



**Uddannelses- og  
Forskningsministeriet**

**Prækvalifikation af videregående uddannelser - Datavidenskab**

Udskrevet 27. april 2026

## Kandidat - Datavidenskab - IT Universitetet i København

Institutionsnavn: IT Universitetet i København

Indsendt: 01/02-2019 07:30

Ansøgningsrunde: 2019-1

Status på ansøgning: Godkendt

[Afgørelsesbilag](#)

[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?

Ja

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

Nej

### Uddannelsestype

Kandidat

### Uddannelsens fagbetegnelse på dansk fx. kemi

Datavidenskab

### Uddannelsens fagbetegnelse på engelsk fx. chemistry

Data Science

### Den uddannedes titel på dansk

Cand.scient. i datavidenskab

### Den uddannedes titel på engelsk

Master of Science (MSc) in Data Science

**Hvilket hovedområde hører uddannelsen under?**

Naturvidenskab

**Hvilke adgangskrav gælder til uddannelsen?****Generelle krav for optagelse på IT-Universitetets kandidatuddannelser:**

Universitetsbacheloruddannelse eller en professionsbacheloruddannelse. Ved anden uddannelse (f.eks. en international uddannelse), der kan sidestilles med en universitetsbacheloruddannelse eller en professionsbacheloruddannelse, vil IT-Universitetet foretage en individuel vurdering af den pågældende uddannelse. Fælles forudsætning for optagelse på alle kandidatuddannelserne er et sprogkrav om engelsk på B-niveau.

**Specifikke adgangskrav for optagelse på kandidatuddannelsen i datavidenskab:**

For at opfylde de faglige forudsætninger på uddannelsens obligatoriske kurser inden for datavidenskab, skal ansøgere have en bacheloruddannelse eller tilsvarende, i datavidenskab eller tilsvarende, der har dækket følgende emner:

- Lineær algebra og optimering (Linear algebra and optimization)
- Algoritmer og datastrukturer (Algorithms and data structures)
- Introduktion til maskinlæring (Introduction to machine learning)
- Anvendt statistik (Applied statistics)- Grundlæggende programmering: funktionel programmering, objektorienteret programmering (Introduction to programming (functional programming, object oriented programming)
- Introduktion til databaser (introduction to databases)
- Grundlæggende grafteori (basic graph theory)

Mindst 80% af emnerne må have været dækket i løbet af bacheloruddannelsen.

**Retskrav:**

Følgende bacheloruddannelser har retskrav til kandidatuddannelsen:

- Bachelor i datavidenskab (IT-Universitetet)

**Bacheloruddannelser der opfylder adgangskravene:**

- Bachelor i datavidenskab (Århus Universitet)
- Bachelor i datavidenskab (Aalborg Universitet)
- Bachelor i machine learning og data science (Københavns Universitet)
- Bachelor i kunstig intelligens og data (DTU)

**Er det et internationalt uddannelsessamarbejde, herunder Erasmus, fællesuddannelse og lign.?**

Nej

**Hvis ja, hvilket samarbejde?****Hvilket sprog udbydes uddannelsen på?**

Engelsk

**Er uddannelsen primært baseret på e-læring?**

Nej

**ECTS-omfang**

120

**Beskrivelse af uddannelsens formål og erhvervsigte**

IT-Universitetet i København (ITU) ansøger om at oprette en kandidatuddannelse (cand.scient.) i Data Science (datavidenskab), som bygger på kernekompetencer lig dem, der kan erhverves på ITU's bacheloruddannelse i Data Science. Således har den nye uddannelse, som uddybet nedenfor, fokus på stærke tekniske kompetencer med et dybt teoretisk fundament. Kandidaterne skal varetage højt specialiserede jobfunktioner inden for analyse og anvendelse af data, hvor der pt. er stor mangel på arbejdskraft.

Der er stor uopfyldt efterspørgsel efter højt kvalificerede medarbejdere med data science kompetencer såvel nationalt som regionalt, og danske virksomheder har svært ved at rekruttere ansatte med de nødvendige dybdegående tekniske kompetencer. Denne efterspørgsel forventes desuden at stige betydeligt over de næste ti år. Se punktet *Underbygget skøn over det nationale og regionale behov for dimittender* for citater fra analyser af disse forhold. Regeringens *Strategi for Danmarks digitale vækst* (Erhvervsministeriet 2018) fremhæver virksomheders analyse og anvendelse af egne og offentlige data som et strategisk område med stort potentiale, men også som et område hvor danske virksomheder halter bagefter virksomheder i nabolandene.

Uddannelsens fokus på stærke tekniske kompetencer med et dybt teoretisk fundament understøttes af en repræsentativ gruppe af aftagere, der dækker et bredt udsnit af brancher, såvel private som offentlige institutioner, som vi har involveret i designet af uddannelsen. Se punkterne *Hvordan er det konkret sikret, at den nye uddannelse matcher det påviste behov?* for analyser af disse forhold og *Hvilke aftagere har været inddraget i behovsundersøgelsen?* Helt generelt bifaldt de involverede aftagere det overordnede design: "The panel finds that it is very good that the program ensures that the theoretical skills are also combined with practical application. The applied approach is very useful for employers and companies when hiring graduates and the "Profound and Applied" should be the ITU MSc in Data Science profile in relation to other Universities MSc Data Science programs." (se *Minutes employers meeting 24 October* i *Dokumentation af efterspørgsel på uddannelsesprofil*). Se også uddannelsens samlede mål for læringsudbytte i *The Academic Profile and Objectives for the Programme* i *Programme Design* i *Dokumentation af efterspørgsel på uddannelsesprofil* s.56.

Uddannelsen vil drage fordel af et stærkt forskningsmiljø på ITU inden for både datalogi og data science og andre fagområder, der er relevante for en solid data science faglighed. Institut for Datalogi på IT-Universitetet har rekrutteret meget aktivt i de seneste tre år, både bredt i datalogi og mere specifikt i data science. Medio 2019 vil data science gruppen omfatte fire faste videnskabelige medarbejdere inden for sprogteknologi (natural language processing, NLP), tre inden for netværksanalyse (fx store sociale grafer), og et par stykker inden for machine learning. Yderligere vækst forventes i de kommende år. Alle har en stærk teknisk baggrund i datalogi, matematisk fysik eller teoretisk statistik, og flertallet har udenlandske phd-grader (England, Holland, Irland, Italien, Kroatien, Østrig) og et stærkt internationalt netværk. Det betyder, at de studerende vil befinde sig i et studiemiljø, hvor de lærer den nyeste data science i forskning og i praksis.

Aftagergruppen har undervejs givet ITU feedback, som er blevet indarbejdet i det endelige uddannelsesdesign. Som det ses af efterfølgende repræsentative skriftlige svar fra aftager Jens Lund (konsulent, tidligere Nordea - Executive Director, Head of Analytics and Business Design) finder han, at ITU har formået at indarbejde aftagernes behov: *"Overall, I find that you have managed to find a very good balance in the proposed design of a MSc in Data Science education"* og at ITU med uddannelsen kan være med til at sikre Danmarks indflydelse på data science området og ikke overlade det til store tech-giganter eller lande som Kina: *"In my opinion, it is a field that will be growing exponentially over the coming years, and it will be of importance for Denmark as a society that we become part of this development in order to not be completely overtaken by big tech giants (Google, Apple, Amazon, Facebook, ...) or countries like China, that is spending enormous amounts of resources in this area and moves forward at a quick pace."* (mail 13. december 2018).

Efterspørgslen efter data science kompetencer vokser på mange områder inden for erhvervsliv og forskning. Data science "has become increasingly important across all industries" (er blevet stadig vigtigere på tværs af alle brancher over hele verden) jf. LinkedIn Workforce Report fra 2018 (LinkedIn 2018). Dette gælder også i Danmark, hvor en analyse af danske jobannoncer inden for data science viser en stor efterspørgsel i mange forskellige brancher (Plank 2018) (Se *Analysis of the Danish Data Science Job Market* i *Dokumentation af efterspørgsel på uddannelsesprofil*). Kandidater i data science er blandt de højest uddannede it-medarbejdere og bidrager afgørende til konkurrencedygtighed og vækst. Deres kompetencer er efterspurgt i næsten alle erhverv (finans, forsikring, kommunikation, transport, detailhandel, robotindustri, automatisering, produktion, kemisk industri, turisme, medier og mange andre brancher) samt i digitaliseringen af den offentlige sektor.

Uddannelsens erhvervssigte er, at kandidaterne først og fremmest vil finde beskæftigelse i det private erhvervsliv, og en mindre del også i den offentlige sektor, begrundet i *Markedsundersøgelse kandidatuddannelse i data science* udfærdiget af Rambøll Management Consulting december 2018 for ITU (Rambøll 2018) (Se rapporten i *Dokumentation af efterspørgsel på uddannelsesprofil*). Rapportens 10 interviews med såvel private som offentlige aftagere (Rambøll 2018, s. 7) viser for eksempel: *"at der eksisterer et aktuelt efterspørgselsbehov på data scientist understøttes også af de kvalitative interviews. Denne tendens går igen i flere af undersøgelsens interviews. André Stelsig Rogaczewski fra Netcompany påpeger, at: "IT-branchen på nuværende tidspunkt kan opsuge alle kandidater"*. Se punktet *Underbygget skøn over det nationale og regionale behov for dimittender*.

Formålet med den nye uddannelse er at uddanne kandidater i Data Science, som kan varetage højt specialiserede jobfunktioner, hvor der pt. er mangel på arbejdskraft. Uddannelsen optager studerende, der har en bachelorgrad i Data Science eller lignende, fra ITU, fra andre danske universiteter eller fra udenlandske universiteter. ITU's egen BSc i Data Science (hvor ansøgere forventes hovedsageligt at komme fra, se *Forventet optag på de første 3 år af uddannelsen*) har hidtil haft et stort antal højt kvalificerede ansøgere, og efterspørgslen efter uddannelsen stiger. I 2017 var antallet af førsteprioritetsansøgere til ITU's BSc i Data Science næsten tre gange højere end der var kapacitet til at optage. Antallet af førsteprioritetsansøgere steg derefter med 30% i 2018 (se *Forventet optag på de første 3 år af uddannelsen*). Kvaliteten af ansøgere er meget høj for en teknisk uddannelse, med en adgangskvotient for kvote 1 på 9,0 i 2018.

En rundspørge blandt andetårsstuderende på BSc i Data Science på ITU (med spørgsmål om fremtidige uddannelsesplaner) med en svarprocent på over 55%, viste at 87.5% af de studerende ville fortsætte på en MSc i Data Science på ITU, hvis det bliver en mulighed (se *Overgang BSc in Data Science til MSc in Data Science* i *Dokumentation af efterspørgsel på uddannelsesprofil*).

Et stort antal højt kvalificerede ansøgere samt resultatet af rundspørgen viser, at den foreslåede MSc i Data Science har potentiale til at levere de efterspurgte kandidater i Data Science.

Der er klart plads til og behov for yderligere en kandidatuddannelse i Data Science i Danmark, dels fordi den giver stærke tekniske kompetencer med dybt teoretisk fundament som efterspurgt i erhvervslivet, og dels fordi der er et stort umættet behov på arbejdsmarkedet for kandidater med data science kompetencer i Danmark. Behovet er særlig stort i hovedstadsområdet, som vist af en analyse af danske jobannoncer (Plank 2018). Den ansøgte kandidatuddannelse bygger på kernekompetencer erhvervet i en forudgående bacheloruddannelse i Data Science. For en beskrivelse af relationen til de eksisterende kandidatuddannelser, se punktet *Beskriv ligheder og forskelle til beslægtede uddannelser, herunder beskæftigelse og eventuel dimensionering*.

Ved at oprette den nye uddannelse ønsker vi at uddanne flere data science kandidater, der er i stand til at arbejde på engelsk på et højt fagligt niveau. Den foreslåede uddannelse vil blive undervist på engelsk af flere grunde, som yderligere uddybet i *Kort redegørelse for det nationale og regionale behov for den nye uddannelse*.

For det første er arbejds sproget i it-virksomheder meget ofte engelsk, hvilket også afspejles i en analyse af danske jobannoncer (Plank 2018), hvor 66,3% af 1.390 data science orienterede jobopslag var på engelsk. Selv på offentlige arbejdspladser, hvor arbejds sproget ofte er dansk, opfattes data science som en engelsksproget disciplin. Lars Kjeldgaard (Senior Data Scientist, Skatteforvaltningen) skriver "*We also support an English-language program. Data Science is an international discipline by nature, and all of the literature on Data Science is in English*" (mail 21. december 2018).

For det andet kan en engelsksproget uddannelse være med til at tiltrække internationale talenter til det danske jobmarked, og erfaringer fra sammenlignelige kandidatuddannelser på ITU viser, at 63,4% af de udenlandske kandidater 1-2 år efter dimission er bosatte og beskæftigede i Danmark.

For det tredje er der mangel på danske forskere i data science, og uden dem er det svært at levere forskningsbaseret uddannelse af høj kvalitet på dansk. De videnskabelige medarbejdere på ITU's institut for datalogi kommer pt. fra godt 15 forskellige lande, og vi vil fortsat være nødt til at rekruttere forskere og undervisere fra udlandet.

Selv om uddannelsen undervises på engelsk, vil et flertal af kandidaterne være dansktalende, da vi forventer at uddannelsen overvejende rekrutterer fra ITU's eksisterende bacheloruddannelse i data science, der fremover vil have Dansk A som adgangskrav. Dette imødekommer de aftagere, især i den offentlige sektor, der foretrækker dansktalende kandidater. Rambølls markedsundersøgelse konkluderer netop at aftagerne foretrækker kandidater med både gode dansk- og engelskkundskaber (Rambøll 2018 s.10). Ydermere er det en fordel for de danske studerende at uddannelsen undervises på engelsk, da data science grundlæggende er en international disciplin, hvor arbejds sproget er engelsk. Ifølge aftagerrepræsentant Kaare Brandt Petersen, Advanced Data Analysis, Implement er det: "*Kritisk for de studerendes arbejdskompetencer, at lære at begå sig fagligt på engelsk og det er meget væsentligt, at de studerende fra et tidligt tidspunkt lærer at bruge engelsk som et dagligdags arbejdsprog med kolleger / medstuderende*"(mail 10. december 2018).

### Uddannelsens struktur og konstituerende faglige elementer

Data Science er en tværfaglig disciplin per definition. Den foreslåede kandidatuddannelse i Data Science på ITU er designet efter tre færdighedsakser (se nedenfor), for at de studerende kan tilegne sig stærke tekniske kompetencer. Den foreslåede kandidatuddannelse i Data Science omfatter 90 ECTS obligatoriske studieaktiviteter og 30 ECTS valgfrie studieaktiviteter (kurser eller projekter). De obligatoriske studieaktiviteter indeholder 37,5 ECTS kurser, 7,5 ECTS tværfagligt projekt, og afsluttes med 15 ECTS forskningsprojekt samt 30 ECTS kandidatspeciale. I alt 120 ECTS konstituerende elementer.

Uddannelsen bygger videre på den studerendes bacheloruddannelse i Data Science eller tilsvarende.

Kandidatuddannelsen er organiseret i fire semestre med følgende indhold:

- Første semester: Tre obligatoriske kurser à 7,5 ECTS og et 7,5 ECTS tværfagligt projekt.
- Andet semester: To obligatoriske kurser à 7,5 ECTS og to valgfrie studieaktiviteter à 7,5 ECTS.
- Tredje semester: En 15 ECTS valgfri videregående studieaktivitet, og et 15 ECTS forskningsprojekt, som forberedelse til kandidatspecialet. Projektet indeholder en obligatorisk aktivitet, hvor de studerende skal udarbejde en reflekterende rapport om de seneste etiske diskussioner inden for data science.
- Fjerde semester: 30 ECTS kandidatspeciale.

Se grafisk fremstilling i bilag *Programme Design* i *Dokumentation af efterspørgsel på uddannelsesprofil* s.53.

De obligatoriske kurser på første og andet semester er designet i henhold til tre færdighedsakser (som skitseret i vedhæftede *Programme design*):

1. Det matematiske og algoritmiske fundament for Data Science.
2. Udforskning af heterogene data, avanceret visualisering og involvering af interessenter i projektbaserede kurser
3. Kommunikationsfærdigheder og etiske overvejelser i forbindelse med Data Science.

Kurserne er som følger.

Akse 1:

- *Algorithm Design* (7,5 ECTS): hvor den studerende lærer at designe algoritmer til effektivt at løse givne it-problemer.
- *Advanced Applied Statistics & Multivariate Calculus* (7,5 ECTS): hvor den studerende lærer avanceret multivariat statistisk analyse og inferens.

Akse 2:

- *Data in the Wild: Processing heterogeneous data* (7,5 ECTS): hvor den studerende lærer at finde, bruge og kombinere data og at vælge forskellige dataanalysemetoder (f.eks., machine learning, SQL).
- *Data Science in Production* (7,5 ECTS): hvor den studerende får erfaring med at arbejde med komplekse kodebaser og produktionsniveau kode, at integrere deres egen kode til at løse et data science problem i en eksisterende kodebase, opbygge en test suite til denne kode og få erfaring med software engineering praksis i den virkelige verden.
- *Advanced Data Visualization* (7,5 ECTS): hvor den studerende bygger og skræddersyer visualiseringer til faglig kommunikation og projekter i data science.

Akse 3:

- indeholder et *specialeforberedende forskningsprojekt* 15 ECTS og et 30 ECTS *Kandidatspeciale*.

Strukturen er vist grafisk i bilag *Programme Design* (s.3) i *Dokumentation af efterspørgsel på uddannelsesprofil*. Alle undervisningsmaterialer er på engelsk; der findes ikke tilstrækkelig avancerede materialer på dansk. Der benyttes flere forskellige programmeringssprog, dog fortrinsvis Python, R og Java.

Den studerende kan vælge 15 ECTS *valgfri kurser* på andet semester fra en vifte af specialiserede kurser. Nogle eksempler på sådanne kurser er Operations Research, Computer Systems Performance, Practical Concurrent and Parallel Programming og Security (anvendt informationssikkerhed).

Der vil være et variabelt antal *videregående valgfri kurser*. Studerende skal vælge 15 ECTS kurser, herunder: Algoritmer (Algorithms), Datahåndtering (Data Systems), Maskinlæring (Machine Learning), Sprogteknologi og Deep Learning (Natural Language Processing and Deep Learning), Robotteknologi (Robotics) og Netværksteori (Network Science).

I det 7,5 ECTS *tværfaglige projekt* på første semester skal den studerende samarbejde med studerende fra andre uddannelser og bidrage med de særlige forkundskaber og kompetencer, som en kandidatstuderende i Data Science kan tilføre projektet, for eksempel konstruktion af en maskinlæringsprototype eller udkast til et systemdesign.

Det specialeforberedende 15 ECTS *forskningsprojekt* på tredje semester skal sikre, at den studerende får det bedst mulige udbytte af det efterfølgende kandidatspeciale. I projektet skal den studerende derfor definere og afgrænse emnet for kandidatspecialet, finde en specialevejleder og en eventuel virksomhedspartner. Den studerende vil typisk udforske problemstillingen gennem et litteraturstudie, konstruktion af en prototype, indsamling af data eller lignende. Desuden indeholder dette kursus en obligatorisk aktivitet, hvor de studerende skal udarbejde et reflekterende review report om de seneste diskussioner om etik inden for datalogi, maskinlæring og kunstig intelligens. Der tilstræbes en synergi med det videregående valgfri kursus, hvis anden del den studerende følger sideløbende. Det specialeforberedende projekt dokumenteres med en skriftlig rapport og en mundtlig eksamen.

*Kandidatspecialet* 30 ECTS på fjerde semester skal sikre væsentlig progression i forhold til de forudgående faglige elementer og skal dokumentere, at den studerende selvstændigt, men under forskningsbaseret vejledning, kan beskrive og analysere en kompleks faglig problemstilling inden for data science og foreslå, implementere og vurdere løsninger til denne under anvendelse af videnskabelige teorier og metoder.

### Begrundet forslag til taxameterindplacering

Da uddannelsen kernefaglighed hører under det naturvidenskabelige område, foreslår ITU, at uddannelsen indplaceres på takst 3 i lighed med øvrige godkendte kandidatuddannelser i datavidenskab.

### Forslag til censorkorps

Censorkorps for datalogi

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

Dokumentation af efterspørgsel på uddannelsesprofil.pdf

### Kort redegørelse for det nationale og regionale behov for den nye uddannelse

Der er stor uopfyldt efterspørgsel efter højt kvalificerede data science medarbejdere såvel nationalt som regionalt, og efterspørgslen forventes desuden at stige betydeligt. Det er især kandidater med dybe tekniske kompetencer inden for data science, der er meget efterspurgt. Se punktet *Underbygget skøn over det nationale og regionale behov for dimittender* for citater fra analyser af disse forhold. Regeringens egen *Strategi for Danmarks digitale vækst* (Erhvervsministeriet 2018) fremhæver virksomheders analyse og anvendelse af egne og offentlige data som et strategisk område med stort potentiale, men også som et område hvor danske virksomheder halter bagefter virksomheder i nabolandene. Som nævnt i samme punkt som ovenstående har danske virksomheder svært ved at rekruttere ansatte med de nødvendige dybdegående tekniske kompetencer inden for data science.

Der er en række grunde til at den nye uddannelse er engelsksproget:

For det første er arbejds sproget i it-virksomheder meget ofte engelsk, for eksempel for at lette samarbejdet med outsourcete teams i udlandet. Det engelsksprogede behov understøttes af en analyse af danske jobannoncer (Plank 2018)(se *Analysis of the Danish Data Science Job Market* i *Dokumentation af efterspørgsel på uddannelsesprofil*) inden for data science. Analysen, foretaget over 2,5 måned september-november 2018, fandt 1.390 data science orienterede jobopslag på landsplan, hvoraf 66,3% var på engelsk. Det betyder at selv dansktalende dimittender skal kunne begå sig fagligt på engelsk. Selv på offentlige arbejdspladser, hvor arbejds sproget ofte er dansk, opfattes data science som en engelsksproget disciplin. Lars Kjeldgaard (Senior Data Scientist, Skatteforvaltningen) skriver "*We also support an English-language program. Data Science is an international discipline by nature, and all of the literature on Data Science is in English.*" (mail 21. december 2018). Det stemmer også fint med IT-Branchens IT-Barometer 2018, der indeholder citater såsom "*4 ud af 5 virksomheder mener, at udenlandsk arbejdskraft er vigtig eller meget vigtig*", "*33,8% har allerede udenlandsk arbejdskraft og 15,8% forventer at få det*" (IT Branchen 2018).

For det andet vil en engelsksproget kandidatuddannelse tiltrække studerende, der både er mere kvalificerede og med højere diversitet, også kønsmæssigt, til fordel for studiet, fagområdet og de efterfølgende arbejdsgivere. I den forbindelse kan vi sammenligne med en eksisterende teknisk MSc uddannelse, nemlig ITU's MSc Software Development, der er engelsksproget. I perioden 2009-2014 var knap halvdelen af uddannelsens kandidater udenlandske. En analyse viser at flertallet af de udenlandske kandidater efter uddannelsen bor og arbejder i Danmark. (De øvrige muligheder er: bor i udlandet; er uden for arbejdsmarkedet; søger arbejde). Mere præcist er 63,4% af de udenlandske kandidater 1-2 år efter dimission bosatte og beskæftigede i Danmark. Med andre ord bliver 2 ud af 3 af vores udenlandske kandidater i Danmark, og ikke blot er de interesserede i at arbejde her, de arbejder faktisk. Dette viser at vores optag af udenlandske studerende i meget høj grad kommer danske virksomheder til gode inden for dette felt hvor der er mangel på kvalificeret arbejdskraft. Se punktet *Underbygget skøn over det nationale og regionale behov for dimittender*. Ganske vist er beskæftigelsestallet endnu højere for danske kandidater, idet 85,6% af disse bor og arbejder i Danmark. Men eftersom udlandet har båret udgifterne til de udenlandske studerendes opvækst, folkeskole, gymnasium og i de fleste tilfælde også bacheloruddannelse, udgør de udenlandske kandidater en betragtelig gevinst for den danske samfundsøkonomi og for statens finanser, jf. beregninger baseret på den samfundsøkonomiske model DREAM (2012). (Beskæftigelsestallene ovenfor stammer fra Uddannelses- og Forskningsministeriets datavarehus, baseret på Danmarks Statistik.) Beskæftigelsesgraden er beregnet af ITU's Analyseenhed på samme måde som i ministeriets "Reduktion i tilgangen til engelsksprogede uddannelser", 22. august 2018, der satte grænsen ved 40%. (Se *Beskæftigelsesgrad K-SDT dim 2009-2014* i *Dokumentation af eftersøgning på uddannelsesprofil*).

