*).* Da studierne anvender forskellige definitioner af data science/data scientists kan det medvirke til en udfordring i at sammenligne studierne, hvilket Rambøll har haft for øje i den analytiske fortolkning.

[10] Altbach, Philip G. 2010: The State of the Rankings. Link: https://www.insidehighered.com/views/2010/11/11/state-rankings, besøgt 19-04-2016.

[11] For uddybning af kritikken se f.eks. Blanchflower 2011: . Link: http://www.newstatesman.com/blogs/david-blanchflower/2011/09/world-university-faculty, eller Redden 2013: Link: https://www.insidehighered.com/news/2013/05/29/methodology-qs-rankings-comes-under-scrutiny

I den indledende håndsøgning har vi taget udgangspunkt i studier udpeget af interne og eksterne eksperter på området, og vi har ydermere anvendt 'snowballing-metoden', hvor referencelister på inkluderede studier er afsøgt for at opspore yderligere relevante studier. De identificerede studier har, som ved uddannelserne i *delopgave I,* gennemgået en nærmere både relevans – og kvalitetsvurdering ud fra studiets metodedesign, antal respondenter, geografi, udgivelsesår og udgiver. Den indledende søgning har givet anledning til et genbesøg af kriterierne vedrørende metodisk design, da større kortlægninger af høj evidens samt peer review-artikler er en mangelvare. Vores vurdering er derfor, at data science endnu ikke er et genstandsfelt for decideret forskning, hvorfor de inkluderede studier primært er undersøgelser udarbejdet af konsulent – eller rådgivningsvirksomheder. Oplysninger herom er i vid udstrækning blevet rapporteret i kortlægningen for at sikre gennemsigtighed.

Kortlægningens fokus er på efterspørgslen i Danmark, men både den indledende søgning og efterfølgende relevans og kvalitetsvurdering har vist, at der er gennemført et fåtal danske undersøgelser om efterspørgslen efter data scientists i Danmark. Derfor bygger kortlægningen også på en stor andel internationale studier.

KVALITATIVE INTERVIEWS MED AFTAGERSEGMENT

For at supplere og kvalificere kortlægningen af den danske efterspørgsel efter data scientists, er der desuden foretaget 11 kvalitative interviews med repræsentanter fra danske virksomheder og organisationer. I nedenstående figur ses en oversigt over fordelingen af respondenter i brancher/sektorer:



Rambøll har været i dialog med ITU om udvælgelsen af interviewpersonerne, og har modtaget en bruttoliste med personer og virksomheder med relevans for undersøgelsen. Da data science er et område, der er relevante indenfor flere og flere sektorer, har det været en prioritet at afdække et bredt felt potentielle aftagere indenfor forskellige brancher og sektorer.

Alle interviews har været semistrukturerede interviews, der er udført med udgangspunkt i en interviewguide med følgende tematikker:

- Nuværende erfaring og fremtidig ønske til arbejdet med data science
- Efterspurgte medarbejderprofiler og kompetencer
- Ønsker til potentiel uddannelse indenfor data science.

Respondenterne har primært haft stilling som chef eller direktør i virksomhedens teknologiske og analytiske afdelinger, mens respondenterne i de resterende interview ligeledes har været chefer, partnere eller direktører samt haft kendskab til data science på anden måde(se bilag 3 for oversigt over respondenter).

BILAG 3 – OVERSIGT OVER RESPONDENTER

- Adam Lebech, Direktør, Dansk Industri
- Birgitte Hass, Direktør, IT-branchen
- Jens Jakob Svanholt, Senior Vice President of PD and Decision Modeling, Danske Bank
- Peter Sergio Larsen, Chief Big Data & Data Science Specialist, Nordea
- Karsten Lund, Creative Director, LEGO
- Erik Verdoner, Business Intelligence Specialist, Copenhagen Airports
- Henrik Lauritzen, CTO, Audience Project
- Johan Kjeldgaard Pedersen, Direktør, Data2impact
- Lisbeth Nielsen, Direktør, Sundhedsdatastyrelsen
- Søren Ilsøe Overgaard, Underdirektør og Head of Technology, Data & Security, SKAT
- Anders Kofod Petersen, Head of Data Science and Engineering Lab, Alexandra Instituttet

BILAG 4 - REFERENCELISTE

A. T. Kearney, 2015: kan hentes på: The 2015 Leadership Ecellence in Analytic Practices (LEAP) Study – Answering the Digital and Analytics Talents Gap: The New 'Trilinguals'. Kan hentes på: http://www.atkearney.co.uk/documents/10192/7010822/LEAP+Answering+the+Digital+and+Analytics+Talent+Gap-The+New+Trilinguals.pdf/d05a1c57-023d-452c-9001-02d4a3b0737d, besøgt 26-04-2016

Aftaletekst, 2015: Aftale om vækstplan for digitalisering af Danmark. Kan hentes på: https://www.evm.dk/aftaler-og-udspil/15-02-26-aftale-om-vaekstplan-for-digitalisering, besøgt 26-04-2016

Altbach, Philip G. 2010: The State of the Rankings. Link: https://www.insidehighered.com/views/2010/11/11/state-rankings, besøgt 19-04-2016

Altbach, Philip G. 2012: The Globalization of College and University Ranking. Kan hentes på: http://www.tandfonline.com/doi/abs/10.1080/00091383.2012.636001, besøgt 19-04-2016

Blanchflower, David 2011: The QS World University Rankings are a load of old baloney. Link: http://www.newstatesman.com/blogs/david-blanchflower/2011/09/world-university-faculty, besøgt 19-04-2016

Burtch Work, 2014: Salaries of Data Scientist. Kan hentes på: http://www.burtchworks.com/files/2014/07/Burtch-Works-Study_DS_final.pdf, besøgt 27-04-2016

CrowdFlower, 2016: What skills should data scientists have in 2016? Kan hentes på: https://www.crowdflower.com/what-skills-should-data-scientists-have-in-2016/, besøgt 27-04-2016

Davenport,Tom, 2015: Why Data Storytelling is So Important, and Why We're So Bad at It. Link: http://deloitte.wsj.com/cio/2015/02/10/why-data-storytelling-is-so-important-and-why-were-so-bad-at-it/, besøgt 27-04-2016

EMC, 2011: Data Science Revealed: A Data-Driven Glimpse into the Burgeoning New Field. Kan hentes på: http://www.emc.com/collateral/about/news/emc-data-science-study-wp.pdf, besøgt 27-04-2016

Erhvervsstyrelsen, 2013: Big Data som vækstfaktor i dansk erhvervsliv – potentialer, barrierer og erhvervspolitiske konsekvenser. Kan hentes på: https://erhvervsstyrelsen.dk/sites/default/files/big-data-som-vaekstfaktor.pdf, besøgt 26-04-2016

Europa-Kommissionen, 2014: Communication from the commission to The European Parliament, The Council, The European Economic and Social Committee and The Committee of the Regions. Kan hentes på: https://ec.europa.eu/digital-single-market/en/news/communication-data-driven-economy, besøgt 25-04-2016

Europa-Kommissionen, 2016: Big Data Value Public-Private Partnership. Link: https://ec.europa.eu/digital-single-market/en/big-data-value-public-private-partnership, besøgt 25-04-2016

Gartner, 2015 Survey Analysis: Hadoop Adoption Drivers and Challenges. Kan hentes på: http://www.gartner.com/newsroom/id/3051717, besøgt 27-04-2016

Harvard Business Review, 2012a: Big data: The Management Revolution. Link: https://hbr.org/2012/10/big-data-the-management-revolution/ar, besøgt 25-04-2016

Harvard Business Review, 2012b: Data Scientist: The Sexiest Job of the 21st Century. Link: https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/, besøgt 25-04-2016

IBM, 2012: Bringing Big Data to enterprise. Link: https://www-01.ibm.com/software/in/data/bigdata/, besøgt: 25-04-2016

IDC, 2011: Worldwide Big data technology and service – 2012-2015 forecast. Kan hentes på: https://ec.europa.eu/digital-single-market/en/news/worldwide-big-data-technology-and-services-2012-2015-forecast, besøgt 25-04-2016

IDC, 2014: Market Analysis Perspective: Nordic Cloud and Industry Transformation, 2014— Enabling Technologies for the 3rd Platform. Kan hentes på: http://www.idc.com/getdoc.jsp?containerId=NCCS56W, besøgt 26-04-2016

IDC, 2015: Worldwide Big data technology and service – 2015-2019 forecast. Kan hentes på: http://www.idc.com/getdoc.jsp?containerId=259532, besøgt 25-04-2016

IIA, 2016: 2016 Analytics Predictions & Priorities. Kan hentes på: http://iianalytics.com/analytics-resources/2016-predictions, besøgt 27-04-2016