For det tredje er der mangel på danske forskere i data science. De videnskabelige medarbejdere på ITU's institut for datalogi kommer pt. fra godt 15 forskellige lande, og vi vil fortsat være nødt til at rekruttere forskere og undervisere fra udlandet for at kunne levere forskningsbaseret uddannelse af høj kvalitet.

Vi er opmærksomme på, at der i Danmark findes videregående uddannelser, der uddanner relativt mange udenlandske studerende, der efterfølgende ikke finder beskæftigelse i Danmark, og ønsker at understrege, at det modsatte gælder for ITU's uddannelser i softwareudvikling. Desuden forventes optaget af udenlandske studerende at være begrænset, da ITU's eksisterende bacheloruddannelse i Data Science, hvorfra hovedparten af kandidatuddannelsens studerende vil komme, fra sommer 2019 får Dansk A som optagelseskrav. At uddannelsen undervises på engelsk vil også i høj grad komme de dansktalende studerende til gode, da data science grundlæggende er en international disciplin, hvor arbejdssproget er engelsk. Medlem af aftagergruppen Kaare Brandt Petersen, Advanced Data Analysis - Implement udtaler: "Når jeg ansætter medarbejdere er det afgørende at de kan læse, tale og skrive engelsk på et ubesværet niveau – også inden for fagområderne. Derfor er det kritisk for de studerendes arbejdskompetencer, at lære at begå sig fagligt på engelsk og det er meget væsentligt, at de studerende fra et tidligt tidspunkt lærer at bruge engelsk som et dagligdags arbejdsprog med kolleger / medstuderende" (mail 10. december 2018).

### **Underbygget skøn over det nationale og regionale behov for dimittender**

Vi forventer, at der fra den nye kandidatuddannelse i Data Science vil dimittere knapt 40 kandidater i 2022, stigende til knapt 50 kandidater i 2024, hvoraf de fleste forventes at komme fra ITU's egen BSc uddannelse i Data Science. Se evt. punktet *Rekrutteringsgrundlag og videreuddannelsesmuligheder* senere i denne ansøgning.

Som dokumenteret nedenfor er der i Danmark en betydelig efterspørgsel efter højtuddannede it-medarbejdere, og heriblandt er det især medarbejdere med dybdegående teknisk viden indenfor data science der er eftertragtede. Den umiddelbare efterspørgsel efter netop data science specialister er størrelsesordenen større end hvad ITU forventer at kunne uddanne. På længere sigt forventes efterspørgslen desuden at stige betydeligt, og dimittenderne fra den nye uddannelse kan forvente meget lav arbejdsløshed.

Rambøll Management har for ITU udarbejdet en markedsundersøgelse om behovet for kandidater i Data Science. Baseret på en gennemgang af et antal eksisterende behovsanalyser konkluderer Rambølls markedsundersøgelse blandt andet, at *"Der er stor og stigende efterspørgsel på data science-kompetencer [...] I Danmark vurderes det udækkede efterspørgselsbehov i 2020 at være på op mod 6.000 ITEK-kandidater, hvoraf 3.000 af disse skyldes fremdrift inden for særligt Big Data."*, hvilket netop er det behov uddannelsen i Data Science vil hjælpe med til at dække (Rambøll 2018 s. 5). Rapporten fremhæver også manglen allerede i dag ved at skrive *"Efterspørgslen efter data scientists er så stor, at flere danske virksomheder har haft forgæves rekrutteringsforsøg."* Med et længere sigte mod 2030 beskriver rapporten, at der forventes udækket efterspørgsel på IT-specialister med lange videregående uddannelser på 13.000. Med henvisning til The World Economic Forum forventes jobs inden for kunstig intelligens-, Big Data- og machine learning etc. at gå fra at udgøre 17% i 2018 til 33% af IKT-arbejdsstyrken i 2022 (Rambøll 2018 s. 4). Dermed forventes en betydelig del af efterspørgslen at være efter kandidater med data science kompetencer. Ved at sammenligne med beslægtede uddannelser, konkluderer rapporten at der forventes lav ledighed blandt dimittenderne fra uddannelsen. Blandt uddannelser, der indeholder elementer af data science, er der en ledighedsfrekvens på mellem 2 og 7 pct., målt tre kvartaler efter endt uddannelse. Til sammenligning er landsgennemsnittet på 18 pct (Rambøll 2018 s. 5). Udover de eksisterende uddannelser vil de to godkendte, men endnu ikke udbudte data science kandidatuddannelser i Jylland, hhv Århus og Aalborg universitet, bidrage med tilsammen ca 100 kandidater fra 2023. Hvis man sammenholder den lave ledighed i de eksisterende uddannelser med redegørelsen for, at der frem mod 2020 vil være brug for 6.000 ITEK kandidater, heraf 3.000 indenfor big data, kan man konkludere, at der, selv med de kommende data science kandidater fra Århus og Aalborg, vil være et stort udækket merbehov for kandidater, som ITU's nye uddannelse i data science kan være med til at dække med 40-50 kandidater om året.

En analyse lavet af ITU af danske jobannoncer (Plank 2018) inden for data science, foretaget over 2,5 måned september-november 2018, fandt 1.390 data science orienterede jobopslag på landsplan, hvilket er betydeligt større end det forventede antal dimittender fra den foreslåede uddannelse. 53% af annoncerne var til stillinger i hovedstadsområdet, hvor ITU og flertallet af landets større softwarevirksomheder er placeret (Plank 2018). De aftagervirksomheder, som vi har talt med i forbindelse med udformning af uddannelsen, er derfor primært fra hovedstadsområdet. Se *Hvilke aftagere har været inddraget i behovsundersøgelsen?*

IT-Branchen laver hvert år en spørgeskemaundersøgelse blandt medlemmerne med statistik det danske IT jobmarked. I årets undersøgelse (IT barometer for 2018) kan man læse *"Næsten 2 ud af 3 it-virksomheder (63,2%) ser manglen på de rette it-kompetencer, som den største barriere for vækst i branchen", og at "39,5% [af virksomhederne] vil arbejde med Big Data"* i det kommende år.

**Hvilke aftagere har været inddraget i behovsundersøgelsen?**

Nærværende punkt beskriver blot, hvilke aftagere der har været inddraget. Punktet *Hvordan er det konkret sikret, at den nye uddannelse matcher det påviste behov?* nedenfor redegør i detaljer for, hvordan dette er foregået, og hvilke ændringer aftagernes kommentarer har affødt.

ITUs Executive Employers Panel, der bl.a. rådgiver ITU om den samlede uddannelsesportefølje, blev på panelets ordinære møde den 23. oktober 2018 præsenteret for og diskuterede planerne om en kandidatuddannelse i data science på ITU. Følgende panelmedlemmer deltog i mødet:

- Kaare Danielsen, Direktør og ejer, Jobindex
- Jan Peter Larsen, Udviklingsdirektør, Bankernes EDB Central (BEC)
- Carsten Gomard (panelets formand), Bestyrelsesformand, Netcompany
- Birgitte Hass, Direktør, IT-Branchen Pernille Geneser, CIO, Bestseller
- Lars Green Lauridsen, Senior Vice President, COWI
- Henrik T. Krøyer, Head of Mainframe, Danske Bank Charlotte Mark, Managing Director, Microsoft Development Center Copenhagen

I forbindelse med udviklingen af forslag til en kandidatuddannelse i Data Science ønskede ITU at inddrage aftagerne så tidligt som muligt i processen. For at sikre inddragelse af en repræsentativ aftagergruppe i udviklingsprocessen indvilgede Bodil Biering (Head of Development at Blackwood Seven) i at være formand for aftagergruppen og sammen med ITU at sammensætte en gruppe af aftagere, der repræsenterer forskellige brancher, såvel offentlige som private.

Der blev afholdt aftagermøde den 24. oktober 2018 på ITU med deltagelse af følgende aftagere:

- Bodil Biering (Chairman) (Blackwood Seven, Head of Development),
- Kaare Brandt Petersen (Implement Consulting Group, Advanced Data Analysis),
- Søren Ilsøe (Deloitte, Director, Analytics & Information Management),
- Jakob Elming (Doolittle Translation/Danish Defense, Data Scientist),
- Kaare Danielsen (Jobindex, stifter og direktør),
- Camilla Kerlaug (Spenderlog, CEO and founder),
- Jens Lund (former Nordea, Head of Model Monitoring and Performance, Risk Models),
- Lars Kjeldgaard (Udviklings og forenklingsstyrelsen, SKAT, Senior Data Scientist),
- Florian Laws (Tradeshift, Software Engineer - Machine Learning).

Endvidere har følgende aftagere gennem kvalitative interviews om behovet for en kandidatuddannelse i data science bidraget til rapporten *Markedsundersøgelse Kandidatuddannelse i Data Science* udført af Rambøll Management Consulting for ITU i efteråret 2018:

- Rikke Zeeberg Digitaliseringsstyrelsen, Direktør
- Søren Reeberg Nielsen Styrelsen for Dataforsyning og Effektivisering, Vicedirektør
- Peter Tibert Stoltze Kontorchef for Metode og Analyse Danmarks Statistik
- Karen Marie Lyng Afdelingschef for Datakvalitet og Indhold Sundhedsdatastyrelsen
- Birgitte Hass IT-Brancheforeningen, Direktør
- Torben Noer IBM Global Data Science and Business Analytics, Technical Sales Leader
- André Stelsig Rogaczewski Netcompany A/S, Adm. direktør
- Simon H. Galsgaard DSV Road Danmark, Executive Vice President
- Stig Geer Pedersen Topdanmark, Seniorprojektleder for Machine Learning og RPA
- Børn Lunding Sandqvist Tryg, Direktør for Pricing & Modelling

Den 8. januar 2019 var forslaget om ny uddannelse i høring hos ITU's aftagerpanel for uddannelserne under Department of Computer Science. Følgende aftagere deltog i mødet:

- Karsten Stanek Pedersen (Chairman), Director, KSP Consulting
- Casper Hovard, Senior Engagement Manager, KSP Nordic
- Mille Østerlund, Chef for civil rådgivning, CFCS - Center for Cybersikkerhed
- Niels Hallenberg, Vice President, Technical Foundation, Simcorp
- Morten Zohnesen, Managing Architect, Netcompany
- Thomas Hartmann, Head of Enterprise Architecture, ATP
- Bodil Biering, Head of Development, Blackwood Seven
- Mikkel Muhldorff Sigurd, Optimization Manager, Maersk Line

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

I forbindelse med udviklingen af forslag til en kandidatuddannelse i Data Science ønskede ITU at inddrage aftagerne så tidligt som muligt i processen. Herunder følger først en kronologisk oversigt over aftagerinvolvering og derefter en uddybning af involveringen, gruppe for gruppe.

**Kronologisk oversigt over aftagerinvolvering**

13. september 2018: Udvælgelse af data science-specifikt aftagergruppe samt planlægning af møde i samarbejde med Bodil Biering, formand for gruppen.
23. oktober 2018: Møde i ITU's Executive Employers' Panel. Præsentation og høring af kandidatuddannelsens *Programme Design*.
24. oktober 2018: Halvdagsmøde med data science-specifikt aftagergruppe med diskussioner og input til *Programme design*.
25. oktober - 1. december 2018: ITU reviderer *Programme Design* på baggrund af input fra aftagere.
1. december - 13. december 2018: Rambøll interviewer 10 relevante aftagere om behov og efterspurgte kompetencer, i forbindelse med udarbejdelsen af markedsundersøgelsen.
10. december 2018: Revideret *Programme Design* sendes til data science-specifik aftagergruppe til skriftlige kommentarer.
8. januar 2019: Aftagerpanelet for uddannelserne under Department of Computer Science giver på deres ordinære møde input og kommentarer til *Programme Design* for uddannelsen.

### Data science specifik aftagergruppe

For at sikre inddragelse af en repræsentativ gruppe af relevante aftagere i udviklingsprocessen indvilgede Bodil Biering (Head of Development at Blackwood Seven og medlem af ITU's ordinære aftagerpanel for uddannelserne under Department of Computer Science) i september 2018 i, sammen med ITU, at sammensætte en aftagergruppe og samtidig være formand for gruppen. ITU mødtes med Bodil Biering den 13. september 2018 og besluttede, hvilke aftagere der skulle inviteres til aftagergruppen. Aftagerne udvalgte med henblik på at dække markedet for de kommende kandidater så bredt som muligt. Således inviteredes både offentlige og private virksomheder fra forskellige brancher (se deltagere i *Hvilke aftagere har været inddraget i behovsundersøgelsen*). Endvidere planlagdes involveringen af aftagerne på dette møde.

Den 24. oktober 2018 afholdt gruppen et halvdagsmøde med forudgående skriftligt oplæg fra ITU. Se dagsorden og referat fra mødet i bilag *Minutes Employers Meeting 24 October 2018* i *Dokumentation af efterspørgsel på uddannelsesprofil*.

Sammen med dagsordenen sendte ITU udkast til *Programme Design* indeholdende uddannelsens:

- overordnede struktur
- overordnede mål og mål for læringsudbytte
- obligatoriske kursers omfang og overordnede indhold
- ansøgermålgruppe og adgangskrav
- jobprofil

*Programme Design*-dokumentet er blevet brugt i hele udviklingsprocessen af uddannelsen og findes derfor i flere versioner. Den uploadede version er fra december 2018 og er derfor en version, der netop afspejler input fra aftagerside i udviklingen af uddannelsen.

På mødet gennemførtes en høring af uddannelsens design og faglige indhold på baggrund af ovenstående.

Helt generelt bifaldt aftagerne det overordnede design: "*The panel finds that it is very good that the program ensures that the theoretical skills are also combined with practical application. The applied approach is very useful for employers and companies when hiring graduates and the "Profound and Applied" should be the ITU MSc in Data Science profile in relation to other Universities MSc Data Science programs.*"

ITU bad aftagerne om feedback og, som det fremgår af referatet, (se bilag *Minutes Employers Meeting 24 October 2018*) havde aftagerne specifikke kommentarer til følgende aspekter af *Programme Design*:

i) Tekniske færdigheder og kompetencer, ii) afgrænsninger af jobprofil, iii) forretningskompetencer (hvad betyder business intelligence), og iv) etik og privacy.

Tilbage meldingen fra mødet blev brugt til at revidere uddannelsens *Programme Design*. Sammenfatning af input fra aftagergruppen (yderligere detaljer i bilag *Minutes Employers Meeting 24 October 2018*) og revisioner af *Programme Design* var:

i) Aftagerne fandt at: "*it is very good that the program ensures that the theoretical skills are also combined with practical application*" (se *Minutes from employers meeting 24 October 2018* i *Dokumentation af efterspørgsel på uddannelsesprofil*).

Aftagerne fandt også færdigheder inden for softwareudvikling vigtige. For at imødekomme sidstnævnte blev der tilføjet et kursus *Data Science in Production*, et projektbaseret kursus, hvor de studerende får erfaring med software engineering i praksis.

ii) Aftagerne konstaterede, at der i forbindelse med jobprofilerne også er mulighed for, at en data scientist bevæger sig mellem profil A og B (se *Programme Design* s. 2). De to afgrænsede jobprofiler udgør ikke en streng dikotomi. ITU har derfor opdateret jobprofilen, så det afspejler sig i den.

iii) Aftagerne fandt at det foreslåede kursus *Business Intelligence* var lidt for arkaisk og foreslog et mere relevant indhold. Det er vigtigt, at kandidater har en grundlæggende forretnings- og organisationsforståelse, så de kan fungere som "oversættere" imellem de forskellige interessenter, men det må ikke være på bekostning af det tekniske indhold. Et nyt kursus *Data Science in Production*, der imødekommer ovenstående behov erstatter *Business Intelligence* kurset.

iv) Etik: ITU spurgte aftagerne, hvordan etiske problemstillinger skal indgå i uddannelsen. Aftagerne fandt det vigtigt at etik indgår som del af uddannelsen, men havde ikke nogen endelig mening om det skulle være i et selvstændigt kursus eller som del af andre kurser. De opfordrede ITU til at inddrage etiske problemstillinger, når det er relevant, måske især i forbindelse med kurser, der bruger store datasæt. Uddannelsens tredje semester indeholder et 15 ECTS projekt med en obligatorisk aktivitet, hvor de studerende skal udarbejde en reflekterende rapport om de seneste diskussioner om etik inden for data science. Desuden gjorde ITU opmærksom på at etik allerede indgår på bacheloruddannelsen i kurset Security and Privacy og Second Year Project.

Efter ITU havde revideret *Programme Design* blev det i starten af december 2018 sendt til aftagergruppen til orientering og yderligere kommentarer.

Som det ses af svar fra medlem Jens Lund (konsulent, tidligere Nordea - Executive Director, Head of Analytics and Business Design) finder han, at ITU har formået at indarbejde aftagernes behov: "*Overall, I find that you have managed to find a very good balance in the proposed design of a MSc in Data Science education*" og at ITU med uddannelsen kan være med til at sikre at Danmarks indflydelse på data science området og ikke overlade det til store tech-giganter eller lande som Kina: "*In my opinion, it is a field that will be growing exponentially over the coming years, and it will be of importance for Denmark as a society that we become part of this development in order to not be completely overtaken by big tech giants (Google, Apple, Amazon, Facebook, ...) or countries like China, that is spending enormous amounts of resources in this area and moves forward at a quick pace.*"

Endvidere finder han at uddannelsen matcher det efterspurgte behov for data science kandidater: "*I also find, that the proposed programme will position itself uniquely in the space of similar programmes at other universities in Denmark. It does so by having a very strong focus on the implementation/computer science and algorithmic aspects without compromising on the understanding of the underlying mathematical/statistical models. From a hiring perspective, I have always looked for candidates with a strong computer science background/experience, as it is something that is essential to make the theoretical models work in practice in production systems.*" (mail 13. december 2018)

Svaret fra aftager Lars Kjeldgaard (Senior Data Scientist, Skatteforvaltningen) understøtter også programdesignet og viser at aftagernes feedback er blevet indarbejdet i uddannelsesdesignet: "*Overall: The revised program looks great. In addition, the replacement of the "Business Intelligence" course with "Data Science in Production" is a big improvement. Production-level code competencies are key in modern Data Science, and it is one of our top priorities when we hire Data Scientists.*" (mail 21. december 2018).

Ingen af gruppens medlemmer kom med forslag til yderligere ændringer.

### **ITU Executive Level Employers panel**

Udover de uddannelsesspecifikke aftagerpaneler har ITU også et Executive Level panel, der rådgiver ITU om den samlede uddannelsesportefølje. Se *Hvilke aftagere har været inddraget i behovsundersøgelsen*. Den 23. oktober 2018 blev panelet præsenteret for og diskuterede planerne om en kandidatuddannelse i data science på ITU. Som forberedelse havde panelet modtaget *Programme Design*. Som det fremgår af referatet bifaldt panelet at uddannelsen "*har sin tyngde i tekniske/matematiske emner og bygger oven på bacheloruddannelsen i Data Science. Endvidere panelet fandt det også vigtigt, at de studerende stifter bekendtskab med etiske problemstillinger og lærer, hvad forretningsforståelse handler om.*" (se *Uddrag af referater fra aftagerpanelmøder i Dokumentation af efterspørgsel på uddannelsesprofil*).

Kommentarer fra Executive Level aftagerpanelets møde viser at de støtter til uddannelsens tekniske fokus og at de etiske problemstillinger indgår i 3. semester, således førte panelets input ikke til ændringer i uddannelsesdesignet.

### **Aftagerpanel for uddannelser under Department of Computer Science**

Endelig havde ITU's aftagerpanel for uddannelser under Department of Computer Science inden mødet den 8. januar 2019 modtaget forslaget til *Programme Design* for den nye uddannelse. På mødet blev uddannelsesdesignet præsenteret og panelet fandt at uddannelsens indhold og struktur i meget høj grad imødekommer et eksisterende behov. Aftagerne var meget glade for kurset "*Data Science in Production*". Kurset adresserer en meget hyppig hindring nyuddannede kandidater kan have med at anvende deres viden. Panelet diskuterede også etik og security, som begge er vigtige emner og er del af uddannelsen.

Drøftelserne med aftagerpanelet bekræftede det foreslåede *Programme Design* og førte ikke til yderligere ændringer.

### **Aftagerinterviews i forbindelse med Markedsundersøgelse Kandidatuddannelse i Data Science**

Ud over den ovenfor beskrevne aftagerinvolvering er der i forbindelse med udarbejdelsen af *Markedsundersøgelse Kandidatuddannelse i Data Science* udført af Rambøll Management Consulting for ITU i efteråret 2018 udført 10 kvalitative interviews om behovet for en kandidatuddannelse i data science og hvilke kompetencer der efterspørges (Se Rambøll 2018 i *Dokumentation af efterspørgsel på uddannelsesprofil*). Rapportens konklusioner og anbefalinger om behov for data science kompetencer stemmer overens med den ansøgte uddannelses udformning: "*Dybe tekniske kompetencer efterspørges mere end tværgående kompetencer.*" (Rambøll 2018 s. 4)

Flere af de interviewede aftagere mangler allerede nu medarbejdere med data science kvalifikationer: "*Flere af de adspurgte aftagerorganisationer og -virksomheder oplever dermed allerede en udækket efterspørgsel på data scientists. På baggrund af undersøgelsens desk study og kvalitative interviews kan det konkluderes, at der er klare tegn på, at der allerede inden for nærmeste fremtid vil være et stort udækket behov for højtuddannede IKT-kandidater, som netop besidder nogle af de kompetencer, som ITU's potentielle kandidatuddannelse i Data Science forventes at kunne imødekomme.*" (Rambøll 2018 s. 7).

### **Konklusion**

Samlet set har aftagerhøringerne været særdeles værdifulde og ført til en række justeringer i uddannelsesforslaget. Vi er derfor overbeviste om, at der er overensstemmelse mellem uddannelsens indhold og aftagernes behov, hvilket bekræftes af Bodil Biering, formand for aftagergruppen i forbindelse med den supplerende høring af aftagergruppen i december 2018: "*it looks good, all the main points from the panel have been integrated: production code and systems, ethics, security.*" (mail 5. december 2018).

### **Beskriv ligheder og forskelle til beslægtede uddannelser, herunder beskæftigelse og eventual dimensionering.**

På baggrund af ministeriets ønske om at universiteterne koordinerer og diskuterer udbuddet af uddannelser inden for data science, informerede ITU 16. november 2018 på et nationalt institutledermøde for datalogi og it-uddannelser de øvrige universiteter om, at ITU planlægger at ansøge om en kandidatuddannelse i data science. Det blev også forklaret at kandidatuddannelsen bygger oven på ITU's eksisterende bacheloruddannelse i data science, og således har et andet (teknisk dybere, mindre forretningsorienteret) sigte end en række af de eksisterende eller foreslåede kandidatuddannelser i data science (på CBS, AU BSS, DTU Management, KU Samfundsvidenskab), og en anden målgruppe end kandidatuddannelserne i datalogi på AU, AAU, KU, ITU og SDU. Den ligner snarere de i 2017 prækvalificerede men endnu ikke udbudte kandidatuddannelser i data science ved AAU og AU.

Den ansøgte kandidatuddannelse bygger på kernekompetencer erhvervet i en forudgående bacheloruddannelse i Data Science eller lignende: lineær algebra, programmering, klassiske algoritmer og datastrukturer, machine learning, statistics, data management og graph theory, beskrevet i punktet *Hvilke adgangskrav gælder til uddannelsen?* Oven på disse kernekompetencer bygger den foreslåede uddannelse en række centrale data science fag (fx algorithm design, advanced applied statistics and calculus, processing of heterogeneous data, data science in production, and advanced visualization). Endvidere tilbyder uddannelsen en række videregående fag og forskningsemner såsom advanced machine learning, natural language processing and deep learning, advanced algorithms, data systems design, robotics osv., som afspejler forskningen på ITU. Samlet giver disse forhold den nye uddannelse et fokus på stærke tekniske kompetencer med dybt teoretisk fundament. Dette støttes af en repræsentativ gruppe af aftagere, som vi har involveret i designet af uddannelsen. Se punktet *Hvordan er det konkret sikret, at den nye uddannelse matcher det påviste behov?* for analyser af disse forhold.

En rapport udarbejdet af Rambøll Management sammenligner den foreslåede nye uddannelse med eksisterende og nye kandidatuddannelser i datavidenskab og konkluderer blandt andet: "**Få danske uddannelser giver en lignende kombination af data science-kompetencer.** Kortlægningen af de eksisterende danske kandidatuddannelser inden for data science har resulteret i, at der i dag blot er 11 kandidatuddannelser, der kan anses som indeholdende elementer, der overlapper med ITU's egen potentielle kandidatuddannelse i Data Science. En nærmere analyse af de identificerede beslægtede uddannelser tegner et billede af, at det kun i mindre grad er muligt at opnå den efterspurgte kombination af kompetencer, som der tillægges en data scientist i interviews. Typisk har uddannelserne et begrænset antal af relevante data science-fag, som oftest kun kan gennemføres ved, at den studerende vælger at specialisere sig inden for et bestemt område. Det er dermed kendetegnene for de eksisterende beslægtede kandidatuddannelser inden for data science, at de får karakter af en specialiseringsvej fremfor en hovedvej, der også indbefatter tilegnelsen af et bredere sæt kompetencer hos de studerende, uden at gå på kompromis med den teknologiske viden og færdigheder." (Rambøll 2018, s. 4)

De gennemførte analyser af danske jobannoncer (se punktet *Underbygget skøn over det nationale og regionale behov for dimittender*) viser at der er en større efterspørgsel efter kandidater med data science kompetencer i hovedstadsområdet end i resten af landet. En analyse af eksisterende og prækvalificerede programmer på dette område viser, at der er to kandidatuddannelser i hovedstadsområdet, som indeholder elementer af data science. Kandidatuddannelsen på CBS har et overordnet fokus på forretningsaspektet (Rambøll 2018, s. 13) og Københavns Universitet's kandidatuddannelse i Social datavidenskab har fokus på at kombinere samfundsvidenskab og datavidenskab (Rambøll 2018, s. 14). For yderligere detaljer om ligheder og forskelle til beslægtede uddannelser se Rambøll s. 12-15.

Beskæftigelsesudsigterne for kandidater i Data Science er fremragende, som angivet i rapporten udarbejdet af Rambøll Management: "**Lav ledighed for beslægtede uddannelser.** For de kandidatuddannelser, som er inkluderet i denne udbudsafdækning, er det meget bemærkelsesværdigt, at samtlige uddannelser har en lav dimittendledighed på mellem to og syv pct. Dette skal holdes op mod det overordnede gennemsnit på tværs af alle kandidatuddannelser i Danmark, som er på ca. 18 pct. Tallene er for dimittender i 2016 og er målt tre kvartaler efter fuldførelse, hvilket er de seneste opgjorte tal for Uddannelses- og Forskningsministeriets aktuelle ledighed blandt dimittender fra videregående uddannelser. En så lav dimittendledighed for samtlige uddannelser i udbudsafdækningen tyder på et stort behov for de kompetencer, der er relateret til data science." (Rambøll 2018, s. 15). Selv når man medtager kandidater fra de to godkendte, men endnu ikke udbudte, kandidatuddannelser i data science på hhv. Aalborg (55-60 kandidater fra 2023) og Århus universitet (30-40 kandidater fra 2023 efterfølgende stigende til 50-65) sammen med ITU's forventede 50 kandidater pr. år er det ikke nok til at dække merbehovet på 6.000 ITEK kandidater, heraf 3.000 indenfor big data, frem mod 2020. Se *Underbygget skøn over det nationale og regionale behov for dimittender*. Endvidere skal det påpeges at ITU's analyse af danske jobannoncer viser at 53% af annoncerne var til stillinger i hovedstadsområdet, hvor ITU og flertallet af landets større softwarevirksomheder er placeret (Plank 2018).

Kandidater i Data Science, der er interesserede i forskning, kan fortsætte i phd-studier. Langt de fleste tager dog arbejde i it-industrien eller i virksomheder, der i væsentlig grad anvender data science.

### Rekrutteringsgrundlag og videreuddannelsesmuligheder

Da optagelseskravet til den nye kandidatuddannelse er en bacheloruddannelse i Data Science eller lignende med opfyldelse af 80% af de specifikke faglige krav (se punktet *Hvilke adgangskrav gælder til uddannelsen?*) forventer vi at kandidatuddannelsen hovedsageligt vil rekruttere fra a) ITU's egen bacheloruddannelse i Data Science, og b) fra 2022 et mindre antal ansøgere fra andre danske universiteter: *Bachelor i Data Science* fra Aalborg Universitet og Århus Universitet, *bachelor i Kunstig intelligens og data* fra DTU samt *bachelor i machine learning og data science* fra Københavns Universitet. c) Endelig forventer vi et mindre antal internationale ansøgere, da der endnu ikke findes mange internationale bacheloruddannelser i Data Science i Europa.

Da rekrutteringsgrundlaget hovedsageligt vil bestå af ITU's egne bachelorer, forventer vi ikke at det vil påvirke de øvrige universiteters kandidatuddannelsers rekrutteringsgrundlag i nævneværdig grad. Men i lighed med andre kandidatuddannelser, som f.eks. i datalogi, vil der være en mindre del af de studerende fra de forskellige danske universiteter, der vælger at læse deres kandidatuddannelse på et andet universitet, og at der derfor snarere er tale om en udveksling af studerende.

### Forventet optag på de første 3 år af uddannelsen

Ansøgere vil hovedsageligt komme fra ITU's egen bacheloruddannelse i data science.

For årene 2020, 2021 og 2022 forventes et optag på 47, 50 og 57 studerende (efter tidligt frafald).

Den stigning der er i optagelsestallene for kandidatuddannelsen fra 2020 til 2022 vil komme fra ITU's egen bacheloruddannelse og vi forventer ikke at ansøgerantallet fra de øvrige danske bacheloruddannelser og internationale bacheloruddannelser vil vokse.

Med baggrund i det store udækkede behov for kandidater, der er beskrevet tidligere i ansøgningen under (titel på punkt) og den store ansøgerinteresse (se nedenfor) planlægger ITU frem mod 2020 at øge optaget til bacheloruddannelsen i data science til 100, hvilket også vil give en stigning i optaget på kandidatuddannelsen.

Interessen for ITU's bacheloruddannelse i data science har fra starten været stor, og fra 2017 til 2018 er det samlede ansøgerantal steget med 14%, og førsteprioritetsansøgningerne er steget med 30%. Ansøbertallene var i 2017 391 heraf 142 med første prioritet og i 2018 446 heraf 185 med første prioritet og der er optaget 56 studerende i 2017 og 64 studerende i 2018 regnet efter tidligt frafald. Derudover er kvaliteten af ansøgere og steget, da adgangskvotienten for kvote 1 fra 2017 til 2018 steget fra 8,0 til 9,0.

**Hvis relevant: forventede praktikaftaler**

N/A

**Øvrige bemærkninger til ansøgningen**

N/A

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

Ja

**Status på ansøgningen**

Godkendt

**Ansøgningsrunde**

2019-1

**Afgørelsesbilag - Upload PDF-fil**

A6 - Godkendelse af KA i Datavidenskab - ITU.pdf

**Samlet godkendelsesbrev - Upload PDF-fil**



# KANDIDATUDDANNELSE I DATAVIDENSKAB

PRÆKVALIFIKATIONSANSØGNING

1. februar 2019



## Ansøgning om prækvalifikation af ny uddannelse

På vegne af IT-Universitetet i København fremsendes hermed ansøgning om oprettelse af ny kandidatuddannelse i datavidenskab.

Dette dokument indeholder:

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

### 1. Yderligere dokumentation for behov og relevans for kandidatuddannelsen

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

- **Markedsundersøgelse kandidat i Data Science (side 4)(Relevante sider 6-20 (1-15))**  
I efteråret 2018 gennemførte Rambøll Management for IT-Universitetet en markedsundersøgelse, der i to dele kortlagde 1. Analyse af efterspørgslen efter data scientists, som skal afdække arbejdsmarkedets behov gennem henholdsvis: a. Et systematisk desk study af volumen i markedet, b. 10 kvalitative interviews med centrale aftagerorganisationer og -virksomheder.  
2. Udbudsdekning af lignende uddannelser. såvel eksisterende kandidatuddannelser som godkendte og ansøgte ,em endnu ikke udbudte. Vurdering af plads til ny kandidatuddannelse.
- **Analysis of the Danish Data Science Job Market. (side 35)**  
IT-Universitetet har endvidere i en 2 ½ måneds periode i 2018 indsamlet relevante data science jobopslag og kan gennem analyse af de indsamlede data dokumentere et eksisterende behov for data science kandidater.
- **Beskæftigelsesgrad for engelske og danske dimittender fra IT-Universitetet i perioden 2009-2014 (side 38)**  
ITU har som sammenligningsgrundlag beregnet beskæftigelsesgrad for dimittender på baggrund af UFM's DVH baseret på data fra Danmarks Statistik.
- **Overgang BSc in Data Science til MSc in Data Science (side 42)**  
Undersøgelse blandt nuværende data science-bachelorere om ønsket kandidatuddannelsen.
- **Minutes employers meeting 24 October 2018 (side 43)**  
Referat af halvdagsmøde med repræsentativ data science aftagergruppe, som har været inddraget i udviklingen af uddannelsen.
- **Uddrag af referater fra aftagerpanelmøder (side 49)**  
Uddrag af referater fra ITU's Executive Employers' Panel 23. oktober 2018 og aftagerpanel for uddannelser under Computer Science Department 8. januar 2019, der understøtter ansøgning om ny uddannelse.

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

- **Programme Design, december 2018 (side 51)**  
Brugt i høring af aftagere og intern udvikling på IT-Universitetet
- **Referencer (side 58)**  
Henvisninger med links til de rapporter m.m., der er citeret fra i ansøgningen.

Til  
**IT-Universitetet**

Dokumenttype  
**Afrapportering**

Dato  
**December 2018**

# MARKEDSUNDERSØGELSE KANDIDATUDDANNELSE I DATA SCIENCE



Rambøll  
Olof Palmes Allé 20  
DK-8200 Aarhus N

T +45 5161 1000  
F +45 5161 1001  
<https://dk.ramboll.com>

Rambøll Management Consulting  
A/S  
CVR NR. 60997918

## INDHOLD

<b>Indledning</b>	<b>2</b>
Baggrund for undersøgelsen	2
Undersøgelhedsdesign	2
Læsevejledning	3
<b>Sammenfatning og primære resultater</b>	<b>4</b>
<b>Del I: Analyse af efterspørgslen efter Data scientists</b>	<b>6</b>
Der er stor og stigende efterspørgsel på data science-kompetencer	6
Efterspørgslen på data scientists stiger frem mod 2030	7
Tekniske kompetencer er mere efterspurgt end tværgående kompetencer	9
Fremover vil efterspørgslen på særligt tekniske kompetencer stige	10
Kandidater skal både have gode dansk- og engelskkundskaber	10
Praktikophold er værdifulde, men kan være på bekostning af specialiserede kompetencer	11
<b>Del II: Udbudsdækning af lignende uddannelser</b>	<b>12</b>
Redegørelse for og analyse af lignende uddannelser	12
Nye uddannelser på vej inden for Data Science	14
Få danske uddannelser giver den efterspurgte kombination af kompetencer	14
Lav ledighed for beslægtede uddannelser	15
<b>Bilag og metode</b>	<b>16</b>
BILAG 1: Desk study	17
BILAG 2: Overblik over interviews inkl. metode	21
BILAG 3: Liste over inkluderede uddannelser inkl. metode	22

## 1. INDLEDNING

I denne rapport præsenterer Rambøll Management Consulting (Rambøll) resultaterne af en markedsundersøgelse om uddannelse og arbejdsmarkedet for data science. Undersøgelsen er gennemført på opdrag af IT-Universitetet (ITU) i perioden oktober-december 2018.

Formålet med undersøgelsen er at kortlægge den eksisterende viden om udbuddet for kandidatuddannelser inden for data science samt at give overblik over den forventede udvikling af arbejdsmarkedet og efterspørgslen på data scientists i Danmark. Markedsundersøgelsen skal dermed kvalificere grundlaget for ansøgning om oprettelsen af en kandidatuddannelse i Data Science på IT-Universitetet.

### 1.1 Baggrund for undersøgelsen

ITU påtænker at oprette en kandidatuddannelse i Data Science. I 2017 oprettede ITU bacheloruddannelsen i Data Science, hvor Rambøll forud for oprettelsen foretog en markedsundersøgelse af behovet for en bacheloruddannelse i Data Science. Nærværende markedsundersøgelse bygger dermed oven på den tidligere markedsundersøgelse og ser på den nyeste litteratur om efterspørgslen på arbejdskraft inden for data science samt udbuddet af lignende kandidatuddannelser.

Baggrunden for, at det er relevant at belyse behovet for en kandidatuddannelse i Data Science, skal særligt ses i lyset af den teknologiske udvikling, hvor flere og større datamængder – såkaldte Big Data – bliver skabt og gjort tilgængelige. Samtidig efterspørger regeringen en bedre udnyttelse af de muligheder, som Big Data giver. Erhvervsministeriet har påpeget, at der er mangel på uddannelsesprogrammer, der fokuserer på spændet mellem dataanalyse og forretningsudvikling.

Ovenstående faktorer peger på, at der allerede er og fremadrettet i højere grad vil være behov for kandidater, som har kompetencer inden for data science. Dette er udgangspunktet for nærværende markedsundersøgelse.

### 1.2 Undersøgellesdesign

Markedsundersøgelsen består af to dele, som bidrager til at analysere aftagerbehovet for en kandidatuddannelse i Data Science samt udbuddet af lignende kandidatuddannelser. Konkret indeholder de to delundersøgelser følgende elementer:

- 1) En efterspørgselsafdækning, som skal afdække arbejdsmarkedets behov gennem henholdsvis:
  - a. Et systematisk desk study af volumen i markedet
  - b. 10 kvalitative interviews med centrale aftagerorganisationer og -virksomheder.
- 2) En udbudsafdækning, der skal sammenligne indholdet af ITU's forslag til en kandidatuddannelse i Data Science med indholdet på andre kandidatuddannelser.

Disse to delundersøgelser afdækkes i henholdsvis del I og del II.

*Del I* afdækker arbejdsmarkedets behov inden for data science. Konkret afdækkes det, hvor stor efterspørgslen på kandidater inden for data science er på nuværende tidspunkt, og hvordan efterspørgslen forventes at udvikle sig frem mod 2030. Desuden afdækkes det, hvilke konkrete kompetencer der efterspørges af kandidater i data science.

Undersøgelsens desk study består af 10 centrale studier – herunder både internationale og nationale. Antallet af inkluderede studier er relativt lille, hvilket primært skyldes, at data science er et forholdsvist nyt felt. Desuden er det prioriteret at inddrage nyere studier, som bidrager til at bygge ovenpå markedsundersøgelsen i 2016, hvilket også begrænser antallet af inkluderede studier. Undersøgelsens desk study bidrager dermed til at skabe overblik over den aktuelle og fremadrettede efterspørgsel på kandidater i Data Science.

De 10 kvalitative interviews med repræsentanter fra centrale aftagerorganisationer og -virksomheder bidrager særligt til at besvare, hvilke konkrete kompetencer som arbejdsmarkedet efterspørger hos kandidater i Data Science.

*Delopgave II* undersøger udbuddet af eksisterende og nye danske uddannelser, der indeholder elementer af data science. De identificerede uddannelsers indhold sammenholdes med det planlagte indhold på den kandidatuddannelse i Data Science, som ITU ønsker at udbyde. Kortlægningen af uddannelser resulterede i 11 inkluderede kandidatuddannelser, som indeholder tilstrækkelige elementer af data science.

I nedenstående figur ses en oversigt over markedsundersøgelsens samlede datagrundlag.

**Figur 1-1: Oversigt over markedsundersøgelsens samlede datagrundlag**



### 1.3 Læsevejledning

Udover dette første kapitel består rapporten af tre øvrige kapitler samt tre bilag.

#### Kapitel 2 – Sammenfatning og primære resultater

Her præsenteres en opsamling af undersøgelsens primære resultater på tværs af rapportens delopgaver.

#### Kapitel 3 – Delopgave I

Dette kapitel indeholder den samlede analyse af efterspørgslen på data scientists. Kapitlet belyser den aktuelle efterspørgsel på data scientists, forventninger til udviklingen i efterspørgslen samt hvilke konkrete kompetencer der efterspørges.

#### Kapitel 4 – Delopgave II

Dette kapitel indeholder en udbudsafdækning af eksisterende og nye uddannelser, som indeholder elementer af data science.

## 2. SAMMENFATNING OG PRIMÆRE RESULTATER

Nærværende undersøgelse afdækker markedets behov for en kandidatuddannelse i Data Science, som et element i IT-Universitetets akkrediteringsansøgning for oprettelsen heraf. Efterspørgsels behovet er af afdækket dels ved gennemgang af tidligere undersøgelser og dels ved 10 aftager interview i forskellige brancher og sektorer. Udbuddet er tilvejebragt ved gennemgang af øvrige uddannelser med tilstrækkelige elementer af Data Science.

IT-Universitet startede en bacheloruddannelse i Data Science i 2017. I forbindelse med akkrediteringsansøgningen blev der da gennemført en tilsvarende markedsundersøgelse. En del af denne afdækning vil således tage afsæt heri, og fremlægge opdatering, hvor det er relevant.

### **Efterspørgslen på data scientists er stor og forventes at stige**

På tværs af både studier og kvalitative interviews er der identificeret stor efterspørgsel på de kompetencer, som en data scientist besidder. Denne tendens er gældende både på internationalt plan og i Danmark. Efterspørgslen efter data scientists er så stor, at flere danske virksomheder har haft forgæves rekrutteringsforsøg, hvor den primære årsag har været manglende kompetencer blandt ansøgerne<sup>1</sup>. Efterspørgslen efter data scientists forventes at stige i alle sektorer frem mod 2030, og særligt IKT-branchen giver udtryk for en markant stor efterspørgsel til data scientist-jobs. Det er i 2016 estimeret, at der i 2030 vil være et udækket efterspørgselsbehov på 13.000 IKT-specialister<sup>2</sup> med en lang videregående uddannelse<sup>3</sup>. Der er ikke fundet nyere opgørelser over behovet for data scientists eller grundlag for at udfordre dette tal. Af undersøgelsens interviews fremgår det, at manglen på kvalificeret arbejdskraft gælder både i Danmark og på internationalt plan hvorfor kandidaterne forventes at ville indgå i den globale kamp om talenter og kompetencer.

### **Dybe tekniske kompetencer efterspørges mere end tværgående kompetencer**

Undersøgelsen af kompetenceefterspørgslen indikerer, at der fremadrettet særligt vil blive lagt vægt på tekniske og matematiske kompetencer inden for fx statistik og programmering. I flere interview bliver der også udtrykt ønske om tværfaglige kompetencer som fx forretningsforståelse og kommunikation, men det er understreget, at behovet for kontekstforståelsen styrker kandidaten. De tværfaglige kompetencer må dog ikke fjerne fokus fra de dybere tekniske kompetencer. Virksomheder efterspørger således især kandidater, som kan håndtere store og komplekse datasæt, og aktuelt efterspørges kandidater, som forstår og kan udvikle selvstående algoritmer og finde mønstre i data. Det er dermed særligt kompetencer inden for Big Data, artificial intelligence og machine learning, som virksomhederne efterspørger hos data scientists, suppleret med et minimum af branche-, sektor- og forretningsforståelse. På verdensplan forventes jobs inden for disse områder desuden at udgøre 33 pct. af jobs inden for IKT-branchen i 2022<sup>4</sup>.

### **Få danske uddannelser giver en lignende kombination af data science-kompetencer**

Kortlægningen af de eksisterende danske kandidatuddannelser inden for data science har resulteret i, at der i dag blot er 11 kandidatuddannelser, der kan anses som indeholdende elementer, der overlapper med ITU's egen potentielle kandidatuddannelse i Data Science. En nærmere analyse af de identificerede beslægtede uddannelser tegner et billede af, at det kun i mindre grad er muligt at opnå den efterspurgte kombination af kompetencer, som der tillægges en data scientist i interviews. Typisk har uddannelserne et begrænset antal af relevante data

<sup>1</sup> Andreas Højbjerg & Rasmus Mikkelsen: "Mismatchproblemer i Norden" (2017).

<sup>2</sup> 'IKT-specialister' omfatter specialister inden for informations- og kommunikationsteknologi, herunder data scientists.

<sup>3</sup> Højbjerg Brauer Schultz: "Virksomheders behov for digitale kompetencer" (2016)

<sup>4</sup> World Economic Forum: "The Future of Jobs Report" (2018).

science-fag, som oftest kun kan gennemføres ved, at den studerende vælger at specialisere sig inden for et bestemt område. Det er dermed kendetegnene for de eksisterende beslægtede kandidatuddannelser inden for data science, at de får karakter af en specialiseringsvej fremfor en hovedvej, der også indbefatter tilegnelsen af et bredere sæt kompetencer hos de studerende, uden at gå på kompromis med den teknologiske viden og færdigheder.

### **Lav ledighed blandt nye dimittender fra de beslægtede uddannelser**

De seneste opgjorte tal fra Uddannelses- og Forskningsministeriets ledighedstal i 2016 viser, at dimittender fra uddannelser, der indeholder elementer af data science, har en ledighedsfrekvens på mellem to pct. og syv pct., målt tre kvartaler efter endt uddannelse. Til sammenligning er det overordnede gennemsnit på tværs af alle kandidatuddannelser i Danmark på ca. 18 pct.<sup>5</sup> Dette underbygger behovet for de kompetencer, som en potentiel kandidatuddannelse i Data Science vil kunne tilføre.

<sup>5</sup> <https://ufm.dk/uddannelse/statistik-og-analyser/faerdiguddannede/aktuel-ledighed>

### 3. DEL I: ANALYSE AF EFTERSPØRGSLEN EFTER DATA SCIENTISTS

Formålet med del I er at skabe overblik over den nuværende og kommende efterspørgsel på kandidater i Data Science, herunder også hvilke konkrete kompetencer arbejdsmarkedet efterspørger. Del I er afdækket gennem et desk study og 10 kvalitative interviews med repræsentanter fra centrale potentielle aftagerorganisationer og -virksomheder af kandidater i Data Science.

Efterspørgselsafdækningen har ledt til følgende fund:

1. Der er stor mangel på højtuddannede kandidater med data science-kompetencer
2. Efterspørgslen på højtuddannede kandidater med data science-kompetencer stiger fremadrettet
3. Tekniske og matematiske kompetencer (statistik, matematik, programmering) vægtes højere end tværfaglige kompetencer (forretningsforståelse, kommunikation).