Interxion, 2013: Big Bata - Beyond the hype. Kan hentes på: http://www.interxion.com/globalassets/documents/whitepapers-and-pdfs/enterprise/big-data/WP_BigData_dk_0313.pdf, besøgt 26-04-2016

ITEK, 2015: Prognose 2020 -ITEK-branchens behov for it- og elektronikkandidater i 2020. Kan hentes på: http://digital.di.dk/SiteCollectionDocuments/ITOS/Publikationer/GC/Prognose2020-ITEK-branchensbehov_for_it_kandidater.pdf, besøgt 26-04-2016

Marszal, Andrew 2012: University rankings: which world university rankings should we trust? The Telegraph. Link: http://www.telegraph.co.uk/education/universityeducation/9584155/University-rankings-which-world-university-rankings-should-we-trust.html, besøgt 19-04-2016

Mckinsey, 2011: Big data: The next frontier for innovation, competition, and productivity, Kan hentes på: http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation, besøgt 25-04-2016

Mckinsey, 2013: Disruptive technologies: Advances that will transform life, business, and the global economy. Kan hentes på: http://www.mckinsey.com/business-functions/business-technology/our-insights/disruptive-technologies, besøgt 25-04-2016

O'Reilly, 2012: Analyzing the Analyzers – An Introspective Survey of Data Scientists and Their Work. Kan hentes på: http://cdn.oreillystatic.com/oreilly/radarreport/0636920029014/Analyzing_the_Analyzers.pdf, besøgt 27-04-2016

O'Reilly, 2015: 2015 Data Science Salary Survey - Tools, Trends, What Pays (and What Doesn't) for Data Professionals. Kan hentes på: http://duu86o6n09pv.cloudfront.net/reports/2015-data-science-salary-survey.pdf, besøgt 27-04-2016

OECD, 2015: Data-Driven Innovation - Big Data for Growth and Well-Being, Kan hentes på: http://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm, besøgt 25-04-2016

Partil, DJ, 2011: Building data science teams. Kan hentes på: http://radar.oreilly.com/2011/09/building-data-science-teams.html, besøgt 27-04-2016

QS, 2015: QS World University Rankings: Methodology. Link: http://www.topuniversities.com/university-rankings-articles/world-university-rankings/qs-world-university-rankings-methodology, besøgt 19-04-2016

Redden, Elizabeth, 2013: Scrutiny of QS Rankings. Link: https://www.insidehighered.com/news/2013/05/29/methodology-qs-rankings-comes-under-scrutiny, besøgt 19-04-2016

Regeringen, 2015: Redegørelse for Danmarks digitale vækst 2015. Kan hentes på: https://www.evm.dk/publikationer/2015/15-05-13-redegorelse-om-danmarks-digitale-vaekst-2015, besøgt 26-04-2016

RJMetrics, 2015: The State of Data Science. Link: https://rjmetrics.com/resources/reports/the-state-of-data-science/, besøgt 26-04-2016

SAS 2013b: Big Data Analytics: Adoption and Employment Trends, 2012–2017. Kan hentes på: http://www.sas.com/offices/europe/uk/downloads/bigdata/eskills/eskills.pdf, besøgt 26-04-2016

SAS,2013a. Big Data Analytics Adoption and Employment Trends, 2012-2017. Kan hentes på: http://www.sas.com/offices/europe/uk/downloads/bigdata/eskills/eskills.pdf, besøgt 26-04-2016

SAS, 2014: Big Data Analytics - Assessment of Demand for Labour and Skills 2013–2020. Kan hentes på: https://www.thetechpartnership.com/globalassets/pdfs/research-2014/bigdata_report_nov14.pdf, besøgt 26-04-2016

SIRA, 2014: European Big Data Value Strategic Research & Innovation Agenda. Kan hentes på: http://www.nessi-europe.eu/Files/Private/EuropeanBigDataValuePartnership_SRIA__v099%20v4.pdf, besøgt 25-04-2016

White House, 2014: Big data: Seizing opportunities, preserving values. Kan hentes på: https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf, besøgt 25-04-2016

WISE, 2014: Big and open data in Europe: A growth engine or a missed opportunity?. Kan hentes på: https://www.microsoft.com/global/eu/RenderingAssets/pdf/2014%20Jan%2028%20EMEA%20Big%20and%20Open%20Data%20Report%20-%20Final%20Report.pdf, besøgt 26-04-2016