Da det ikke er muligt at identificere konkrete tal på efterspørgslen efter data scientists i Danmark, er der taget udgangspunkt i studier, som viser efterspørgslen på IKT<sup>6</sup>- og ITEK<sup>7</sup>-kompetencer. De kompetencer, som kandidater i Data Science vil besidde, er netop kompetencer, som hører under disse kategorier.

#### 3.1 Der er stor og stigende efterspørgsel på data science-kompetencer

I litteraturen – både internationalt og nationalt - tegnes et klart billede af, at der er et markant behov for højtuddannede IKT-specialister. Flere studier viser, at det særligt er data science-kompetencer, som er i vækst inden for IKT- og ITEK-branchen. I 2017 var der fx mere end 87.000 ansatte i it-branchen, hvilket er resultatet af en markant stigning de seneste år. Tallet steg med omkring 2.500 fra 2016 til 2017<sup>8</sup>. Der har dermed allerede været en markant stigning siden 2016, hvor efterspørgslen efter data scientists blev vurderet i forbindelse med oprettelsen af BSc i Data Science.

Det vurderes, at der i nær fremtid på internationalt plan vil være stor vækst i efterspørgslen på arbejdskraft med IKT-kompetencer<sup>9</sup>. Konkret vurderes det, at IKT-branchen er den sektor, der på verdensplan frem mod 2020 vil være i størst vækst. I Danmark vurderes det udækkede efterspørgselsbehov i 2020 at være på op mod 6.000 ITEK-kandidater, hvoraf 3.000 af disse skyldes fremdrift inden for særligt Big Data<sup>10</sup>. Et andet studie vurderer det udækkede efterspørgselspotentiale i Danmark i 2020 til at være 4.000 IKT-specialister<sup>11</sup>. Begge studier indikerer dermed et større udækket efterspørgselspotentiale i Danmark allerede i 2020.

*”Det er min erfaring, at hvis jeg slår en data scientist stilling op, er det muligt at ansætte en relevant kandidat på baggrund af ansøgerfeltet. Der vil typisk være ca. 30 ansøgere, hvoraf de fleste vil have en vis kompetence indenfor området. Der vil typisk være 1-2 kandidater, der lever op til vores høje krav (kompetence, erfaring, personlige egenskaber), og som vi ønsker at ansætte”.*

Bjørn Sandqvist, Tryg

<sup>6</sup> Informations- og Kommunikationsteknologi

<sup>7</sup> IT, tele, elektronik og kommunikation

<sup>8</sup> <https://itb.dk/articles/branchen-i-tal-1>

<sup>9</sup> World Economic Forum: "The Future of Jobs Report" (2018)

<sup>10</sup> Dansk Industri: "Prognose 2020 – ITEK-branchens behov for it- og elektronik kandidater i 2020" (2015)

<sup>11</sup> Højbjerg Brauer Schultz: "Virksomheders behov for digitale kompetencer" (2016)

En undersøgelse om rekrutteringsudfordringer i Norden fra 2016 viser desuden, at hvert fjerde rekrutteringsforsøg inden for IKT-branchen i Danmark var forgæves, og at den primære årsag hertil vurderes at være mangel på rette tekniske kompetencer<sup>12</sup>. At der eksisterer et aktuelt efterspørgselsbehov på data scientist understøttes også af de kvalitative interviews.

*”Jeg kan sagtens forestille mig, at vores rekruttering de næste par år resulterer i, at vi får ansat en række medarbejdere, der har en uddannelse som data scientist med i bagagen”. Peter Stoltze, Danmarks Statistik.*

Denne tendens går igen i flere af undersøgelsens interviews. André Stelsig Rogaczewski fra Netcompany påpeger, at: *”IT-branchen på nuværende tidspunkt kan opsuge alle kandidater”*. Flere af de adspurgte aftagerorganisationer og -virksomheder oplever dermed allerede en udækket efterspørgsel på data scientists. På baggrund af undersøgelsens desk study og kvalitative interviews kan det konkluderes, at der er klare tegn på, at der allerede inden for nærmeste fremtid vil være et stort udækket behov for højtuddannede IKT-kandidater, som netop besidder nogle af

de kompetencer, som ITU’s potentielle kandidatuddannelse i Data Science forventes at kunne imødekomme.

### 3.2 Efterspørgslen på data scientists stiger frem mod 2030

Den store efterspørgsel på kandidater med kompetencer inden for data science forventes at være stigende frem mod 2030 – både på internationalt plan og i Danmark. Ifølge Europa-Kommissionen er efterspørgslen efter specialister inden for digital teknologi i EU steget med fire pct. årligt det seneste årti<sup>13</sup>. Det tyder på, at der også fremadrettet kan forventes en stor stigning i efterspørgslen inden for IKT-branchen i en dansk kontekst.

Ovenstående tendens understøttes også af fremskrivninger for det danske arbejdsmarked. Frem mod 2030 vurderes det, at der vil være en potentiel jobvækst på 90.000-150.000 arbejdspladser direkte relateret til udvikling, implementering og løbende opdateringer af automatiseringsteknologi. På baggrund af en uddannelsesfremskrivning vurderes det dog, at uændrede uddannelsesmønstre vil lede til, at kun 70.000 i 2030 vurderes at have en tilstrækkelig avanceret og analytisk baggrund til at varetage disse jobs<sup>14</sup>. Det efterlader ifølge Højbjerg Brauer Schultz et udækket efterspørgselspotentiale på IKT-specialister med lange videregående uddannelser, som forventes at stige fra 4.000 i 2020 til 13.000 i 2030<sup>15</sup>.

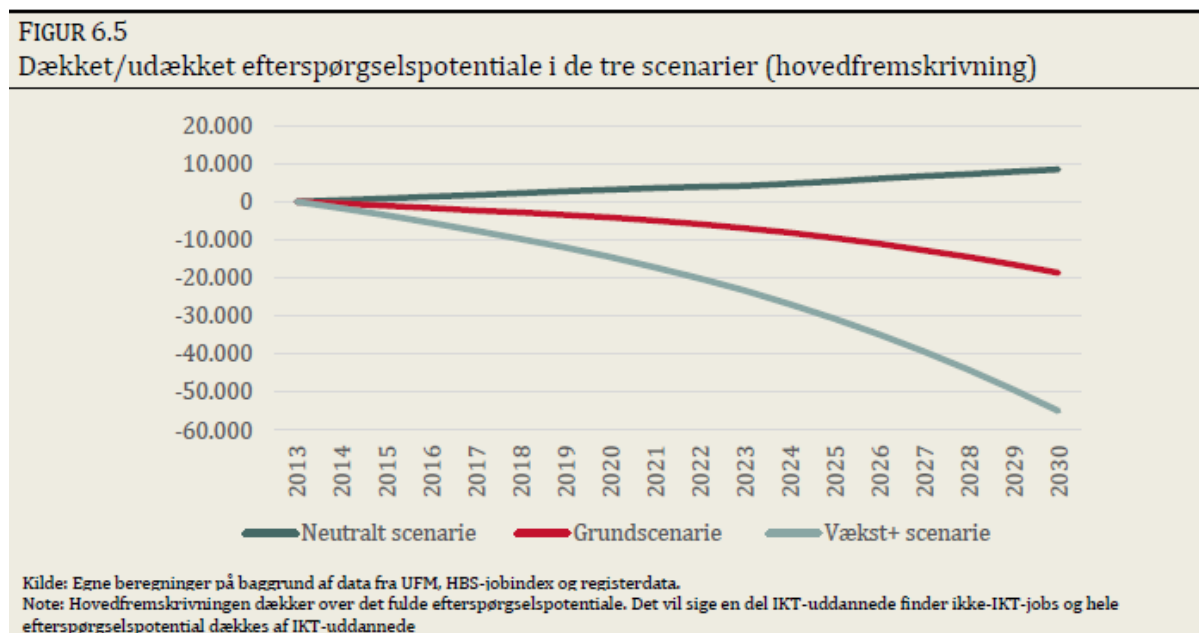
<sup>12</sup> Andreas Højbjerg & Rasmus Mikkelsen: ”Mismatchproblemer i Norden” (2017)

<sup>13</sup> European Commission: ”A New Skills Agenda for Europe” (2016)

<sup>14</sup> McKinsey: ”Automatiseringens effekter på det danske arbejdsmarked” (2017)

<sup>15</sup> Højbjerg Brauer Schultz: ”Virksomheders behov for digitale kompetencer” (2016)

Figur 3-1: Udviklingen i udækket efterspørgselspotentiale på IKT-specialister



Note: Undersøgelsens konklusion er baseret på grundscenariet, som kilden også selv anvender i deres konklusion. Figuren viser det udækket efterspørgselspotentiale for alle IKT-specialister. Dette vurderes at være 19.000 i 2030, hvoraf IKT-specialister med en lang videregående uddannelse forventes at udgøre 13.000. Det neutrale scenarie betyder, at efterspørgslen følger udviklingen i den generelle efterspørgsel; dette svarer til en årlig vækst på 0,4 pct. Grundscenariet betyder, at den fremtidige efterspørgsel følger udviklingen i beskæftigelse for IKT-specialister; dette svarer til en årlig vækst på 2,2 pct. "Vækst+"-scenariet betyder, at den fremtidige efterspørgsel fremskrives ud fra de seneste års tendenser i beskæftigelsen tillagt en ekstra vækstrate baseret på udviklingen i jobopslagsdata; dette svarer til en årlig vækstrate på omkring fire pct.

På tværs af de kvalitative interviews er den klare tendens også, at de adspurgte aftagerorganisationer og -virksomheder i høj grad forventer, at efterspørgslen på data scientists vil stige i fremtiden. Peter Stoltze fra Danmarks Statistik påpeger, at muligheden for at inddrage nye datakilder fremover er med til at øge efterspørgslen på data scientists. Også André Stelsig Rogaczewski fra Netcompany forudsiger "en voldsomt stigende efterspørgsel". Der er dermed en forventning om en stor stigning i efterspørgslen på data scientists – og dette gælder både for den offentlige og private sektor.

Der er dermed klare indikationer på, at der både på internationalt plan, europæisk plan og i Danmark forventes en markant stigning i efterspørgslen efter kandidater med de kompetencer, som data scientists besidder. Flere af de adspurgte virksomheder påpeger også, at manglen på arbejdskraft med de nødvendige kompetencer gør, at de enten er nødt til at søge udenlandsk arbejdskraft eller flytte opgaver til udlandet. Birgitte Hass fra IT-brancheorganisationen påpeger, at det allerede er et problem, da "der flytter kandidater og opgaver ud af landet på grund af mangel på data scientists". Det understøttes af flere af de adspurgte aftagervirksomheder, som fortæller at de allerede flytter opgaver ud af landet eller kan blive nødt til at gøre det i fremtiden. Den udækkede efterspørgsel er dog ikke kun et problem i Danmark. Birgitte Hass fra IT-brancheorganisationen tilføjer, at: "udbuddet af data scientists er lige så knap i udlandet. Derfor trækker udlandet også i dataspecialister fra Danmark". Indeværende undersøgelses interviews understreger dermed, at

"Der forventes en øget efterspørgsel på medarbejdere med kompetencer indenfor data science, da vi også er begyndt at orientere os mod nye datakilder som fx sensordata, satellitdata og data fra sociale medier". Peter Stoltze, Danmarks Statistik.

den udækkede efterspørgsel allerede udgør en udfordring for danske virksomheder, og at manglen på kvalificeret arbejdskraft både i Danmark og på internationalt plan medfører, at både kandidater og opgaver flytter ud af landet. Desuden påpeges det, at det også er vanskeligt at rekruttere udenlandsk arbejdskraft, da data scientists er i meget høj kurs på det globale marked.

### 3.3 Tekniske kompetencer er mere efterspurgt end tværgående kompetencer

På baggrund af undersøgelsens interviews tegner der sig et billede af, at de centrale aftagere af kandidater i Data Science har forskellige forventninger til, hvilke kompetencer en data scientist skal besidde. Overordnet kan der skelnes mellem følgende tre forskellige forventninger til vægtningen af kompetencer hos en data scientist:

1. Højt specialiserede tekniske og matematiske kompetencer (statistik, matematik, programmering)
2. Højt niveau af tekniske kompetencer samt kendskab til forretningsforståelse
3. Højt niveau af både tekniske kompetencer og forretningsforståelse.

Nogle af virksomhederne har en forventning om, at en data scientist primært skal have meget dybdegående matematiske og statistiske kompetencer. Fx mener Stig Geer Pedersen fra Topdanmark, at *”en data scientist mindst skal være på niveau med matematikere, statistikere og dataloger i forhold til de tekniske og statistiske kompetencer”*. Her hersker også en klar holdning til vægtningen af tekniske kompetencer og tværfaglige kompetencer, idet Stig Geer Pedersen fra Topdanmark uddyber, at: *”kommunikation og forretningsforståelse skal nedprioriteres, så de er højt specialiserede i deres arbejde med data”*.

Andre af de adspurgte virksomheder forventer, at en data scientist både har gode tekniske kompetencer, men samtidig også har et basalt kendskab til forretningsforståelse. Her fremhæves det også, at forretningsforståelse er vigtigt, men at det i højere grad kan tillæres i virksomhederne efter endt uddannelse.

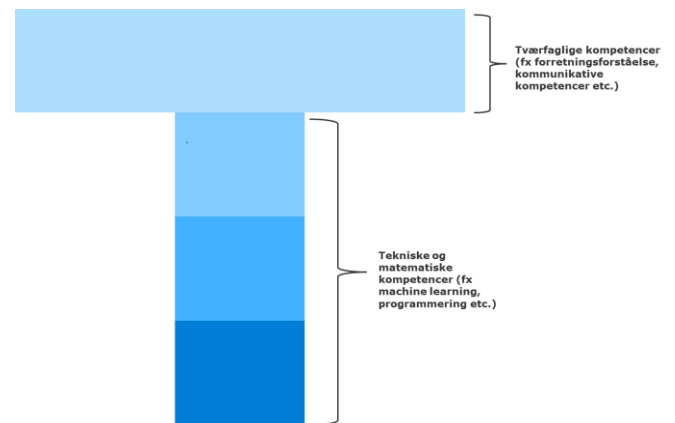
Den sidste kategori af forventninger til en data scientist dækker over de organisationer og virksomheder, som påpeger, at en data scientist både kombinerer et højt teknisk niveau med et godt kendskab til forretningsforståelse. IBM er blandt de aftagere, som netop mener, at den tværfaglige profil er det, der adskiller en data scientist fra andre lignende kandidater, som citatet fra Torben Noer fra IBM i boksen til højre understreger. Flere af de adspurgte aftagere er enige i dette. Bjørn Sandqvist fra Tryg påpeger, at *”de (kandidaterne, red.) skal også have (eller kunne tilegne sig) forretningsforståelse for at kunne indgå i virksomheden og skabe værdi”*.

*”En data scientist skal kombinere to profiler. Der er selvfølgelig de matematiske og statistiske kompetencer, og så er der de forretningsmæssige kompetencer. Kun at besidde de førstnævnte kompetencer gør dig ikke til en data scientist”*. Torben Noer, IBM

Denne sondring mellem, hvad aftagerorganisationer og -virksomheder lægger vægt på, er også relevant i forhold til at forstå, hvad der adskiller deres forventninger til en kandidat i Data Science fra deres forventninger til en bachelor i Data Science. De, som vægter de tekniske kompetencer højest, mener, at det netop er den højt specialiserede tekniske viden, som adskiller en kandidat fra en bachelor. De, som til gengæld mener, at data scientists også bør have kendskab til forretningsforståelse og kommunikation, mener, at disse kompetencer bør tilegnes på kandidatuddannelsen, og at det dermed er den mere tværfaglige profil, der adskiller en kandidat i Data Science fra en bachelor i Data Science.

Udover tværfaglige kompetencer som forretningsforståelse og kommunikation påpeger nogle af de adspurgte aftagervirksomheder og -organisationer også vigtigheden af etiske kompetencer i forbindelse med behandlingen af personfølsomme data. Rikke Zeberg fra Digitaliseringsstyrelsen siger, at *”Det er vigtigt at forstå kompleksiteter, dilemmaer og etik i arbejdet med persondata”*. Også Karen Marie Lyng fra Sundhedsdatastyrelsen fremhæver vigtigheden af viden om arbejdet med persondata, hvilket er særligt relevant i lyset af persondatalovgivningen, da *”Det er vigtigt at vide, hvilke regler der gælder i forhold til at arbejde med persondata”*.

Blandt aftagerorganisationer og -virksomheder er der dermed forskel i forventningerne til, hvilke kompetencer en data scientist skal besidde. De adspurgte virksomheder og organisationer mener, at en Data Science-uddannelse bør fokusere på at give de studerende et højt niveau af tekniske og matematiske kompetencer og til en vis grad suppleret med forretningsforståelse og kommunikative kompetencer. At der lægges vægt på de tekniske og matematiske kompetencer, understøttes desuden af, at flere af de adspurgte virksomheder nævner, at de på nuværende tidspunkt primært har ansatte med baggrunde inden for matematik, statistik, fysik, ingeniørvidenskab og lignende til at varetage arbejde med data science.



Det tegner sig således et billede af, at der er efterspørgsel efter data science-kandidater med såkaldte "T-formede kompetencer". Den vertikale del af T'et i figuren repræsenterer de dybdegående tekniske og matematiske kompetencer som fx matematik, statistik og programmering. Den horisontale del af T'et i figuren repræsenterer mere tværgående kompetencer, herunder forretningsforståelse og kommunikative kompetencer.

### 3.4 Fremover vil efterspørgslen på særligt tekniske kompetencer stige

Studier viser, at det inden for IKT-branchen på internationalt plan særligt er jobs inden for Big Data, machine learning og artificial intelligence, som er i vækst. Konkret estimeres det, at jobs inden for disse områder vil gå fra at udgøre 17 pct. af IKT-arbejdsstyrken i 2018 på verdensplan til at udgøre 33 pct. i 2022<sup>16</sup>. Der forventes dermed en kraftig stigning i efterspørgslen på disse kompetencer allerede inden for nærmeste fremtid. Det er også disse kompetencer, som primært fremhæves i undersøgelsens interviews. Birgitte Hass fra IT-brancheforeningen fremhæver fx, at *”Der skal være undervisning i selv-lærende algoritmer, som selv finder mønstre i data. Man skal også kunne vedligeholde algoritmerne, herunder forstå og kvalitetssikre dem”*. Derudover fremhæves det, at man skal kunne være i stand til at håndtere store og mere komplekse datasæt – og at denne evne bliver vigtigere og vigtigere fremadrettet.

### 3.5 Kandidater med både gode dansk- og engelskkundskaber foretrækkes

I forhold til spørgsmålet om der foretrækkes en dansk- eller engelsksproget uddannelse, er der blandende meninger blandt de adspurgte aftagerorganisationer og -virksomheder. For aftagere, som har danske kunder og løser opgaver for offentlige virksomheder og institutioner i Danmark, er det en forudsætning, at deres ansatte kan kommunikere godt på dansk og har forståelse for dansk kultur og samfund i samspillet med kunder. Andre af de adspurgte virksomheder udtaler, at de foretrækker en engelsksproget uddannelse, hvilket primært skyldes, at de er internationale virksomheder med afdelinger i udlandet. Størstedelen af aftagerorganisationerne og -virksomhederne foretrækker dog, at data science-kandidaterne besidder både gode dansk- og

<sup>16</sup> World Economic Forum: "The Future of Jobs Report" (2018)

engelskkundskaber. De er åbne overfor, at undervisningen også kan foregå på engelsk, som giver mulighed for at få et internationalt forskningsmiljø. Rikke Zeberg fra Digitaliseringsstyrelsen fremhæver, at *”Det er afgørende, at kandidaten kan læse, skrive og tale dansk – og forstå dansk kontekst. Det er ikke afgørende, om uddannelsen er defineret som dansk eller engelsk, men det er afgørende for den offentlige sektor, at uddannelsen er på et højt internationalt niveau, og dermed også, at der er et internationalt forskningsmiljø”*. Dette opnås formentlig i højere grad gennem en engelsksproget uddannelse, hvor ITU har mulighed for at tiltrække internationale forskere til uddannelsen.

Det er således ikke selve undervisningssproget, som er afgørende for aftagerorganisationer og -virksomheder. Det er snarere kandidaternes evner og kompetencer til at arbejde ind i en dansk kontekst. Hovedparten foretrækker, at kandidaterne har forståelse for dansk kultur og kan indgå i samspillet med danske kunder og samarbejdspartnere. De skal samtidig være i stand til at arbejde på engelsk på et højt fagligt niveau. Det er endvidere væsentligt, at majoriteten af kandidaterne er tilgængelige for det danske arbejdsmarked, og derved bidrage til at dække den store efterspørgsel, som flere aftagerorganisationer og -virksomheder oplever i dag.

### **3.6 Praktikophold er værdifulde, men kan være på bekostning af specialiserede kompetencer**

Selv om de adspurgte virksomheder generelt vægter højt specialiserede tekniske kompetencer til fordel for tværfaglige kompetencer, er størstedelen af de adspurgte organisationer og virksomheder generelt positive over for, at kandidatuddannelsen kan indeholde praktikophold. Der er således en efterspørgsel blandt organisationer og virksomheder efter både højt specialiserede tekniske kompetencer og erfaring fra praktikophold. Dette skyldes, at et praktikophold typisk vil forbedre de studerendes forretningsforståelse og kommunikative evner. På grund af tidsmæssige begrænsninger er der dog risiko for, at et praktikophold gennemføres på bekostning af de mere specialiserede evner, som de studerende kunne tilegne sig på uddannelsen, hvis de ikke er i praktik. For de organisationer og virksomheder, som til gengæld ønsker, at kandidaterne har et basalt kendskab til forretningsforståelse, anses praktik for at være en oplagt måde at imødekomme dette kompetencebehov på.

## 4. DEL II: UDBUDSDÆKNING AF LIGNENDE UDDANNELSER

Formålet med del II er at kortlægge de eksisterende kandidatuddannelser inden for Data Science og således arbejde med anden delundersøgelse: *En udbudsafdækning af kandidatuddannelser inden for Data Science*. Kapitlet er inddelt i følgende tre afsnit:

- I *første afsnit* fremlægges de identificerede danske kandidatuddannelser, der rummer tilstrækkeligt med centrale elementer inden for data science. Udover at redegøre for resultaterne af syntesen, analyseres også de udvalgte uddannelsers overlap med en potentiel kandidatuddannelse i Data Science.
- I *andet afsnit* præsenteres nye danske kandidatuddannelser inden for Data Science, som enten er blevet (foreløbigt) godkendt i ansøgningsrunde - efterår 2017, eller som har indsendt ansøgning i ansøgningsrunde - efterår 2018.
- I *tredje afsnit* vurderes det afslutningsvist, hvorvidt der er plads til en ny kandidatuddannelse i Data Science, ved at sammenholde de identificerede uddannelsers indhold med det indhold, som potentielt vil indgå i ITU's kandidatuddannelse i Data Science.

Figur 4-1: Oversigt over datagrundlag for del II



Note: Den samlede oversigt over de inkluderede uddannelser, herunder deres beskrivelser, fremgår af bilag 3.

I ovenstående figur ses datagrundlaget for del II, der består af 20 kandidatuddannelser, der indeholder elementer af data science, som fordeles på henholdsvis 11 inkluderede og ni ekskluderede uddannelser. De 11 inkluderede uddannelser er udvalgt på baggrund af deres overlap med en uddannelse inden for Data Science og indgår derfor i en mere dybdegående analyse af udbuddet for kandidatuddannelser inden for Data Science i Danmark. De ekskluderede uddannelser indeholder ikke tilstrækkeligt mange elementer af data science i uddannelsens profil. Som det fremgår af figuren, er der samtidig blevet lavet en underopdeling i form af hhv. eksisterende kandidat- eller civilingeniøruddannelser samt nye kandidatuddannelser. Endeligt fremgår de inkluderede uddannelser nederst i figuren.

### 4.1 Redegørelse for og analyse af lignende uddannelser

Rambøll har systematisk gennemgået danske uddannelser, der indeholder centrale aspekter af data science. Der er taget udgangspunkt i både eksisterende og nye kandidat- og civilingeniøruddannelser. Vi har udvalgt de identificerede uddannelser ud fra det indholdsmæssige fokus, de har, herunder det fagudbud, der er på uddannelsen, både fag, der er obligatoriske, men ligeledes fag, der indgår som en del af et specifikt specialiseringsområde eller kan tilvælges som valgfag. Specifikt har vi kigget efter fag, som forventes at give kompetencer inden for data science, jf. det, der bliver efterspurgt i del I, samt centrale elementer fra ITU's egen potentielle kandidatuddannelse i Data Science.

Uddannelser, der er inkluderet under den første screeningsproces, men ekskluderet efter den dybdegående relevansvurdering, indeholder dermed ikke tilstrækkeligt med teknologiske fag, som er påkrævet i arbejdet med data science, også kaldet Big Data. Generelt har de ekskluderede uddannelser ikke haft en stor nok hovedvægt på det at arbejde inden for data science, men indeholder kun et enkelt eller få elementer af de fag, som udgøres af en uddannelse i Data Science. Uddannelsernes hovedvægt er udledt med udgangspunkt i mængden af relevante fag inden for data science, herunder belastningen af faget, opgjort som antallet af ECTS-point. Desuden har en efterfølgende kvalitetsvurdering af uddannelsens relevans sikret en vis grad af overlap af elementer inden for data science, da der kan udledes paralleller mellem de kompetencer, der opnås på den inkluderede uddannelse og en uddannelse i Data Science. På denne baggrund er det således muligt at påpege indholdsmæssige tendenser inden for hver af de inkluderede uddannelser, og herved sammenligne indholdet med ITUs planer for en fokuseret og specialiseret overbygning til bacheloren i Data Science

### **Karakteristika for de inkluderede uddannelser**

Det, der har karakteriseret udbuddet af uddannelser, som er inkluderet i denne analyse, er det generelt store teknologiske fokus. Dette område er gennemgribende for hovedparten af uddannelserne, som ofte indeholder programmeringsfag samt arbejder med algoritmer. På tværs af de identificerede relevante uddannelser ses således også en overvægt af civilingeniøruddannelser. Det kendetegnende ved disse uddannelser er, at de alle er præget af at være meget teknologiske, vurderet ud fra antallet af ECTS-point, som udgør de fag, der relaterer sig til enten it, software, programmering eller lignende.

To af de seks civilingeniøruddannelser omhandler software (*Softwareudvikling på ITU og Software Engineering på SDU*) og er således ikke uddannelser, der fokuseret arbejder inden for data science. For at opnå kompetencegivende indhold på disse uddannelser kræves det, at den studerende specialiserer sig inden for det eneste relevante fagområde på uddannelsen, som indeholder elementer af data science. Tilvælges dette, indeholder uddannelsen dog flere relevante fag, der giver den studerende viden inden for mange områder, herunder maskinlæring, algoritmer, statistik, databaser, og visualisering. Uddannelsen i *Software Engineering på SDU* indeholder desuden et fag i Data Science (10 ECTS-point), som er obligatorisk. De resterende fagområder på de to software-uddannelser indeholder ikke nævneværdige elementer ift. data science, da fokus typisk er på udvikling af software.

Blandt de inkluderede uddannelser er der tre uddannelser (*Business Administration and Information Systems på – Data Science på CBS, Erhvervsøkonomi – Business Intelligence på AU, og Business Analytics på DTU*), som har et overordnet fokus på forretningsaspektet inden for data science. Her handler det bl.a. om, hvordan data kan anvendes til at løse virksomheders udfordringer og problematikker, med specielt udgangspunkt i forretningsdata. *Business Analytics på DTU*, som er en civilingeniøruddannelse, er den mest teknologisk orienteret uddannelse af de tre, hvor obligatoriske fag som eksempelvis Machine Learning and Data Mining (5 ECTS-point) indgår som en del af uddannelsens opbygning. Den mindst tekniske uddannelse, som er inkluderet i denne analyse, er *Erhvervsøkonomi med speciale i Business Intelligence på AU*. Uddannelsen har et stort fokus på at anvende data til beslutningstagning hos virksomheder, hvilket faguddannelsen også bærer præg af. For at uddannelsen kan blive konkurrencedygtig ift. en kandidat inden for Data Science kræves det, at hele uddannelsens tredje semester (valgfag på 30 ECTS-point) benyttes til at supplere med de mere teknologisk tunge fag inden for data science. Den sidste uddannelse af de tre, *Business Administration and Information Systems – Data Science på CBS*, er som navnet antyder, en forretnings- og kommunikationsorienteret uddannelse inden for data science. Uddannelsen er ny i 2018 og har et fagligt indhold, der befinder sig i krydsfeltet mellem de to ovennævnte uddannelser. Havde uddannelsen været mere teknologisk, og således

mindre orienteret på det forretnings- og kommunikationsmæssige indhold, ville parallelterne have været mere tydelige.

En anden kandidatuddannelse, der er værd at fremhæve, er kandidatuddannelsen i *Datalogi*, som udbydes på en række af de danske universiteter (KU, SDU, AU, AAU og ITU). 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. Opbygningen på kandidatuddannelsen i datalogi fremstår forholdsvis ens på tværs af de forskellige uddannelsesinstitutioner, da de alle giver mulighed for specialisering inden for områder, der relaterer sig til data science, enten i form af valgfag på uddannelsen, eller som en formaliseret retning, hvilket er tilfældet på KU, AAU og ITU. Det kendetegner derfor kandidatuddannelserne i datalogi, at der er stor valgfrihed til egen specialisering, og flere kandidater indeholder specifikke fag inden for data science.

#### 4.2 Nye uddannelser på vej inden for Data Science

Som en del af udbudsafdækningen er også nye danske kandidatuddannelser inden for Data Science blevet undersøgt. Her er det både blevet undersøgt, hvilke danske Data Science-uddannelser som er blevet (foreløbigt) godkendt i ansøgningsrunde - efterår 2017, samt hvilke uddannelser der har indsendt ansøgninger i ansøgningsrunde - efterår 2018. Resultatet af denne søgning er, at AU og AAU begge har fået en foreløbig godkendelse til en ny kandidatuddannelse i Datavidenskab (Data Science), og KU har ansøgt om en ny kandidatuddannelse i Social Datavidenskab (Social Data Science).

På AAU indeholder *kandidatuddannelsen i Datavidenskab* selvsagt en del elementer inden for data science. Uddannelsens opbygning bærer desuden præg af valgfrihed, som er indbygget i de tre første semestre. Fokus er på håndtering og beregning med avancerede og komplekse datatyper. Derudover opnås kompetencer inden for machine learning, intelligent dataanalyse, infrastruktur til Big Data, avanceret statistiske dataanalyser samt forståelse for at udføre dataanalyser inden for forretning, industri og ingeniør- og sundhedsvidenskab.

*AU's kandidatuddannelse i Datavidenskab* er indholdsmæssigt meget beslægtet med AAU's, men selve opbygningen af uddannelsen adskiller sig. På AU er der obligatoriske kurser såsom Advanced Statistical Learning og Large scale optimization (begge 10 ECTS-point). Derudover vælges et specialiseringsforløb, hvilket enten være en datalogisk/datasikkerheds orienteret specialisering, en statistisk orienteret specialisering, en teknisk orienteret specialisering, eller endelig en økonomisk og forretningsforståelsesmæssig specialisering. Rammen for specialiseringen udgør 30-60 ECTS-point, hvilket dermed bestemmer omfanget af kandidatspecialet.

På *KU's kandidatuddannelse i Social Datavidenskab* er der et stort fokus på at kombinere både samfundsvidenskab og datavidenskab. Dette giver en bred profil med både stærke analytiske og teknologiske kompetencer. Færdigheder inden for programmering, machine learning og natural language processing (NLP) er blot en del af undervisningen inden for data science. Derudover opnås samfundsvidenskabelige færdigheder som en kombination af både kvantitative og kvalitative empiriske metoder.

#### 4.3 Få danske uddannelser giver den efterspurgte kombination af kompetencer

På tværs af de inkluderede kandidatuddannelser, der er beslægtet med ITU's kandidat i Data Science, tegner der sig et billede af et specialiseret område inden for den overordnede uddannelsesretning. Det er et gennemgående tema for de fleste af civilingeniøruddannelserne samt uddannelsen i Datalogi, at de studerende får konkrete teknologiske værktøjer og viden, der

gør dem i stand til at bearbejde og håndtere store datamængder samt udvikle software og større it-systemer. Det betyder også, at fag, der relaterer sig mere specifikt til data science, sjældent er særligt fremtrædende. Både datalogi- og civilingeniøruddannelserne giver dermed i højere grad generelle kompetencer inden for det teknologiske og datalogiske område, hvor det så er gjort muligt at specialisere sig mere specifikt inden for data science. Derfor bør det fremhæves, at det især er de teknologiske fag, der dominerer, og der kun er begrænset mulighed for at tage fag inden for eksempelvis forretningsforståelse og derigennem mulighed for at tilegne sig andre ikke-teknologiske kompetencer, der kan være relevante for en data scientist. For de øvrige kandidatuddannelser gør det modsatte sig i flere tilfælde gældende. Her er der generelt mangel på de mere teknologiske færdigheder, som er helt fundamentale for en data scientist, hvorfor det bliver et krav, at dette kan tilvælges uddannelsen i tilstrækkelig grad. Det skal dermed understreges, at data science får karakter af en teknologisk specialiseringsvej fremfor en hovedvej, der også indbefatter tilegnelsen af et bredere sæt kompetencer for de studerende. Det vurderes således, at det kun i mindre grad er muligt for de beslægtede kandidatuddannelser at give de studerende de nødvendige kompetencer inden for data science.

I forlængelse heraf skal det nævnes, at hvis ikke ITU's kandidatuddannelse i Data Science oprettes, vil ITU's bachelorer i Data Science finde andre overbygninger, som ikke bygger direkte oven på ITU's egen bacheloruddannelse. Dette giver mindre sandsynlighed for, at vi får de stærkeste kandidater med kompetencer inden for data science ud, hvilket ikke er overensstemmelse med de tilkendegivelser, der er udledt fra interviews med aftagervirksomheder og -organisationer. Det skal imidlertid bemærkes, at ITU's bacheloruddannelse i Data Science udbydes på engelsk, hvorfor en eventuel dansksproget overbygning potentielt kan afholde nogle af ITU's egne bachelorer i Data Science fra at søge ind. Det scenarie flugter ikke med de ønsker, som aftagervirksomheder og -organisationer har, da de i højere grad efterspørger kompetencer, som bygger direkte ovenpå bacheloren.

#### **4.4 Lav ledighed for beslægtede uddannelser**

For de kandidatuddannelser, som er inkluderet i denne udbudsafdækning, er det meget bemærkelsesværdigt, at samtlige uddannelser har en lav dimittendledighed på mellem to og syv pct. Dette skal holdes op mod det overordnede gennemsnit på tværs af alle kandidatuddannelser i Danmark, som er på ca. 18 pct. Tallene er for dimittender i 2016 og er målt tre kvartaler efter fuldførelse, hvilket er de seneste opgjorte tal for Uddannelses- og Forskningsministeriets aktuelle ledighed blandt dimittender fra videregående uddannelser. En så lav dimittendledighed for samtlige uddannelser i udbudsafdækningen tyder på et stort behov for de kompetencer, der er relateret til data science.

## 5. BILAG OG METODE

Nedenfor fremgår undersøgelsens tre bilag, som relaterer sig til undersøgelsens tre datakilder. I hvert bilag præsenteres de metodiske overvejelser, der ligger til grund for arbejdet med de respektive datakilder.

- I bilag 1 ses søgestrategien samt de udvalgte studier til undersøgelsens desk study.
- I bilag 2 fremgår en oversigt over de aftagerorganisationer og -virksomheder, som er blevet interviewet i forbindelse med efterspørgselsafdækningen.
- I bilag 3 ses en oversigt over den metodiske fremgangsmåde for udbudsafdækningen samt en oversigt over inkluderede uddannelser.

## 5.1 BILAG 1: Desk study

I dette bilag fremgår de udvalgte studier i undersøgelsens desk study samt det metodiske grundlag for udvælgelsen af de konkrete studier.

### Søgestrategi for desk study

For at identificere nye studier om arbejdsmarkedets behov siden markedsundersøgelsen i 2016<sup>17</sup> er der opstillet en systematisk søgestrategi med fastlagte søgetermer og inklusionskriterier. Denne systematiske søgestrategi er kombineret med snowballing-metoden, hvor identificerede studiers referencelister også er screenet for relevante studier. Søgestrategien er skitseret i nedenstående boks.

### Litteratur om efterspørgslen på Data Science-kandidater

#### Søgetermer

Data Science, Big Data, statistik, maskinlæring, dataarkitektur, programmering, algoritmer, data management, dataingeniør, data mining, data analytics, datavisualisering, Business Intelligence, Natural Language Processing (NLP), kunstig intelligens, Artificial Intelligence (AI).

Søgetermer suppleres af ord som: arbejdsmarked, beskæftigelse, employment, job opportunities.

#### Inklusionskriterier

Sprog: Dansk og engelsk

Geografi: Hele verden

År: 2014 – nu (størst fokus på studier fra 2016 og senere).

#### Søgeproces

1. Søgeord anvendes i forskellige kombinationer i en søgning på gængse databaser (fx Google).
2. Relevansvurdering af fremkomne studier (tema, årstal etc.)
3. Supplerende søgning pba. de inkluderede studiers referencelister.
4. Etablering af bruttoliste af studier til gennemlæsning.
5. Gennemlæsning af studier – frascreening af irrelevante studier.
6. Syntese udarbejdes.

### Oversigt over udvalgte studier til desk study

På baggrund af ovenstående beskrevne søgestrategi er følgende 10 studier inkluderet i undersøgelsens desk study om efterspørgslen på kandidater i Data Science. I nedenstående tabel er studierne hovedpointer opsummeret.

<b>Forfattere</b>	<b>Højbjerg Brauer Schultz, Kubix, Alexandra Institutet (styregruppe med repræsentanter fra Erhvervs- og Vækstministeriet), Uddannelses- og Forskningsministeriet og Ministeriet for Børn, Undervisning og Ligestilling</b>
Titel	Virksomheders behov for digitale kompetencer
Årstal	2016

<sup>17</sup> Rambøll 2016: Markedsundersøgelse for data science.

Land	Danmark
Hovedpointer	I 2030 vil der være et udækket efterspørgselspotential på 13.000 IKT-specialister med en lang videregående uddannelse.

<b>Forfattere</b>	<b>McKinsey &amp; Company</b>
Titel	Automatiseringens effekter på det danske arbejdsmarked
Årstal	2017
Land	Danmark
Hovedpointer	Der er potentiale for 90.000-150.000 jobs direkte relateret til udvikling og indførelse af automatiseringsteknologi frem mod 2030. Med uændrede uddannelsesmønstre vil der kun være 70.000-80.000 med tilstrækkelig tung analytisk baggrund. Dette kan lede til et opadgående lønpres, som gør det mindre rentabelt for virksomhederne at indføre teknologien, eller øge behovet for at tiltrække højt kvalificeret international arbejdskraft.

<b>Forfattere</b>	<b>Dansk Industri</b>
Titel	Prognose 2020 – ITEK-branchens behov for it og elektronikkandidater i 2020
Årstal	2015
Land	Danmark
Hovedpointer	<p>Der vil være et udækket behov i 2020 på omkring 3.000 kandidater alene i ITEK-branchen. Derudover vil der være et udækket behov på op til yderligere 3.000 kandidater grundet den generelle teknologiske udvikling i samfundet, herunder Big Data, digitalisering og Internet of Things.</p> <p>Prognosen vurderer, at der vil være et stigende behov i branchen og dermed opstå mangel på højtuddannet arbejdskraft. Herudover vurderes det, at flere og flere brancheområder ændrer deres forretningsmodel og udbygger deres varer og services med avanceret elektronik og digitale tjenester.</p> <p>Konkret påpeges, at en langsigtet holdbar løsning kræver både:</p> <ol style="list-style-type: none"> <li>1. At optaget til uddannelserne skal øges markant.</li> <li>2. At profilen til en række uddannelser skal tilpasses til behovet.</li> <li>3. At gennemførelsesprocenten skal øges.</li> <li>4. At der skal kunne tiltrækkes flere udlændinge med disse uddannelser til det danske arbejdsmarked.</li> </ol>

<b>Forfattere</b>	<b>Højbjerg Brauer Schultz</b>
Titel	Virksomheders efterspørgsel efter tekniske og digitale kompetencer (STEM-kompetencer)
Årstal	2018
Land	Danmark
Hovedpointer	Rapporten analyserer private virksomheders efterspørgsel efter tekniske og digitale kompetencer.

	38 procent af jobopslag inden for den private sektor, som er målrettet STEM-kompetencer, vedrører kompetencer inden for IKT. Stigningen i efterspørgslen på STEM-kompetencer er således drevet af en stigende efterspørgsel på IKT-kompetencer.
<b>Forfattere</b>	<b>Andreas Højbjerg &amp; Rasmus Mikkelsen</b>
Titel	Mismatchproblemer i Norden: En kortlægning af rekrutteringssituationen i nordiske væksterhverv
Årstal	2017
Land	Danmark
Hovedpointer	<ol style="list-style-type: none"> <li>1. Rekrutteringsudfordringen er størst inden for <i>IKT</i> og <i>Digital automatiseret produktion</i>. Inden for IKT er næsten hvert tredje rekrutteringsforsøg forgæves.</li> <li>2. Halvdelen af de virksomheder, som har oplevet forgæves rekrutteringsforsøg, angiver, at årsagen var at ansøgerne ikke havde de rette tekniske kompetencer.</li> <li>3. Rekrutteringsudfordringen skyldes både volumenmæssigt mismatch (ikke nok kandidater) og indholdsmæssigt mismatch (ikke de rette kompetencer).</li> <li>4. Inden for IKT og digital automatiseret produktion er det henholdsvis 60 % og 70 % af virksomhederne i Danmark, som mener, at uddannelserne dækker virksomhedens behov.</li> </ol>
<b>Forfattere</b>	<b>Højbjerg Brauer Schultz</b>
Titel	Barrierer for virksomheders dataanvendelse
Årstal	2017
Land	Danmark
Hovedpointer	<p>Der er et behov for at tilpasse udbuddet af uddannelser, så virksomhedernes behov i højere grad dækkes.</p> <p>Omkring hver tredje virksomhed i undersøgelsen angiver, at de har svært ved at rekruttere og tiltrække medarbejdere, som har de rette kompetencer til at understøtte virksomhedens brug af data. Det indikerer, at der kan være behov for at tilpasse uddannelserne/udbyde nye uddannelser, så virksomhedernes kompetencebehov i højere grad dækkes.</p>
<b>Forfattere</b>	<b>The World Economic Forum</b>
Titel	The Future of Jobs Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution
Årstal	2018
Land	Hele verden
Hovedpointer	<p>Arbejdsmarkedsudviklingen frem til 2022 viser, at der inden for IKT-branchen særligt vil efterspørges færdigheder inden for Big Data og cloud teknologi.</p> <p>Den væsentligste barriere for at virksomheder ikke anvender ny teknologi er manglende færdigheder på det lokale arbejdsmarked. Dette angiver 74 % af de adspurgte virksomheder, som en årsag hertil.</p> <p>Jobs inden for kunstig intelligens-, Big Data- og machine learning etc. vil gå fra at udgøre 17 % i 2018 til 33 % af IKT-arbejdsstyrken i 2022.</p>

	Det fremtidige arbejdsmarked efterspørger højt kvalificeret arbejdskraft, herunder kvalifikationer inden for Big Data, artificial intelligence etc.
--	---

<b>Forfattere</b>	<b>European Commission</b>
Titel	Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions: A New Skills Agenda for Europe
Årstal	2016
Land	Bruxelles
Hovedpointer	Efterspørgslen på professionelle inden for digital teknologi er steget med 4 pct. pr. år i det seneste årti.

<b>Forfattere</b>	<b>Erhvervsministeriet</b>
Titel	Strategi for Danmarks digitale vækst
Årstal	2018
Land	Danmark
Hovedpointer	<ol style="list-style-type: none"> <li>1) Danmark har relativt få i arbejdsstyrken med en videregående digital eller teknisk uddannelse sammenlignet med øvrige nordeuropæiske lande.</li> <li>2) Virksomhederne har omfattende rekrutteringsudfordringer inden for en række tekniske, digitale og naturvidenskabelige områder.</li> <li>3) Efterspørgslen efter arbejdskraft med tekniske, digitale og naturvidenskabelige kompetencer vil fortsætte med at stige, og der kan blive en mangel i 2025 på ca. 10.000 ingeniører og naturvidenskabelige kandidater.</li> <li>4) Omkring halvdelen af de virksomheder, som i 2016 har forsøgt at rekruttere IT-specialister, angiver at forsøget var forgæves.</li> </ol>

<b>Forfattere</b>	<b>IT-Branchen</b>
Titel	IT-Branchen i tal
Årstal	2018
Land	Danmark
Hovedpointer	I 2017 var der mere end 87.000 ansatte i it-branchen. Det er steget fra omkring 80.000 i 2014. Konkret er tallet steget med ca. 2.500 fra 2016 til 2017.

## 5.2 BILAG 2: Overblik over interviews inkl. metode

### Kvalitative interviews med aftagersegment

For at supplere og kvalificere kortlægningen af den danske efterspørgsel efter data scientists er der foretaget 10 kvalitative interviews med repræsentanter fra danske aftagervirksomheder og -organisationer. I nedenstående figur ses en oversigt over fordelingen af respondenter i 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 fremtidigt ønske til arbejdet med data science
- Efterspurgte medarbejderprofiler og kompetencer
- Ønsker til potentiel uddannelse indenfor data science.

Respondenterne har primært haft stilling som chef, direktør og projektleder i virksomhedens ledelse eller i relevante teknologiske og analytiske afdelinger. Respondenterne fremgår af nedenstående oversigt.