World Economic Forum, 2012: Big Data, Big Impact: New Possibilities for International Development. Kan hentes på: https://www.weforum.org/reports/big-data-big-impact-new-possibilities-international-development/, besøgt 25-04-2016

Zagenberg & Company, 2011: Kvantificering af værdien af åbne offentlige data. Kan hentes på: https://digitaliser.dk/resource/1021067 , besøgt 26-04-2016

Referencer

(Danmarks Vækstråd, Anbefalinger vedr. datadreven udvikling og vækst. 2015) *Danmarks Vækstråds anbefalinger vedr. datadreven udvikling og vækst*. Marts 2015. Danmarks Vækstråd. http://danmarksvaekstraad.dk/file/555170/Big_Data.pdf

(DI: Stort behov for øget dansk digitalisering. 2016) DI: *Stort behov for øget dansk digitalisering.* Pressemeddelelse (Maj 2016) https://di.dk/DI/Presse/Pressemeddelelser/Pages/DI-Stort-behov-for-oeget-dansk-digitalisering.aspx

(DI ITEK, ITEK-branchens behov for it- og elektronikkandidater i 2020. 2015) *ITEK-branchens behov for it- og elektronikkandidater i 2020*. 2015. http://digital.di.dk/SiteCollectionDocuments/ITOS/Publikationer/GC/Prognose2020-ITEK-branchensbehov_for_it_kandidater.pdf

(Erhvervs og Vækstministeriet: Redegørelse om Danmarks digitale vækst 2016) *Redegørelse om Danmarks digitale vækst 2016.* 2016. Erhvervs og Vækstministeriet.

(a) https://www.evm.dk/~/media/files/2016/16-05-31-redegorelse-f-danmarks-digitale-vaekst.ashx

(b) https://www.evm.dk/publikationer/2016/16-05-31-redegorelse-om-digital-vaekst

(Erhvervsstyrelsen, Big Data som vækstfaktor i dansk erhvervsliv. 2013) *Big Data som vækstfaktor i dansk erhvervsliv – potentialer, barrierer og erhvervspolitiske konsekvenser.* 2013. Erhvervsstyrelsen. https://erhvervsstyrelsen.dk/sites/default/files/big-data-som-vaekstfaktor.pdf

(Erhvervsstyrelsen, digitale kompetencer. 2016) *Kortlægning af virksomhedernes behov for digitale kompetencer.* 2016. Erhvervsstyrelsen. https://erhvervsstyrelsen.dk/kortlaegning-af-virksomhedernes-behov-digitale-kompetencer

(Gyldendal) *Data - virksomhedens nye grundstof*. Mikkel Holm Sørensen og Simon Bentholm (Gyldendal, 2013)

(Harvard Business Review 2012) *Data Scientist: The Sexiest Job of the 21st Century.* Harvard Business Review, (Oktober 2012). https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

(IDC, Worldwide Big data technology and service - 2012-2015 forecast. 2011.) IDC, *Worldwide Big data technology and service* - 2012-2015 forecast. 2011 https://ec.europa.eu/digital-single-market/en/news/worldwide-big-data-technology-and-services-2012-2015-forecast

(IRIS Group, Datadreven vækst i Danmark. 2015) *Datadreven vækst i Danmark--Potentialer og udfordringer ved big data i udvalgte sektorer*. 2015. IRIS Group. http://danmarksvaekstraad.dk/file/555171/Baggrundsrapport_IRIS_Group.pdf

(IRIS Group, Digitalisering af dansk erhvervsliv. 2013) *Digitalisering af dansk erhvervsliv--en undersøgelse af it-anvendelse i danske virksomheder*. 2013. IRIS Group. https://erhvervsstyrelsen.dk/digitalisering-af-dansk-erhvervsliv-en-undersoegelse-af-it-anvendelse-i-danske-virksomheder

(Job Listings Analysis. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*). *Technical Report: An empirical analysis of the current Data Science Job Market in Denmark*. (September 2016). ITU.

(Lisbeth Nielsen, Direktør, Sundhedsdatastyrelsen. 2016). Lisbeth Nielsen. Personal interview with IT University of Copenhagen Communications.

(Mikkel Holm Sørensen, Direktør, KL7. 2016) Mikkel Holm Sørensen. Personal interview with IT University of Copenhagen Communications.