Navn	Virksomhed/organisation	Stilling
Rikke Zeberg	Digitaliseringsstyrelsen	Direktør
Søren Reeberg Nielsen	Styrelsen for Dataforsyning og Effektivisering	Vicedirektør
Peter Tibert Stoltze	Kontorchef for Metode og Analyse	Danmarks Statistik
Karen Marie Lyng	Afdelingschef for Datakvalitet og Indhold	Sundhedsdatastyrelsen
Birgitte Hass	IT-Brancheforeningen	Direktør
Torben Noer	IBM	Global Data Science and Business Analytics Technical Sales Leader
André Stelsig Rogaczewski	Netcompany A/S	Adm. direktør
Simon H. Galsgaard	DSV Road Danmark	Executive Vice President
Stig Geer Pedersen	Topdanmark	Seniorprojektleder for Machine Learning og RPA
Bjørn Lunding Sandqvist	Tryg	Direktør for Pricing & Modelling

### 5.3 BILAG 3: Liste over inkluderede uddannelser inkl. metode

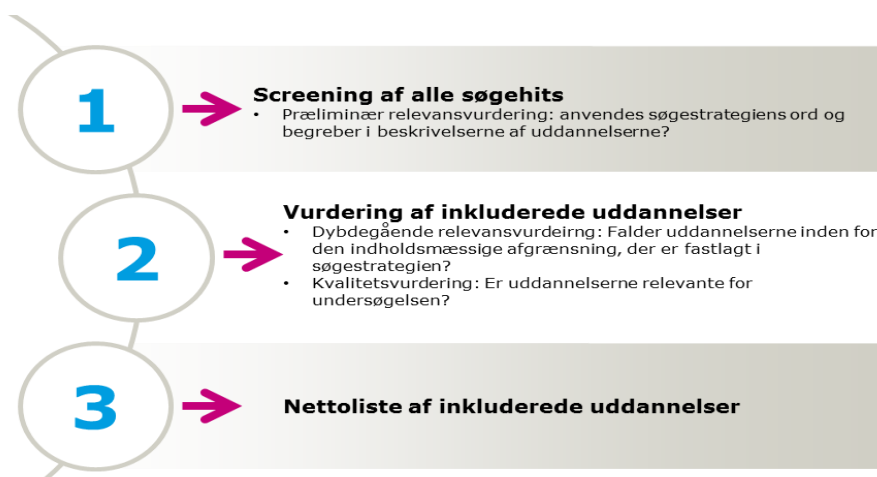
#### Metodisk tilgang til del II

Del II i nærværende undersøgelse har haft til hensigt at kortlægge eksisterende og nye danske uddannelser, der indeholder elementer af data science. Søgestrategien har været essentiel for at sikre en dækkende kortlægning 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 konkurrerende danske kandidatuddannelser er der inden for Data Science?
Søgetermer	Data Science, Big Data, statistik, maskinlæring, dataarkitektur, programmering, algoritmer, data management, dataingeniør, data mining, data analytics, data visualisering, Business Intelligence, Natural language processing (NLP), kunstig intelligens, artificial intelligence (AI)
Inklusionskriterier	<i>Typen af uddannelser:</i> Kandidat- og civilingeniøruddannelser  <i>Sproglig og geografisk afgrænsning:</i> Dansk- og engelsksprogede uddannelser i Danmark (både Øst- og Vestdanmark)
Kilder	Uddannelsesinstitutionernes hjemmesider: Danske universiteter: <a href="http://www.au.dk/">http://www.au.dk/</a> <a href="http://www.aau.dk/">http://www.aau.dk/</a> <a href="http://www.cbs.dk/">http://www.cbs.dk/</a> <a href="http://www.dtu.dk/">http://www.dtu.dk/</a> <a href="http://www.itu.dk/">http://www.itu.dk/</a> <a href="http://www.ku.dk/">http://www.ku.dk/</a> <a href="http://www.ruc.dk/">http://www.ruc.dk/</a> <a href="http://www.sdu.dk/">http://www.sdu.dk/</a>  Uddannelsesguiden: <a href="https://www.ug.dk/">https://www.ug.dk/</a>  Nye uddannelser: <a href="http://pkf.ufm.dk/">http://pkf.ufm.dk/</a>

Den korte tidsramme for kortlægningen har stillet krav til balancen mellem alsidighed og præcision. Vores erfaringer fra tidligere kortlægninger viser, at når alsidighed øges, så mindskes præcision. Det sker med risiko for, at søgeresultaterne falder uden for søgekriterierne. Det er derfor afgørende, at søgningen blev gennemført systematisk og fokuseret fra start. Undersøgelsen af danske uddannelser har derfor fra start haft fokus på de danske universiteters uddannelsesudbud. På den baggrund blev der udarbejdet en bruttoliste af danske uddannelser. Næste skridt var at gennemføre en systematisk og kritisk vurdering af de indsamlede uddannelser og dermed ekskludere eventuelle uddannelser, der faldt uden for den indholdsmæssige afgrænsning.

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<sup>18</sup>. 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, indgik i en dybdegående relevansvurdering og kvalitetsvurdering i andet trin.
- 2 Den dybdegående relevansvurdering i andet trin omfattede en endelig vurdering af, om uddannelserne faldt inden for kortlægningens indholdsmæssige afgrænsning. Her blev der kigget nærmere uddannelsesinstitutionernes eget materiale om de enkelte uddannelser, herunder studieordninger, og hvorvidt dette overlapper med planerne for en kandidatuddannelse i Data Science. Her har vi især skelet til opbygningen af uddannelserne, deriblandt obligatoriske fag på uddannelserne og muligheder for valgfag. Ved de medtagne uddannelser er der derefter undersøgt optag på kandidatuddannelserne samt dimittendledighed, for at give en vurdering af, hvor efterspurgte uddannelserne og dimittenderne er.
- 3 På baggrund af relevans- og kvalitetsvurderingen er der udarbejdet en nettoliste over de inkluderede 11 danske uddannelser inden for data science, som fremgår nedenfor.

### De inkluderede danske uddannelser

De identificerede uddannelser præsenteres i tabelform nedenfor.

Uddannelsens titel	<b>Datalogi</b>
Niveau og uddannelsesinstitution	Kandidatuddannelse på AU, AAU, KU, SDU, og ITU.
Link	<a href="https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/naturvidenskabeligeogkandidatuddannelser/matematikfysikkemiogdatalogi/datalogi">https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/naturvidenskabeligeogkandidatuddannelser/matematikfysikkemiogdatalogi/datalogi</a> <a href="https://studier.ku.dk/kandidat/datalogi/">https://studier.ku.dk/kandidat/datalogi/</a> <a href="https://www.itu.dk/uddannelser/kandidatuddannelser/datalogi">https://www.itu.dk/uddannelser/kandidatuddannelser/datalogi</a> <a href="https://www.sdu.dk/da/Uddannelse/Kandidat/Datalogi">https://www.sdu.dk/da/Uddannelse/Kandidat/Datalogi</a> <a href="http://kandidat.au.dk/datalogi/">http://kandidat.au.dk/datalogi/</a> <a href="https://www.aau.dk/uddannelser/kandidat/datalogi">https://www.aau.dk/uddannelser/kandidat/datalogi</a>

<sup>18</sup> Uddannelsesguiden udgives af Ministeriet for Børn, Undervisning og Ligestilling og indeholder alle anerkendte danske uddannelser.

Grundlæggende indhold	Uddannelsen handler om at analysere og løse komplekse problemer i forbindelse med it. Indeholder programmering, softwarearkitektur, databaser, algoritmer og datastrukturer og systemudvikling. Kompetencer i at skabe og udvikle forståelse for data.
Elementer af data science i uddannelsens profil	<p>Datalogiddannelsen er forholdsvis teknisk i forhold til denne kortlægning. På kandidatuddannelserne er der en række relevante specialiseringsmuligheder.</p> <p>På KU er uddannelsen opbygget af både obligatoriske kurser som programmering, algoritmer og datastrukturer, computersystemer, og machine learning (alle 7,5 ECTS). Derudover kan man bruge de valgfrie og begrænset valgfrie kurser til at specialisere sig inden for data science – ofte omtalt som Big Data, der handler om at uddanne specialister i data-baseret beslutningsstøtte. Der undervises i informationssøgning, machine learning 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.</p> <p>På ITU tilbydes flere forskelle specialiseringer, hvilket udgøres af en organiseret samling af kurser (22,5 ECTS). Her er der mulighed for at specialisere sig i data systems, som giver en dybdegående forståelse for computersystemer og softwareinfrastruktur til at lagre og behandle data, samt hvordan disse designes, implementeres og analyseres. Derudover kan en specialisering indenfor machine learning give både en praktisk og teoretisk forståelse af moderne machine learning-metoder, herunder kunstig intelligens og arbejde med store datamængder.</p> <p>På SDU er der mulighed for stor valgfrihed i opbygningen af uddannelsen. Der er således mulighed for at tage flere fag omhandlede datasystemer. Mest relevant er fagene; Big Data (10 ECTS), algoritmer og datastrukturer (10 ECTS), avancerede datastrukturer (10 ECTS), samt diverse programmeringsfag.</p> <p>På AU er kandidatuddannelsen på engelsk og indeholder 90 valgfri ECTS, hvoraf 60 ECTS kan udgøres af hhv. algoritmer (30 ECTS) og programmering (30 ECTS). Derudover kan relevante fag være machine learning (10 ECTS), avanceret datastrukturer (10 ECTS) og avanceret data management og analyse (10 ECTS).</p> <p>På AAU kan man på andet år specialisere sig indenfor syv forskellige områder. Specialiseringsområder der relaterer sig mest til uddannelsen i Data Science er databaseteknologi, maskinintelligens, og programmeringsteknologi.</p>
Oplysninger relateret til efterspørgsel	<p>Optag (2017): 232</p> <p>Antal dimittender (2016): 166</p> <p>Dimittendledighed (2016 k3): 6 % (gennemsnit for alle uddannelsesinstitutioner)</p>
Uddannelsens titel	<b>Computer Science (IT)</b>
Niveau	Kandidatuddannelse (civilingeniøruddannelse) på AAU
Link	<a href="https://www.ug.dk/uddannelser/universitetsuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingenioeruddannelser/computer-science-it">https://www.ug.dk/uddannelser/universitetsuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingenioeruddannelser/computer-science-it</a> <a href="https://www.aau.dk/uddannelser/kandidat/computer-science-it">https://www.aau.dk/uddannelser/kandidat/computer-science-it</a>
Grundlæggende indhold	Uddannelsen er international og foregår på engelsk. Kompetencer opnås inden for programmering, software engineering, maskinintelligens, datahåndtering, samt indlejrede systemer.
Elementer af data science i uddannelsens profil	<p>Computer Science indeholder nogle af de samme elementer fra data science set fra et teknisk perspektiv. Det specielle ved uddannelsen i computer science er, at det er muligt på 2. semester af kandidaten, at følge ét af fire forskellige spor:</p> <ul style="list-style-type: none"> <li>• Information Technology</li> <li>• Machine Intelligence</li> <li>• Data Engineering</li> <li>• Embedded Systems.</li> </ul> <p>Fælles for alle fire retninger er de obligatoriske fag inden for programmering og software engineering. De øvrige valgfag vælges med udgangspunkt i de valgte spor (alle 5 ECTS). Her kan bl.a. nævnes:</p>

	<ul style="list-style-type: none"> <li>• Advanced Topics in Databases</li> <li>• Machine Intelligence</li> <li>• Advanced Algorithms</li> <li>• Advanced Programming</li> <li>• Data-Intensive Systems</li> </ul>
Oplysninger relateret til efterspørgsel	<p>Optag (2017): <i>Ikke oplyst</i></p> <p>Antal dimittender (2016): <i>Ikke oplyst</i></p> <p>Dimittendledighed (2016 k3): <i>Ikke oplyst</i></p>
<b>Uddannelsens titel</b>	<b>Business Analytics</b>
Niveau	Kandidatuddannelse (civilingeniøruddannelse) på DTU
Link	<a href="https://www.ug.dk/uddannelser/universitetsuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingenioeruddannelser/business-analytics">https://www.ug.dk/uddannelser/universitetsuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingenioeruddannelser/business-analytics</a> <a href="https://www.dtu.dk/uddannelse/kandidat/business-analytics">https://www.dtu.dk/uddannelse/kandidat/business-analytics</a>
Grundlæggende indhold	Kandidatuddannelsen i Business Analytics bygger på de fire kernekompetencer for området: statistik, datalogi (computer science), operationsanalyse (optimering) og management (business)
Elementer af data science i uddannelsens profil	<p>Uddannelsen i Business Analytics er teknisk, og der skabes viden om, hvordan data kan anvendes til at løse virksomheders udfordringer. Der opnås indsigter i, hvordan man bedst anvender nye teknologier til at analysere data og løse problemer inden for økonomi og management. I obligatoriske fag, som eksempelvis machine learning and data mining (5 ECTS), undervises der i metoder til at arbejde med store datamængder. Desuden skal den studerende også undervises i teknologi, økonomi, ledelse og organisation.</p> <p>Derudover kan der valgfrit vælges mellem forskellige fag som:</p> <ul style="list-style-type: none"> <li>• Algorithms and Data Structures (5 ECTS)</li> <li>• Algorithms for Massive Data Sets (7,5 ECTS)</li> <li>• Artificial Intelligence and Multi-Agent Systems (7,5 ECTS)</li> <li>• Social Data Analysis and Visualization (5 ECTS)</li> <li>• Advanced Machine Learning (5 ECTS)</li> <li>• Mathematical Programming Modelling (5 ECTS)</li> </ul> <p>Derudover udbydes flere fag inden for statistik, business mm.</p>
Oplysninger relateret til efterspørgsel	<p>Optag (2017): <i>Ikke oplyst</i></p> <p>Antal dimittender (2016): <i>Ikke oplyst</i></p> <p>Dimittendledighed (2016 k3): <i>Ikke oplyst</i></p>
<b>Uddannelsens titel</b>	<b>Digitale medieteknologier</b>
Niveau	Kandidatuddannelse (civilingeniøruddannelse) på DTU
Link	<a href="https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingenioeruddannelser/digitale-medieteknologier">https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingenioeruddannelser/digitale-medieteknologier</a> <a href="https://www.dtu.dk/uddannelse/kandidat/digitale-medieteknologier">https://www.dtu.dk/uddannelse/kandidat/digitale-medieteknologier</a>
Grundlæggende indhold	Uddannelsen er engelsksprogede og handler om at udvikle services og løsninger inden for informations- og kommunikationsteknologi.
Elementer af data science i uddannelsens profil	<p>Uddannelsen er blandt de mere tekniske. Der er dog også undervisning, der giver en vis forretningsforståelse. Man kan specialisere sig inden for ét ud af fem områder, og i forhold til denne kortlægning af konkurrerende uddannelser, er det hovedsageligt det område der hedder "Data Science", der indeholder elementer relateret til data science. En specialisering indenfor denne retning indeholder bl.a. fagene:</p> <ul style="list-style-type: none"> <li>• Computational tools for Big Data (5 ECTS)</li> <li>• Algorithms for massive data sets (7,5 ECTS)</li> <li>• Social Data Analysis and Visualization (5 ECTS)</li> <li>• Introduction to machine learning and data mining (5 ECTS)</li> </ul>
Oplysninger relateret til efterspørgsel	<p>Optag (2017): 59</p> <p>Antal dimittender (2016): 44</p>

	Dimittendledighed (2016 k3): 7 %
Uddannelsens titel	<b>Business Administration and Information Systems - Data Science</b>
Niveau	Kandidatuddannelse på CBS
Link	<a href="https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/samfundsvidenskabeligekandidatuddannelser/erhvervsøkonomi/business-administration-and-information-systems">https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/samfundsvidenskabeligekandidatuddannelser/erhvervsøkonomi/business-administration-and-information-systems</a> <a href="https://www.cbs.dk/uddannelse/kandidat/msc-in-business-administration-and-information-systems/msc-in-business-administration-and-information-systems-data-science">https://www.cbs.dk/uddannelse/kandidat/msc-in-business-administration-and-information-systems/msc-in-business-administration-and-information-systems-data-science</a>
Grundlæggende indhold	Uddannelsen kombinerer forretningsforståelse og dataanalyse, hvilket giver kompetencer til at håndtere store mængder forretningsdata, som kan anvendes af virksomheder i deres beslutningstagning.
Elementer af data science i uddannelsens profil	Uddannelsen er bygget op omkring evnen til at kunne arbejde med data i et forretningsmæssigt landskab. Dermed er dataanalytiske metoder, programmering, og andre tekniske færdigheder grundlæggende værktøjer for denne uddannelse, som kan anvendes til at skabe værdi for virksomheder. Undervisningen kombinerer således den mere tekniske del af arbejdet med data med en forretningsmæssig/business del. Der er en række obligatoriske fag, som indeholder elementer af data science: <ul style="list-style-type: none"> <li>• Foundations of Business Data Analytics: Architectures, Statistics and Programming</li> <li>• Data Mining, Machine Learning, and Deep Learning</li> <li>• Data Economics</li> </ul> <p>Alle obligatoriske fag udgør hver især 7,5 ECTS. Uddannelsen giver desuden mulighed for praktik på 3. semester, som maksimalt kan udgøre 15 ECTS.</p>
Oplysninger relateret til efterspørgsel	Optag (2017): <i>Ikke oplyst</i> Antal dimittender (2016): <i>Ikke oplyst</i> Dimittendledighed (2016 k3): <i>Ikke oplyst</i>
Uddannelsens titel	<b>Matematisk Modellering og Computing</b>
Niveau	Kandidatuddannelse (civilingeniøruddannelse) på DTU
Link	<a href="https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingeniøruddannelser/matematisk-modellering-og-computing">https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingeniøruddannelser/matematisk-modellering-og-computing</a> <a href="https://www.dtu.dk/uddannelse/kandidat/matematisk-modellering-og-computing">https://www.dtu.dk/uddannelse/kandidat/matematisk-modellering-og-computing</a>
Grundlæggende indhold	Uddannelsen foregår på engelsk og giver viden om anvendt matematik og matematisk modellering, der undervises i moderne computerudstyr og håndtering af store datamængder.
Elementer af data science i uddannelsens profil	Uddannelsen er blandt de mere tekniske uddannelser og giver mulighed for at vælge fokusområder inden for flere forskellige tekniske områder. Blandt de mest relevante fokusområder er "machine learning and signal processing" og "industrial and applied statistics", hvor begge fokusområder har fag, der indeholder elementer relateret til data science. <p>Fag der kan indgå i en profil inden for området "machine learning and signal processing":</p> <ul style="list-style-type: none"> <li>• Advanced Machine Learning (5 ECTS)</li> <li>• Computational tools for Data Science (5 ECTS)</li> <li>• Algorithms for Massive Data Sets (7,5 ECTS)</li> <li>• Model-based machine learning (5 ECTS)</li> <li>• Computational Data Analysis (5 ECTS)</li> </ul> <p>Fag der kan indgå i en profil inden for området "industrial and applied statistics":</p> <ul style="list-style-type: none"> <li>• Computational Data Analysis (5 ECTS)</li> <li>• Advanced dataanalysis and statistical modelling (5 ECTS)</li> <li>• Statistical quality control (5 ECTS)</li> <li>• Flere statistiske fag (alle 5 ECTS)</li> </ul>
Oplysninger relateret til efterspørgsel	Optag (2017): 116

	Antal dimittender (2016): 74 Dimittendledighed (2016 k3): 2 %
Uddannelsens titel	<b>Software Engineering</b>
Niveau	Kandidatuddannelse (civilingeniøruddannelse) på SDU
Link	<a href="https://www.ug.dk/uddannelser/universitetsuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingeniøruddannelser/software-engineering">https://www.ug.dk/uddannelser/universitetsuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/civilingeniøruddannelser/software-engineering</a> <a href="https://www.sdu.dk/da/Uddannelse/Kandidat/SoftwareEngineering">https://www.sdu.dk/da/Uddannelse/Kandidat/SoftwareEngineering</a>
Grundlæggende indhold	Uddannelsen bygger på forståelsen af software, samt hvordan man kan udvikle og forbedre nye og eksisterende softwareprodukter og -tjenester. Der opnås ligeledes kompetencer i implementering af it samt viden om styrker og svagheder ved forskellige softwareløsninger i forhold til virksomheder og brugere.
Elementer af data science i uddannelsens profil	Uddannelsen udgøres af fem fagområder, hvoraf et af områderne er datavidenskab. Datavidenskab handler om udvinding af viden fra data. Fokus er især på indsamling, filtrering, behandling, dannelse og distribution af data. Der undervises i relevante fag, der giver den studerende viden inden for maskinindlæring, algoritmer, statistik, databaser og visualisering.  Uddannelsen indeholder desuden et fag i Data Science (10 ECTS), som er obligatorisk for alle. Foruden dette fag indeholder de resterende fire fagområder ikke nævneværdige elementer ift. data science, da fokus hovedsageligt er på udvikling af software mm.
Oplysninger relateret til efterspørgsel	Optag (2017): 42  Antal dimittender (2016): 11  Dimittendledighed (2016 k3): 6 %
Uddannelsens titel	<b>Erhvervsøkonomi – Business Intelligence</b>
Niveau	Kandidatuddannelse på AU
Link	<a href="https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/samfundsvidenskabeligekandidatuddannelser/erhvervsøkonomi/erhvervsøkonomi-candmerc">https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/samfundsvidenskabeligekandidatuddannelser/erhvervsøkonomi/erhvervsøkonomi-candmerc</a> <a href="http://kandidat.au.dk/businessintelligence/">http://kandidat.au.dk/businessintelligence/</a>
Grundlæggende indhold	Uddannelsen er engelsksproget og omhandler hvordan man kan forbedre beslutningstagningen i virksomheder gennem analyse af data. Der opnås kompetencer til at arbejde med bl.a. planlægning og implementering af business intelligence løsninger.
Elementer af data science i uddannelsens profil	Uddannelsen er ikke blandt de mest tekniske i denne kortlægning. Til gengæld giver den indsigt i, hvordan man opsætter data til at være relevante i beslutningstagningen i virksomheder. Det kræver både evner i at indsamle og behandle de store mængder data, som i dag produceres af virksomheder. På uddannelsen er der fag, der relaterer sig til data science: <ul style="list-style-type: none"> <li>• R and SQL for Business Analytics (5 ECTS)</li> <li>• Data Warehousing (5 ECTS)</li> <li>• Data Mining for Business Decisions (10 ECTS).</li> </ul> Derudover består uddannelsens tredje semester af valgfag (30 ECTS) på AU eller andre universiteter, som man f.eks. kunne benytte til at supplere med de mere tekniske fag, hvilket der er rig mulighed for.
Oplysninger relateret til efterspørgsel	Optag (2017): <i>Ikke oplyst</i>  Antal dimittender (2016): <i>Ikke oplyst</i>  Dimittendledighed (2016 k3): <i>Ikke oplyst</i>
Uddannelsens titel	<b>Softwareudvikling (design)</b>
Niveau	Kandidatuddannelse på ITU
Link	<a href="https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/ituddannelser/softwareudvikling-design">https://www.ug.dk/uddannelser/bachelorogkandidatuddannelser/kandidatuddannelser/tekniskvidenskabeligekandidatuddannelser/ituddannelser/softwareudvikling-design</a> <a href="https://www.itu.dk/uddannelser/kandidatuddannelser/softwareudvikling-design">https://www.itu.dk/uddannelser/kandidatuddannelser/softwareudvikling-design</a>
Grundlæggende indhold	Uddannelsen er teknisk og handler om at designe og programmere software, herunder udnyttelse af programmeringssprog, operative systemer, distribuerede systemer, databasesystemer og modellering. Uddannelsen

	indeholder klassiske datalogiske kurser i algoritmer, datastrukturer og databaser, som suppleres af kurser i organisationsteori, forretningsprocesser og ledelse.
Elementer af data science i uddannelsens profil	<p>På kandidatuddannelsen i Softwareudvikling er der bl.a. mulighed for en specialisering inden for "Business Analytics". Denne specialisering giver en dybdegående viden om dataanalyse og giver mulighed for bl.a. at designe databasesystemer til at analysere Big Data og udføre data mining ved hjælp af klassifikations- og forudsigelsesalgoritmer. Specialiseringen inden for "Business Analytics" indeholder derfor også fag, hvor der indgår nogle af de samme elementer, som inden for data science:</p> <ul style="list-style-type: none"> <li>• Introduction to Database Design (obligatorisk) (7,5 ECTS)</li> <li>• Algorithms and Data Structures (obligatorisk) (7,5 ECTS)</li> <li>• Intelligent Systems Programming (7,5 ECTS)</li> <li>• Big Data Management (7,5 ECTS)</li> <li>• Data Mining (7,5 ECTS)</li> </ul>
Oplysninger relateret til efterspørgsel	<p>Optag (2017): 161</p> <p>Antal dimittender (2016): 107</p> <p>Dimittendledighed (2016 k3): 2 %</p>
Uddannelsens titel	<b>Datavidenskab (Data Science)</b>
Niveau	Kandidatuddannelse på AU og AAU ( <b>foreløbig godkendelse foreligger</b> )
Link	<a href="http://pkf.ufm.dk/flows/3704d145882a4305254cb3e2d11dcde0">http://pkf.ufm.dk/flows/3704d145882a4305254cb3e2d11dcde0</a>
Grundlæggende indhold	Kandidatuddannelsen i datavidenskab giver en kombination af kompetencer fra flere fagområder: Dataopsamling, dataprocessering, datahåndtering, datamining, og maskinlæring fra det datalogiske område samt avanceret statistik fra det matematiske område
Elementer af data science i uddannelsens profil	<p>På AAU indeholder kandidatuddannelsen i Datavidenskab et betydeligt overlap med de elementer, som forventes af indgå i en kandidatuddannelse i Data Science på ITU. Uddannelsen i Datavidenskab bærer præg af valgfrihed, som er indbygget i de tre første semestre. Fokus er på håndtering og beregning med avancerede og komplekse datatyper. Derudover opnås kompetencer inden for machine learning og intelligent dataanalyse, infrastruktur til Big Data, avanceret statistiske dataanalyser, samt forståelse for at udføre dataanalyser indenfor forretning, industri samt ingeniør- og sundhedsvidenskab.</p> <p>På AU er ovenstående beskrivelse også meget passende, men selve opbygningen af uddannelsen adskiller sig fra AAU's. På AU er der et par obligatoriske kurser, hhv. Advanced Statistical Learning og Large Scale Optimization (begge 10 ECTS). Derudover vælges et specialiseringsforløb, hvilket enten være en datalogisk/datasikkerhedsorienteret specialisering, en statistisk orienteret specialisering, en teknisk orienteret specialisering eller endeligt en økonomisk og forretningsforståelsesmæssig specialisering. Rammen for specialiseringen udgør 30-60 ECTS, hvilket dermed bestemmer omfanget af kandidatspecialet.</p>
Oplysninger relateret til efterspørgsel	<p>Optag (2017): <i>Ikke oplyst</i></p> <p>Antal dimittender (2016): <i>Ikke oplyst</i></p> <p>Dimittendledighed (2016 k3): <i>Ikke oplyst</i></p>
Uddannelsens titel	<b>Social Datavidenskab</b>
Niveau	Kandidatuddannelse på KU ( <b>ansøgning indsendt</b> )
Link	<a href="http://pkf.ufm.dk/flows/3704d145882a4305254cb3e2d13b4300">http://pkf.ufm.dk/flows/3704d145882a4305254cb3e2d13b4300</a>
Grundlæggende indhold	Uddannelsen kombinerer samfundsvidenskabelig faglighed med datavidenskabelige færdigheder
Elementer af data science i uddannelsens profil	<p>Uddannelsen fokus på at kombinere både samfundsvidenskab og datavidenskab giver en bred profil med både stærke analytiske og teknologiske kompetencer.</p> <p>Færdigheder indenfor programmering, machine learning og natural language processing (NLP) er blot en del af undervisningen inden for datavidenskab. Derudover opnås samfundsvidenskabelige færdigheder som en kombination af både kvantitative og kvalitative empiriske metoder.</p>
Oplysninger relateret til efterspørgsel	Optag (2017): <i>Ikke oplyst</i>

	Antal dimittender (2016): <i>Ikke oplyst</i>
	Dimittendledighed (2016 k3): <i>Ikke oplyst</i>

**De ekskluderede danske uddannelser:**

- *Computerteknologi*
- *Software*
- *Statistik*
- *Matematik*
- *Matematik-teknologi*
- *Matematik-økonomi*
- *Kommunikation og it*
- *It, kommunikation og organisation*
- *Informationsteknologi*

# Analysis of the Danish Data Science Job Market

Barbara Plank  
Department of Computer Science  
bplank@itu.dk

IT UNIVERSITY OF COPENHAGEN

Technical report, December 14, 2018

## Abstract

*This study analyzes the need of data scientists in Denmark on the basis of a 2018 job post analysis ranging over data from a 2.5 month period. Our analysis provides strong evidence for a high demand for data scientists in Denmark, providing additional insights on geographical dispersion, language considerations and types of sectors.*

## 1 Introduction

Data scientists are in high demand globally and demand outpaces availability [1, 2, 3]. In this report, we seek to provide evidence for the high demand of data scientists in Denmark across three dimensions, by providing empirical results of a job posting analysis.