(Minutes Employers Meeting. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*). *Minutes Employers Meeting.* (May 2016)

(OECD, An OECD horizon scan of megatrends and technology trends in the context of future research policy. 2016) *An OECD horizon scan of megatrends and technology trends in the context of future research policy.* 2016. OECD for Styrelsen for Forskning og Innovation. http://ufm.dk/publikationer/2016/an-oecd-horizon-scan-of-megatrends-and-technology-trends-in-the-context-of-future-research-policy

https://erhvervsstyrelsen.dk/kortlaegning-af-virksomhedernes-behov-digitale-kompetencer

(Rambøll-Markedsundersøgelse. 2016. In upload: *Dokumentation af efterspørgsel på uddannelsesprofil*). Markedsundersøgelse for data science. Rambøll. (May 2016). Report commissioned by ITU.

Uddannelses- og
Forskningsministeriet

IT-Universitetet i København
E-mail: itu@itu.dk

Afgørelse om godkendelse af ny uddannelse

Uddannelses- og forskningsministeren har på baggrund af gennemført prækvalifikation af IT-Universitetets ansøgning om godkendelse af bacheloruddannelse i data science truffet følgende afgørelse:

Godkendelse af bacheloruddannelse i data science

Afgørelsen er truffet i medfør af § 17 i bekendtgørelse nr. 852 af 3. juli 2015 om akkreditering af videregående uddannelsesinstitutioner og godkendelse af videregående uddannelser.

Da ITU er positivt institutionsakkrediteret gives godkendelsen til oprettelse af bacheloruddannelsen uden forudgående uddannelsesakkreditering.

Styrelsen for Videregående Uddannelser kontakter snarest IT-Universitetet med koder fra Danmarks Statistik.

Det er en forudsætning for godkendelsen, at uddannelsen og dennes studieordning skal opfylde uddannelsesreglerne, herunder bekendtgørelse nr. 1328 af 15. november 2016 om bachelor- og kandidatuddannelser ved universiteterne (uddannelsesbekendtgørelsen).

Ansøgningen er blevet vurderet af Det rådgivende udvalg for vurdering af udbud af videregående uddannelser (RUVU). Vurderingen er vedlagt som bilag.

Uddannelsen er omfattet af reglerne i uddannelsesbekendtgørelsen.

KOT-nummer tildeles automatisk fra Den Koordinerede Tilmelding, når udbuddet er oprettet af uddannelsesinstitutionen på Optagelse.dk. Optagelse.dk åbner for udbudsoprettelse i uge 43

Hovedområde:
Uddannelsen hører under det naturvidenskabelige område.

Titel:
Efter reglerne i uddannelsesbekendtgørelsens § 14, stk. 3, fastlægges uddannelsens titel til:

Dansk:	Bachelor (BSc) i datavidenskab
Engelsk:	Bachelor of Science (BSc) in Data Science

19. januar 2017

Styrelsen for Videregående Uddannelser

Bredgade 43
1260 København K
Tel. 7231 7800
Fax 7231 7801
Mail uds@uds.dk
Web www.ufm.dk

CVR-nr. 3404 2012

Ref.-nr. 16/044833-21

Udbudssted:
Uddannelsen udbydes i København.

Sprog:
Ministeriet har noteret sig, at uddannelsen udbydes på engelsk.

Ministeriet bemærker hertil, at det fremgår af § 9, stk. 1, i bekendtgørelse nr. 257 af 18. marts 2015 om adgang til bacheloruddannelser ved universiteterne (bacheloradgangsbekendtgørelsen), at hvis en uddannelse eller væsentlige dele heraf udbydes på engelsk, skal ansøgeren senest inden det tidspunkt, der er fastsat for studiestarten, dokumentere kundskaber i engelsk svarende til mindst engelsk B-niveau.

Normeret studietid:
Efter reglerne i uddannelsesbekendtgørelsens § 13 fastlægges uddannelsens normering til 180 ECTS-point.

Takstindplacering:
Uddannelsen indplaceres til: heltidstakst 3.
Aktivitetsgruppekode: 6255 (bacheloruddannelse i data science).

Censorkorps:
Ministeriet har noteret sig, at uddannelsen tilknyttes censorkorpset for datalogi. Det er muligt at supplere censorkorpset, således at det samlede korps bl.a. dækker alle de fag/fagelementer, der indgår i uddannelsen.