Specifically, we provide empirical evidence for:

- i) the immediate open job market for data scientists in Denmark with respect to its *geographical distribution*,
- ii) a quantification of job postings for data scientists in Denmark in respect to *language considerations* and
- iii) the demand of types of *employers* (public/private etc) seeking data science competencies in Denmark.

Our study on a dataset of job postings collected over a period of 2.5 month shows that data sci-

ence competencies are of high demand locally. This complements recent findings from an earlier empirical study in Denmark [4] and evidence at a global scale, e.g., the LinkedIn workforce report [1].

**Contributions** Our study shows:

- Data scientists are needed throughout Denmark; the largest proportion of job are sought after in the main capital area (40% in the municipality of Copenhagen).
- The job market is looking for English-speaking employees, as evidenced by the fact that 66.3% of the job postings are in English.
- Data science competencies are needed in many sectors, i.e., the public sector, companies and universities.

## 2 Method

### 2.1 Data

We analyze job postings collected in the time frame of 2.5 months, in particular, from September 20 to December 3, 2018. We follow the setup of [4] and use their data-science-related terms (e.g., ‘data science’, ‘predictive analytics’, ‘machine learning’, ‘data mining’, ‘data engineer’ details in the

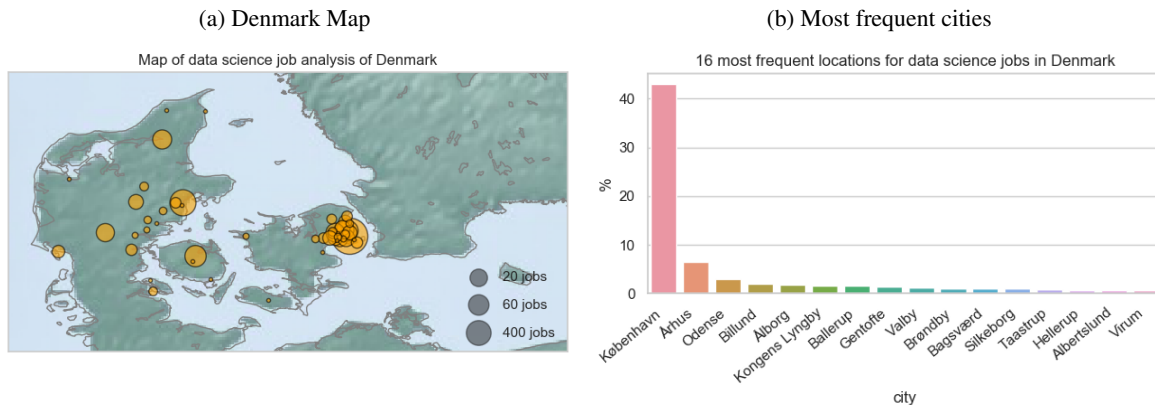


Figure 1: Locations of data science jobs in Denmark.

reference) to query for jobs posted in Denmark.<sup>1</sup>

After duplicate removal (outlined in Section 2.2), the resulting data set consists of 1,390 job advertisements. For each job we have access to job title, company, language of the post (as provided by the data provider) and a short job snippet. We also scrape the entire job posting for further analysis, and extract the job description section of the job post (using BeautifulSoup for localization of the job description and for boilerplate removal).

An example job posting is provided below.<sup>2</sup>

```
Lead Data Scientist for R&D Division
Company: KMD
Location: Ballerup
Snippet:
Do you want to help making the world a
better place, using machine learning
and data mining? You can develop
data science knowledge utilizing
multiple sources...
```

## 2.2 Preprocessing

The first step consists of the removal of duplicates or near-duplicates. To do so, we use a simple bag-of-word analysis and compare similarity of the entire job postings using cosine similarity on a term

<sup>1</sup>From <http://dk.indeed.com/>.

<sup>2</sup>The data is available at: <https://www.itu.dk/people/bapl/dsjobsdk/>

frequency (TF)-inverse document frequency (TF-IDF) representation for the documents. By manual inspection, we set the similarity threshold  $t$  to  $t = 0.7$ . This resulted in the final data set of  $n = 1,390$  job posts (out of 1,900 job posts collected initially), where highly overlapping or near-duplicates were removed.

We further enrich each job post by geolocation information (longitude and latitude).

## Results

The job analysis provides three take-aways.

**Geographical distribution** As illustrated in Figure 1 (a) and (b), there is a high demand for jobs in the main capital region (53% of the job posts are around the greater Copenhagen area<sup>3</sup>). In fact, the highest density of jobs can be found in København, followed by Jutland (Aarhus) and Funen (Odense).

**Language considerations** The analysis provides strong empirical evidence that English is important for the IT job market in Data Science. As shown

<sup>3</sup>Taking into account København, Frederiksberg, Valby, Brøndby, Bagsværd, Taastrup, Gentofte, Hellerup, Albertslund, Virum and Kongens Lyngby.

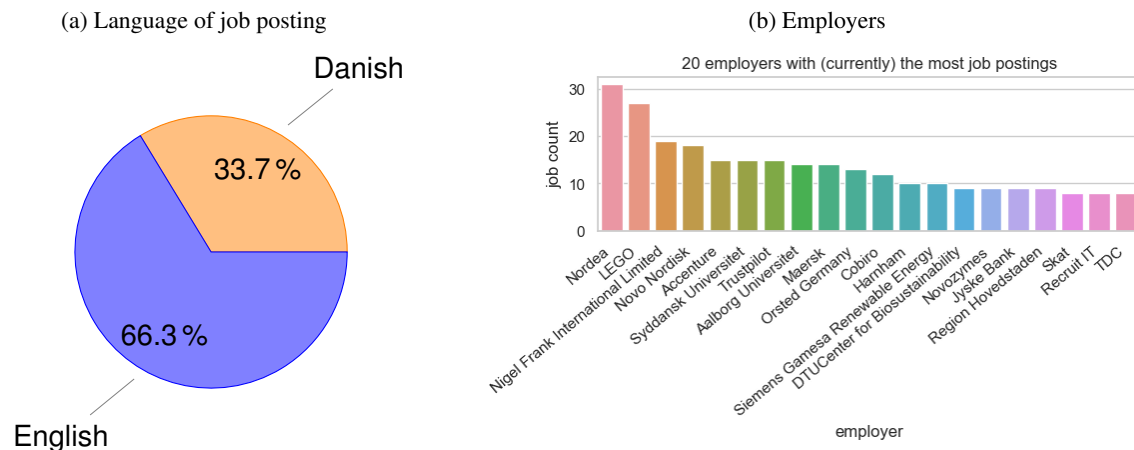


Figure 2: Job profile analysis: language of job posts, top employers (from our sample of 1,390 jobs).

in Figure 2 (a), the majority of job posts (66.3%) has been published in English. Moreover, out of these 921 English job postings, a simple keyword search reveals that 66% of the jobs explicitly require English language skills.<sup>4</sup> This further supports a global market.

**Competencies across sectors** Figure 2 (b) shows the most frequent employers. Amongst the 20 employers that seek the most data science competencies in our sample, we can find the Danish public sector (e.g., Skat, Region Hovedstaden), the banking sector (e.g., Nordea Jyske Bank), international cooperations (LEGO, Trustpilot, Maersk, Novo Nordisk, Novozymes, TDC) and universities (Syddansk Universitet and Aalborg Universitet). This closer look further supports the need for data science competencies across industries, public agencies and universities, as has been identified at a global scale [1] and as shown here, is the case for Denmark as well.

<sup>4</sup>By searching for (in a case-insensitive manner) for the term “(in) English” in the full job descriptions, we found statements such as ‘Fluent in English’, ‘Excellent oral and written communication in English’, ‘academic level in Danish or English’, ‘possess good communications skills in English’.

## References

- [1] August LinkedIn Workforce Report: Data Science Skills are in High Demand Across Industries. <https://news.linkedin.com/2018/8/linkedin-workforce-report-august-2018>.
- [2] The Figure Eight 2017 Data Scientist Report. [https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower\\_DataScienceReport.pdf](https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower_DataScienceReport.pdf).
- [3] The Figure Eight 2018 Data Scientist Report. <https://www.figure-eight.com/figure-eight-2018-data-scientist-report>.
- [4] Natalie Schluter and Rasmus Pagh. An empirical analysis of the current Data Science Job Market in Denmark. Technical report, IT University of Denmark, 2016.

## Beskæftigelsesgrad for engelske og danske dimittender fra IT-Universitetet i perioden 2009-2014

Jane, 11-1-2019

Tabellerne viser for hvert af årene 2009-2014:

- Antal fuldførte kandidater.
- Beskæftigelsesgraden (pct.) opgjort som andelen af dage, hvor dimittenderne har været i arbejde i det andet år (12-23 måneder) efter de dimitterede.
- Dage med ledighed (pct.) er andelen af dage, hvor dimittenderne har været ledige i det andet år (12-23 måneder) efter de dimitterede.
- Dage uden for arbejdsstyrken (pct.) er andelen af dage, hvor dimittenderne har været uden for arbejdsstyrken i det andet år (12-23 måneder) efter de dimitterede.
- Dage udrejst af Danmark (pct.) er andelen af dage, hvor dimittenderne har været registreret som udrejst af Danmark i det andet år (12-23 måneder) efter de dimitterede.

De enkelte dimittenders dage i hver af de fire ovennævnte grupperinger er opgjort.

Hvis en dimittend har været i beskæftigelse i seks måneder, ledig i to måneder og udrejst af Danmark i fire måneder, vil de tilsvarende dage indgå i den samlede beregning af dimittendernes andel af dage i hver af grupperingerne.

Samlet set vil procentandelene af dage i hver af de fire grupperinger give 100 pct. for hvert af årene 2009-2014.

Nedenstående data er trukket fra UFM's DVH i tabellen 'ElevBesk' baseret på data fra Danmarks Statistik.

2014 er det seneste dimittend-år, hvor der er beskæftigelsesdata.

### Alle dimittender (engelsprogede og dansksprogede)

Dimittend-år	Antal fuldførte						Beskæftigelsesgrad: Dage i beskæftigelse (pct.)						Dage med ledighed (pct.)						Dage uden for arbejdsstyrken (pct.)						Dage udrejst af Danmark (pct.)					
	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014
72_Kandidatuddannelser på IT-Universitetet	207	200	243	280	332	358	74,9%	78,0%	75,1%	80,3%	74,4%	80,9%	8,4%	8,0%	11,3%	6,7%	10,4%	6,0%	7,5%	6,4%	6,4%	4,6%	6,4%	5,6%	9,2%	7,6%	7,2%	8,4%	8,9%	7,5%
6265_Softwareudvikling, cand.it.	44	54	53	59	64	86	71,8%	75,2%	71,0%	81,4%	74,8%	76,3%	9,3%	3,4%	14,2%	2,6%	5,5%	3,4%	6,9%	11,0%	7,2%	5,6%	8,7%	6,2%	11,9%	10,3%	7,5%	10,4%	11,0%	14,1%

### Engelsprogede<sup>\*)</sup> (ikke-dansk og ikke-nordisk statsborgerskab) dimittender

Dimittend-år	Antal fuldførte						Beskæftigelsesgrad: Dage i beskæftigelse (pct.)						Dage med ledighed (pct.)						Dage uden for arbejdsstyrken (pct.)						Dage udrejst af Danmark (pct.)					
	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014
72_Kandidatuddannelser på IT-Universitetet	44	30	47	46	59	52	59,1%	63,9%	51,8%	64,4%	48,1%	51,6%	14,9%	11,0%	16,9%	5,6%	11,0%	4,3%	11,8%	12,1%	12,7%	8,9%	15,1%	6,5%	14,2%	13,0%	18,6%	21,1%	25,7%	37,6%
6265_Softwareudvikling, cand.it.	27	22	24	24	33	33	66,4%	76,2%	51,2%	73,5%	59,4%	56,2%	14,1%	4,4%	26,5%	3,8%	7,1%	3,9%	11,2%	11,1%	13,9%	9,6%	16,2%	9,1%	8,3%	8,3%	8,3%	13,1%	17,4%	30,8%

<sup>\*) Engelsksprogede i tabellen ovenfor ligger meget tæt på definitionen af engelsksprogede i UFM's notat 'Reduktion i tilgangen til engelsksprogede uddannelser' dateret 22-8-2018. I dette notat defineres engelsksprogede som 'En engelsksproget studerende eller dimittend er en person, der ikke er dansk eller øvrig nordisk statsborger og som ikke har en tidligere uddannelse i Danmark (grundskole eller ungdomsuddannelse)'. I tabellen ovenfor indgår dimittender, der ikke er danske eller nordiske statsborgere uden hensyntagen til deres tidligere uddannelse, da denne uddannelsesparameter ikke findes i UFM-datasættet. Det er UFM's aktuelle rådgivning at anvende dette tilnærmede datasæt og ved sammenligninger ses, at der er meget tæt overensstemmelse med UFM's data om engelsksprogede i notatet. Dette ses fx ved at sammenligne tallene i de grå celler i tabellen ovenfor med tabel 7 i notatet (se kopi af tabel 7 i fanebladet i dette regneark).</sup>

### Dansksprogede (dansk og nordisk statsborgerskab) dimittender

Dimittend-år	Antal fuldførte						Beskæftigelsesgrad: Dage i beskæftigelse (pct.)						Dage med ledighed (pct.)						Dage uden for arbejdsstyrken (pct.)						Dage udrejst af Danmark (pct.)					
	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014	2009	2010	2011	2012	2013	2014
72_Kandidatuddannelser på IT-Universitetet	163	170	196	234	273	306	79,1%	80,4%	80,6%	83,5%	80,0%	85,9%	6,7%	7,5%	10,0%	7,0%	10,3%	6,3%	6,3%	5,4%	4,9%	3,7%	4,5%	5,4%	7,9%	6,7%	4,4%	5,8%	5,2%	2,4%
6265_Softwareudvikling, cand.it.	17	32	29	35	31	53	80,4%	74,6%	87,4%	86,9%	91,3%	88,8%	1,7%	2,8%	4,1%	1,8%	3,8%	3,1%	0,2%	10,9%	1,7%	2,8%	0,7%	4,3%	17,6%	11,7%	6,9%	8,6%	4,2%	3,8%

## Arbejdsmarkedsstatus for dansk- og engelsprogede dimittender fra IT-Universitetet i 2012-2014

Analyseenheden, 18-12-2018

Tabellerne viser 'dage beskæftiget i pct. af dage i arbejdsstyrken', hvor:

- Dage beskæftiget i procent af dage i arbejdsstyrken er opgjort for fuldførte 12-23 måneder efter fuldførelsestidspunkt.
- Dage i arbejdsstyrken er summen af dage, hvor personen er
- Udregningen af 'dage beskæftiget i pct. af dage i arbejdsstyrken' =  $\text{dage\_besk\_sum}/(\text{dage\_besk\_sum}+\text{dage\_ledig\_sum})\cdot 100$  (enheden er dage i primær tilstand).

Nedenstående data er trukket fra UFM's DVH i tabellen 'ElevBesk' baseret på data fra Danmarks Statistik.

2014 er det seneste dimittend-år, hvor der er beskæftigelsesdata.

### Alle dimittender (engelsprogede og dansksprogede)

Dimittend-år	2012	2013	2014
<b>72_Kandidatuddannelser på IT-Universitetet</b>	<b>92,3 %</b>	<b>87,7 %</b>	<b>93,1 %</b>
6265_Softwareudvikling, cand.it.	96,9 %	93,1 %	95,8 %

### Engelsprogede (ikke-dansk og ikke-nordisk statsborgerskab) dimittender

Dimittend-år	2012	2013	2014
<b>72_Kandidatuddannelser på IT-Universitetet</b>	<b>92,0 %</b>	<b>81,4 %</b>	<b>92,4 %</b>
6265_Softwareudvikling, cand.it.	95,1 %	89,3 %	93,5 %

### Dansksprogede (dansk og nordisk statsborgerskab) dimittender

Dimittend-år	2012	2013	2014
<b>72_Kandidatuddannelser på IT-Universitetet</b>	<b>92,3 %</b>	<b>88,6 %</b>	<b>93,2 %</b>
6265_Softwareudvikling, cand.it.	98,0 %	96,0 %	96,7 %

# Reduktion i tilgangen til engelsksprogede uddannelser - Notat fra UFM

22-08-2018

## Definition af engelsksprogede studerende og dimittender

En engelsksproget studerende eller dimittend er en person, der ikke er dansk eller øvrig nordisk statsborger og som ikke har en tidligere uddannelse i Danmark (grundskole eller ungdomsuddannelse).

**Tabel 6**

Arbejdsmarkedsstatus for engelsk- og dansksprogede dimittender (2012-2014), 2 år efter dimission (pct.) samt engelsk- og dansksprogede dimittender fordelt på hovedområder, antal og andel (pct.)

	Beskæftigelse engelsksprogede	Beskæftigelse dansksprogede	Ledige engelsksprogede	Ledige dansksprogede	Uden for arbejdsstyrken engelsksprogede	Uden for arbejdsstyrken dansksprogede	Udrejste engelsksprogede	Udrejste dansksprogede	Antal engelsksprogede dimittender	Antal dansksprogede dimittender	Andel engelsksprogede dimittender
<b>2012</b>											
Natur, bach.	*	74	*	1	*	0	*	25		78	0
Natur, kandidat	61	84	6	7	10	4	23	6	42	238	15
<b>2013</b>											
Natur, bach.	*	55	*	14	*	6	*	25	*	133	*
Natur, kandidat	47	80	11	10	15	4	26	5	58	274	17
<b>2014</b>											
Natur, bach.	*	79	*	5	*	3	100	12	*	135	*
Natur, kandidat	49	86	4	6	7	5	40	2	49	309	14

## Analyseenhedens udregninger baseret på tabel 6 (til venstre)

Beskæftigelsesgrad Engelsksprogede	Beskæftigelsesgrad dansksprogede	Beskæftigelsesgrad alle dimittender
	98,7	
91,0	92,3	92,1
	79,7	
81,0	88,9	87,5
	94,0	
92,5	93,5	93,3

Anm: \* Antallet af engelsksprogede dimittender er under 10, og der kan derfor ikke beregnes beskæftigelsesgrader.

Det er ikke i tabellerne på uddannelsesniveau muligt at vise beskæftigelsesgrader og tilgang i samme tabel, da der er en tidsmæssig forskydning i de to opgørelser, og der i denne periode har været en kodeoprydning i uddannelseskoderne, som betyder, at man ikke kan sammenkøre opgørelserne. Derfor fremstilles beskæftigelsestal og tilgangs/bestandstal i forskellige tabeller.

**Tabel 7**

Beskæftigelsesstatus blandt engelsksprogede dimittender (2014) to år efter dimission, samt antal engelsksprogede dimittender, fordelt på uddannelser, pct.

Uddannelse	Beskæftigelsesgrad**	Antal engelsksprogede dimittender	Antal dimittender i alt	Antal beskæftigede engelsksprogede
Digital design og interaktive teknologier, bach.	*	*		
Global virksomhedsinformatik, bach.	*	*	39	
Softwareudvikling, bach.	*	*	39	
Digital design og interaktive teknologier, kand.2år	*	*	*	
Digital design og kommunikation, kand.2år	*	*	130	
Digital innovation og management, kand.2år	*	*	89	
Softwareudvikling, kand.2år	55	32	86	18
Spil, kand.2år	29	15	52	

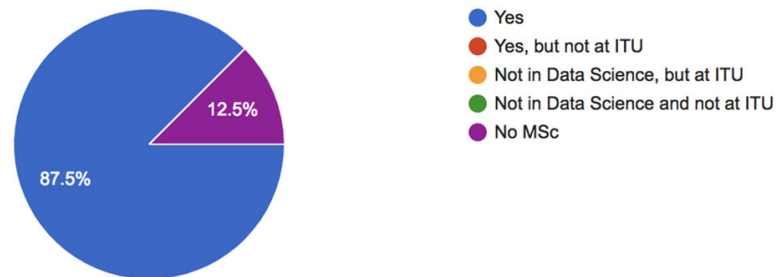
\* Antallet af engelsksprogede dimittender er lavere end 10, og der kan derfor ikke beregnes beskæftigelsesgrader.

\*\*I beregningen af beskæftigelsesgraderne er personer, der er gået i videre uddannelse ekskluderet.