Maksimumramme:
Styrelsen har meddelt ministeriet, at styrelsen ikke ønsker at fastsætte en maksimumsramme for tilgangen til uddannelsen. Universitetet bestemmer derfor selv efter reglerne om frit optag, hvor mange studerende der optages på uddannelsen, jf. § 11, stk. 1, i bekendtgørelse nr. 257 af 18. marts 2015 om adgang til bacheloruddannelser ved universiteterne (bacheloradgangsbekendtgørelsen) med senere ændringer.

Ministeriet har noteret sig, at universitetet har fastsat en maksimumramme på 50 studerende for tilgangen til uddannelsen det første år uddannelsen udbydes (2017).

Adgangskrav:
Adgangskravene til bacheloruddannelsen i datavidenskab vil blive fastsat til:
- Engelsk B
- Matematik A

Herudover har universitetet fastsat, at både Engelsk B og Matematik A hver især skal være bestået med karakteren mindst 6.

Med venlig hilsen

Jørgen Sørensen

Nr. A5 - Ny uddannelse – prækvalifikation (efterår 2016)			**Status på ansøgningen:** *Godkendt*	
Ansøger og udbudssted:	IT-Universitetet (København)			
Uddannelsens type/navn (fagbetegnelse):	Bacheloruddannelse i Data Science			
Den uddannedes titler på hhv. da/eng:	- Bachelor (BSc) i Data Science - Bachelor (BSc) of Science in Data Science			
Hovedområde:	Naturvidenskab		**Genansøgning:** *(ja/nej)*	Nej
Sprog:	Engelsk	**Antal ECTS:**	180 ECTS	
Link til ansøgning på http://pkf.ufm.dk:	http://pkf.ufm.dk/flows/3704d145882a4305254cb3e2d10fb7e1			
Om uddannelsen: indhold og erhvervssigte	**Beskrivelse af den nye uddannelse, dens konstituerende elementer/struktur, erhvervssigte og adgangskrav**			
Beskrivelse af uddannelsen:	Ansøger nævner, at der ikke findes en etableret dansk oversættelse af data scientist/data science, men formålet er at udvikle en tværfaglig bacheloruddannelse for analyse og anvendelse af data. De eksisterende uddannelser rettet mod dataanalyse mangler dele af paletten af relevante kompetencer, og Data Science skal derfor inkludere fagområder "fra datalogi, statistik, kvantitativ sociologi såvel som komplementære kompetencer i teknisk kommunikation og etiske aspekter af databehandling" (ansøgningsskemaet).			
Konstituerende faglige elementer/struktur:	Uddannelsen opbygges af en række obligatoriske moduler svarende til i alt 135 ECTS, retningsspecifikke og valgfrie fag svarende til hver 15 ECTS og et bachelorprojekt på 15 ECTS. Kurserne (engelsksprogede) er bl.a. følgende: Data Science and Programming, Applied Statistics, Data Science in Research, Business and Society, Algorithms and Data Structures, Linear Algebra and Optimization, Data Management, Data Visualization and Data-driven Decision Making, Technical Communication, Security and Privacy.			
Erhvervssigte:	De færdige bachelorer formodes at kunne gå direkte ud på jobmarkedet og arbejde med "generelle data science problemstillinger". Derudover vil de kunne fortsætte på relevante kandidatuddannelser på f.eks. ITU's kandidat i Softwareudvikling eller i Digital Innovation and Management (forretnings-it).			
Adgangskrav:	Dansk eller international gymnasial uddannelse med to specifikke adgangskrav: Matematik A med 6 i gennemsnit. Engelsk B med 6 i gennemsnit.			
Forventet optag på uddannelsen:	50 årligt			
RUVU's vurdering på møde d. 3. november 2016:	RUVU vurderer, at ansøgningen opfylder kriterierne, som fastsat i bekendtgørelse nr. 852 af 3. juli 2015, bilag 4. RUVU har ved vurderingen lagt vægt på, at behovsafdækningen dokumenterer en aktuel efterspørgsel efter data scientists, og at denne efterspørgsel forventes at være stigende fremover. RUVU bemærker, at det er væsentligt, at der er tale om en tværfaglig bacheloruddannelse inden for analyse og anvendelse af (big) data, og at det fremgår, at de færdiguddannede bachelorer forventes at kunne gå direkte ud på jobmarkedet. RUVU noterer, at der er tale om en engelsksproget bacheloruddannelse, hvilket på det pågældende fagområde forekommer relevant i det konkrete tilfælde.			