Kilde: Uddannelses- og forskningsministeriet pba. data fra Danmarks Statistik

### Would you do an MSc in Data Science at ITU if there were one?

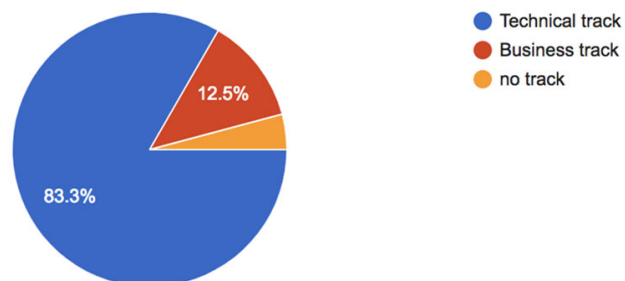
24 responses



Survey 2nd year DS BSc **students** at ITU (September 2018)

### Which track did you intend on taking during your BSc?

24 responses



Survey 2nd year DS BSc **students** at ITU (September 2018)

Students Affairs & Programmes  
Mette Holm Smith  
Direct phone: +45 7218 5087  
E-mail: [meho@itu.dk](mailto:meho@itu.dk)  
November 7, 2018

# Minutes approved employers Panel meeting regarding new Msc in Data Science Wednesday 24 October 2018 9:00-12:30 room 3A20

## Participants:

### Employers:

Bodil Biering (Chairman) (Blackwood Seven), Kaare Brandt Petersen (Implement Consulting Group), Søren Ilsøe (Deloitte), Jakob Elming (Doolittle Translation/Danish Defense), Kaare Danielsen (Jobindex), Camilla Kerlaug (Spenderlog), Jens Lund (former Nordea), Lars Kjeldgaard (Udviklings og forenklingsstyrelsen SKAT), Florian Laws (Tradeshift).

### IT University:

Associate professor Barbara Plank (Head of MSc Data Science Program Development), Associate Professor Natalie Schluter, Associate Professor Pinar Tözün, Associate Professor Troels Bjerre Lund, Professor Peter Sestoft (Head of CS Department), Associate Professor Stine Gotved (Head of Studies), Signe Vægter Rasmussen (Project Manager), Mette Holm Smith (Programme Coordinator, minutes)

### Agenda:

9.00-9.15 Welcome and Round of introductions (focusing on participants' careers and its relation to data science) (Bodil)

9.15-9.25 Purpose and outcome of the meeting (Bodil)

9.25-9.40 Presentation of the program draft including open questions to the design (appendix) (Barbara)

9.40-10.20

Discussion of open questions in relation to the design:

1. Name a couple of job profiles of a data scientist in your company
2. What kind of training / courses do your employees in data science positions ask for?
3. What megatrends do you see in industry?
4. What is Business Intelligence for you? What about business is important to know for a data scientist to interact with stakeholders?
5. Are there tools/software systems that will be used in the long run?
6. Ethics: separate course (7.5 ECTS)? E.g., "Ethics & Privacy" (Social Implications, Anonymization, Diff. Privacy)

10.20-10.30 Break

10.30-10.40 Presentation of job profiles (appendix) (Barbara)

10.40-11.30 Discussion of job profiles

11.30-12.00 Other comments from the panel and wrap up: conclusions from the panel

12.00-12.30 Lunch

## **Welcome and round of introduction, purpose and outcome of the meeting and definition of a data scientist**

Chairman of the employers group Bodil Biering from Blackwood Seven welcomed the employers to the meeting about the development of a MSc Data Science program at ITU. The panel has been put together in order to have a broad representation from both private and public companies and organizations.

After a round of introduction Bodil briefly went over the purpose of the meeting: to get the employers' input on data science competencies in demand in your company or organization as well as the outcome of meeting: a set of recommendations and reflections from employers on the program's fit with industry needs. (See attached slides from the meeting).

Bodil finally looked into the definition of a Data Scientist in the field between computer science, mathematics and statistics and business/domain expertise and presented the definition in ITU's program design. How does this relate to a fast moving world with creation of new jobs? And how do we prepare students for a rapid change? In the Cynefin Framework the data scientist would be placed between the complicated and complex domains. (see attached slides)

## **Presentation of suggested program design**

Associate Professor Barbara Plank (Head of development group) presented the suggested program design and the motivation and choices made for the program.

There is a huge gap in demand for data science skills as shown in the lead example from US in the LinkedIn workforce report from August 2018. The skills are in demand across industries and not only in certain industries. It is a broad demand.

An analysis of harvested job profiles over a month (Sep-Oct) set up by ITU shows 1100 relevant data science jobs with 42 % of the jobs located in the municipality of Copenhagen. Out of the 1100 job postings, 65 % were posted in English.

Data science is a broad field and the BSc program could have many different academic profiles within the ITU academic field of content between arts, business and natural science. The current enrolled BSc students have to choose either a business or technical track on their BSc program and 83.3 % out of 24 2<sup>nd</sup> year students have indicated that they will choose the technical track.

This supports the working group's choice of making a technical program placed within the technical angle of the ITU triangle but ITU would like to have the employers' input on this.

Barbara presented the structure of the program design (see slides) and invited the employers to give input and comments to the program design based on questions deriving from discussions and struggles when designing the program draft:

1. Name a couple of job profiles of a data scientist in your company
2. What kind of training / courses do your employees in data science positions ask for?
3. What megatrends do you see in industry?
4. What is Business Intelligence for you?
5. What about business is important to know for a data scientist to interact with stakeholders?
6. Are there tools/software systems that will be used in the long run?
7. Ethics: separate course (7.5 ECTS)? E.g., "Ethics & Privacy" (Social Implications, Anonymization, Diff. Privacy)

The input and comments were based on the questions and have been structures below:

### **Name a couple of job profiles of a data scientist in your company**

The panel first discussed whether they would prefer a broad profile with some knowledge within many areas or a more specialized profile with a good ability to learn what they have not touched upon in their education.

It also depends of the size of the company. A large company can afford to hire more graduates who are specialized in different areas whereas a start-up will have to look for the broad graduate who can do all thing.

However, as it is today no graduate has the whole DS competence profile.

The panel emphasized that a data scientist from ITU should have good and solid technical skills with a strong foundation in mathematics and statistics.

### **Applied skills**

The panel finds that it is very good that the program ensures that the theoretical skills are also combined with practical application. The applied approach is very useful for employers and companies when hiring graduates and the “Profound and Applied” should be the ITU MSc in Data Science profile in relation to other Universities MSc Data Science programs.

In relation to the profile description, the panel discussed which skills and competencies are important for a graduate to have.

### **Technical skills and competencies**

The panel found that besides the good and solid data science technical skills graduates should also have good software engineering skills in order to write high quality code and be able to bring the program from their desktop and into the business.

It includes skills like systematic testing, design patterns, tools and processes for configuration management, continuous deployment, and basic software architecture.

### **Business skills and competencies**

The panel found that “Business intelligence” sounds a bit archaic and suggested to not use that title for a course but maybe use a title like “Business Analytics” or “Insights and discovery”. In general, the panel agreed that compared to technical skills business understanding is a skill that a graduate can learn better within the first years in a company.

However, they found that graduates should have some kind of basic business understanding and understanding of organizations in order to be able to act as “translators” and bridge between different stakeholders.

ITU could try to integrate the business knowledge into the technical courses. In companies, the technical side has to build a bridge to business and in order to do so the graduates have to understand that what they do both relates to and have impact on business.

It could also be having students do a proper assessment of resources needed for a project. Every time you make projects include the business e.g. GDPR, resistance to change and maybe teach change management to get a business organizational understanding.

Based on experience with teaching technical employees a panel member suggested to teach what is important in business by taking all business strategy and making it simple. Cook it down to 10 hours about what is important.

### **Content already learned on the bachelor**

During the meeting, it was highlighted that Data Science is broad, but the current design has been developed to provide a MSc program designed to build on top of the ITU BSc in Data Science.

The employers commented that if the MSc program relies on content taught at the ITU BSc program it is important to address the prerequisites for students coming with a BSc background from outside ITU.

ITU answered that this was taken into account and that the working group has thoroughly looked into planned and existing programs at AAU, AU, CBS, KU and DTU within the Data Science area.

### **Ethics: separate course (7.5 ECTS)? E.g., “Ethics & Privacy” (Social Implications, Anonymization, Diff. Privacy)**

Barbara asked the panel how ethics should be implemented in the program in relation to importance. Should it be a whole course on ethics? Alternatively, should it be combined as part of another course?

On the bachelor DS there is on the fifth semester a course on Security and Privacy that will be taught by a lawyer and two other courses on the bachelor (Second Year Project and Reflections on DS) also touch upon related content.

In general, the panel found that it is important to include ethics in the program. The panel discussed the frames for what to teach in ethics in relation to legal framework. Some argued that ethics is a very normative subject and that they did not expect the university to teach things outside the legal aspect. Others found that it is important to teach ethics and teach the students to consider the consequences and if the graduates have not learned to reflect it can be legal but out of line.

The panel did not end up with a clear recommendation whether to have a dedicated course on ethics or only include it in other courses. However, they found that it is relevant to include ethics and behavior in all courses and projects on data handling in order to have an awareness, in particular to raise awareness of ethical implication of data processing within Data Science and stipulate reflection (not prescription).

### **Presentation and discussion of Job profiles**

Barbara presented the job profiles

- Profile A [functionality]: data-driven service enhancement and development of new services (focus on functionality)
- Profile B [analytics / integration]:
  - Subprofile 1 [analytics]: analytics/insights
  - Subprofile 2 [integration/knowledge transfer]: data acquisition, data integration, data transfer

Bodil introduced to related profiles: data operator and data owner. Subprofile 2 is what nowadays is often called 'data operator'.

What became apparent in the discussion is that the current dichotomy is not seen to be strict. For example, a data scientist developing a Machine Learning model could be considered Profile A (if it ends as a new feature for a product), while it could also be Profile B/1 (if it is more insight-driven, and less product specific). This also depends on the size of the company, where there could be dedicated separate roles, while there is also the possibility that a data scientist moves between profile A and B. A suggestion was made to rename the two profiles A and B as "product-feature driven" versus "insight-driven", respectively.

Independent of the naming, a data-scientist can be responsible to generate a POC (proof-of-concept). The "one-time solution" aspect was more linked to profile B/1.

### **Wrap up and conclusions**

Bodil thanked the panel for a very productive discussion and much relevant input for continuing the development of the program, which the ITU academic working group will use in the development of a final draft that will be presented to the panel in December in a written hearing.

## Uddrag af:

### Referat af møde i Aftagerpanel for Software

Tirsdag den 8. januar 2019 16:00-19:30 på IT Universitetet, Rued Langgaards Vej 7, 2300 Kbh. S.

#### Ad 3. Presentation of the MSc programme in Data Science, which ITU prepares to apply for.

Barbara præsenterede forslaget:

- Baggrund for at oprette uddannelsen:
  - Der starter lignende kandidatuddannelser op i Århus og Aalborg, men der er også stor efterspørgsel efter kandidater i Københavns området.
  - En overbygning på BSc DS, så disse bachelorer får mulighed for at fortsætte på en uddannelse i direkte forlængelse af deres bachelor.

En undersøgelse blandt 2. års BSc DS studerende viser at størstedelen af de studerende ønsker at fortsætte på en DS kandidatuddannelse (87,5% af 2. års studerende (svarprocent 57%)

BSc DS har to tracks, Business og Teknisk. De fleste studerende (83,3% ønsker det tekniske track.

- Disse to tilbagemeldinger præger designet af kandidatuddannelsen, der er teknisk fokuseret, men også har vægt på det anvendelsesorienterede.
- Målgruppen er ansøgere med en BSc i DS eller lignende. Ansøgerne skal have en solid matematisk og teknisk baggrund.
- Jobprofil: to profiler (se mere information i Programme Design for uddannelsen – sendt ud som bilag):
  - A: Funktionalitet
  - B: Analyse/Integration.

#### Kommentarer fra panelet og evt. svar fra Barbara:

- Balancegangen mellem teknik og forretningsforståelse i forslaget: Panelet drøftede at forretningsforståelse er vigtig at give de studerede, så de kan anvende deres kompetencer på jobmarkedet. Men på den anden side kommer den dybe forretningsforståelse efter uddannelsen i forbindelse med job i virksomhederne. Panelet er glad for kurset Data Science in Production, der netop integrerer data science og forretningsforståelse. Barbara oplyser, at man netop ønsker at integrere virksomheder i uddannelsen i tråd med ITUs ånd gennem projekter, men at den dybe forretningsforståelse er en kompetence de studerende bedst lærer på jobmarkedet efter afslutningen af deres studier.
- Panelet vurderer, at jobprofilerne rammer jobmarkedets behov.
- Panelet spørger til hvad begrebet "Etik" omfatter, og om security er integreret i dette begreb? Etik er overvejelser over sociale og etiske konsekvenser af at udvikle data systemer. Security er noget andet.

Panelet spørger, om der bør være et dedikeret kursus i security? Kandidaterne skal kende til security og have en bred tilgang til dette, men det er ikke dem som f.eks. skal have specialistviden til f.eks. at konstruere security systemer. Barbara nævner at der vil være et valgfag i security, som de studerende kan vælge.

## **Uddrag af:**

Referat af Executive-Level Employers' Panel møde, 23. oktober, 2018, kl. 16:00 – 19:00

### **4. Ny kandidatuddannelse i Data Science (orientering og drøftelse)**

Mads Tofte bød velkommen til Barbara Plank, der kort præsenterede sig selv og gav en status på arbejdet med at lave en ny kandidatuddannelse i Data Science på ITU. Arbejdet bygger bl.a. på en jobprofil-analyse og en undersøgelse blandt studerende. Hun viste slides med udkast til forslag til en kursusplan og modtog panel-medlemmernes kommentarer og besvarede spørgsmål vedrørende planens indhold. Panelet bifaldt, ud uddannelsesdesignet har sin tyngde i tekniske/matematiske emner og bygger oven på bacheloruddannelsen i Data Science; panelet fandt det også vigtigt, at de studerende stifter bekendtskab med etiske problemstillinger og lærer, hvad forretningsforståelse handler om. Carsten Gomard takkede på panelets vegne Barbara Plank for en meget interessant præsentation.

# Master in Data Science at IT University of Copenhagen

Barbara Plank, Natalie Schluter, Pinar Tözün, Troels Bjerre Lund  
December 4, 2018

## Overview

This document describes the design and implementation of a possible Master in Data Science (MSc DS) program at the IT University of Copenhagen. It is designed for students holding a BSc in Data Science or similar.

The proposed MSc Data Science is a 2 year program that aims to provide graduates with comprehensive advanced analytical and technical skills covering all aspects of handling, analyzing, and operationalizing data. In addition, graduates are trained in applying these skills in realistic settings, through projects that include interaction with domain experts.

### *Why study MSc DS at ITU?*

- **Profound & Applied:** The MSc DS combines theoretical foundations with practical applications. The MSc DS graduates will be equipped with computational and statistical skills to design sophisticated solutions to solve applied problems using data.
- **In high demand:** The demand for data science competencies grow in many areas of business and research. Data science “has become increasingly important across all industries” both across the globe [[LinkedIn Workforce Report, 2018](#)] and in Denmark [[Technical report, 2018](#)]. Most large companies have data scientists (e.g., in banking, insurance, logistics, pharma, retail).
- **Focused:** The program and research-based learning environment at ITU are specially focused on core Data Science. The MSc DS courses are specifically designed to deepen the technical knowledge of students in core Data Science areas. ITU is also becoming an important research hub for core Data Science in Denmark, providing students with a fertile study environment in which to learn about state-of-the-art Data Science in research and in practise.

### **Acquired Skills**

The MSc in Data Science focuses on the following skills:

<ul style="list-style-type: none"><li>● Machine learning</li><li>● Statistical Analysis</li><li>● Algorithm Design</li></ul>	<ul style="list-style-type: none"><li>● Data Visualization</li><li>● Mining and Exploring Data</li><li>● Data Science in Production</li></ul>	<ul style="list-style-type: none"><li>● Research Design</li><li>● Communicating Results</li><li>● Ethics</li></ul>
--	---	--

## Program Design

### Design considerations

- The program must have at least 10 ECTS electives, and a 30 ECTS Master's thesis.
- Progression from BSc of Data Science at ITU.
- Possibility of going abroad (in 3rd semester). To allow for this possibility, the obligatory interdisciplinary project course has been allocated in the first semester. Another reason for doing so is to give students a broader perspective in the beginning of the MSc. In the third semester they focus on their research project and advanced elective (locally or abroad).
- Data Science is a broad field. The proposed MSc in Data Science at ITU focuses on **strong technical competencies with profound theoretical foundations**, which is a direction supported by both the employers' group and a survey conducted among current BSc Data Science students (see attached [minutes of the employer meeting](#) for details). Therefore, the program was designed according to the three skill axes outlined above:

- Mathematical and algorithmic foundations (blue; 15 ECTS, first semester)
- Exploration of heterogeneous data, advanced visualization, involvement of stakeholders in project-based courses (green; 22.5 ECTS, first and second semester)
- Communication skills and ethical considerations (grey; this will part of the research project, with obligatory ethical reflection, and the Master's thesis)

The skills will be enriched in dedicated advanced electives where the student chooses a direction to specialize in, for example: Data Systems Design, Advanced Machine Learning or Natural Language Processing.

- **Job profile considerations:**<sup>1</sup>  
Data Science competencies are in high demand, both globally and in Denmark (see our [technical report](#)). In collaboration with the employer group and our job post analysis, we identified two current and future market needs in the Danish labor market that the MSc in Data Science seeks to address. Note that these are *competence profiles*, which do not strictly form a dichotomy. A data scientist might need competencies from one profile, or switch between them, depending on the task. This is particularly the case in larger corporations.
  - **Profile A** [functionality]: data-driven service enhancement and development of new services (enhancement of existing products)
  - **Profile B** [analytics / integration]:
    - Subprofile 1 [analytics]: analytics/insights, e.g., data owner (a data scientist who derives insights from the data)
    - Subprofile 2 [integration/knowledge transfer]: data acquisition, data integration, data transfer; e.g., data engineer

---

<sup>1</sup> Both profiles can also be a respective Research & Development position.

## Proposed course plan

<b>Semester 1</b>	Algorithm Design (7.5 ECTS)	Advanced Applied Statistics & Multivariate Calculus (7.5 ECTS)	Data in the Wild: Processing heterogeneous data (7.5 ECTS)	Interdisciplinary project (7.5 ECTS)
<b>Semester 2</b>	Elective (7.5 ECTS)	Elective (7.5 ECTS)	Data Science in Production (7.5 ECTS)	Advanced Data Visualization (7.5 ECTS)
<b>Semester 3</b>	Advanced elective (15 ECTS course)		Research Project incl. Ethics (15 ECTS)	
<b>Semester 4</b>	Master thesis (30 ECTS)			

### Proposed advanced electives (15 ECTS):

- Advanced Machine Learning (15 ECTS)
- Data Systems Design (15 ECTS)
- Advanced Algorithms (15 ECTS)
- Natural Language Processing and Deep Learning (15 ECTS)
- Robotics (15 ECTS)
- Network Science (15 ECTS)

### Other proposed electives (7.5 ECTS):

- Operations Research (7.5 ECTS)
- Computer Systems Performance (7.5 ECTS)
- Practical Concurrent and Parallel Programming (7.5 ECTS)
- Security (7.5 ECTS)

## Target group

The requirements for being admitted to this program are listed below:

- The applicant must hold a bachelor in Data Science or the like.
- The successful applicant is expected to have covered most of the following topics during his/her bachelor studies:
  - Linear algebra and optimization
  - Algorithms and data structures
  - Introduction to machine learning
  - Applied statistics
  - Introduction to programming (functional programming, object oriented programming)
  - Data management (introduction to databases)
  - Basic graph theory

At least 80% of the topics must have been covered during the applicant's bachelor studies. A graduate from the BSc in Data Science at ITU has covered these topics.

## Application Structure

Besides the formal requirements from the IT University, the applicant must submit a cover letter (at most 2 pages) containing information describing how the list of requirements is met by the applicant (list of courses where they learned about the topics above).

## Description of core courses

Algorithm Design. This course introduces students to techniques for solving complex programming tasks arising in modern IT systems. Focus in the course is on algorithm design and identification of computationally hard problems. The course contains both theoretical analysis and implementation exercises. Contents of lectures includes: Formulating an algorithmic problem, greedy algorithms, graph algorithms, divide and conquer, dynamic programming, network flow, reductions. Book: Algorithm Design, by Eva Tardos and Jon Kleinberg, Addison-Wesley, 2005.

Advanced Applied Statistics & Multivariate Calculus. The course introduces the students to advanced multivariate statistical analysis and inference. In particular, the course gives an introduction to multivariate distributions, multivariate linear models (including mixed model) and latent variable models (factor analysis, mixture models and latent time series model). For these models, methods for model fitting, construction of confidence intervals and model checking will be discussed. In addition, advanced hypothesis testing will be covered, including multivariate tests and multiple testing, and the students will be introduced to experimental design. The course contains both the theoretical foundation for the subjects as well as practical

implementation of the methods in a programming language. Throughout the course concepts from multivariate calculus will be introduced as they are needed.

*Data in the Wild: Processing heterogeneous data.* In this course, the students learn how to collect data (taking ethical and security guidelines into consideration), find complementary datasets in the wild (utilize open data, data from data lakes, private data from a company, etc.), transform the existing data into a new state if necessary, utilize datasets that contain data with different formats including unstructured data, combine and select different data analysis methods (e.g., machine learning, SQL), and iterate over these steps as long as it is necessary for a specific data science or business optimization problem. In parallel, the students also gain practice with discussing, presenting, and visualising their data processing and analysis steps as well as the final results while solving this problem.

*Data Science in Production.* In this project-based course, students work with a real-world data science problem, e.g., they enhance an existing complex data system with a new functionality. In parallel, they have to consider scalability, efficiency, and usability aspects of the added functionality. Students gain experience to interact with complex code bases and production-level code, to understand code written by other people, integrate their code in an existing code base to solve a data science problem, curate a test suite for this code, and gain experience with software engineering practices in the real world (e.g., agile development, continuous delivery, development, operations).

*Advanced Data Visualization.* In this course students build customized visualizations for ad hoc science communication projects through the extension of standard visualization libraries (e.g., D3.js), usage of programming languages explicitly developed for visualization, and advanced web-based technologies for high-quality graphics (e.g., WebGL). This course progresses from the Data Visualization bachelor course, in which students learn the basics of human perception and how to use standard libraries and visualization techniques to communicate efficiently and effectively. Advanced Data Visualization goes beyond standard techniques, and aims to provide the student to innovate in the field of data visualization.

By the end of the course, the student should have the tools and skills to build visualizations such as: <https://www.youtube.com/watch?v=4glhRkCcD4U> or <http://webverse.org/>

*Interdisciplinary Project.* This is a course where students from different MSc programmes at ITU collaborate on a project. For students with data science skills, there will be ample opportunities to contribute to interdisciplinary projects in business IT, computer game analytics, design and communication.

*Research Project including Ethics.* This project is a precursor activity for preparing students for their MSc thesis. The project requires students to find a supervisor and thesis topic at the beginning of the second year. It includes research design, including doing a literature review, collecting data, and preliminary implementation for the thesis. Moreover, this course contains a mandatory activity in which students will be required to prepare a reflective review paper on

recent discussions on Ethics within Data Science, Machine learning and Artificial Intelligence. At the end of the project a report and oral exam will be graded. Passing it is a requirement to proceed to the MSc thesis.

## **Programme Title and Objectives**

### *Programme Title*

**Section 1.** A student, who has completed the programme, has the right to use the title *candidatus/candidata scientiarum (cand.scient.) i datavidenskab.*

*Subsection 2.* The title in English is Master of Science (MSc) in Data Science.

### *Programme Objectives*

**Section 2.** The purpose of this programme is to provide students with the scientific qualifications to identify, formulate, solve and reflect on complex problems relating to data science.

*Subsection 2.* The programme prioritises the student's ability to assess, apply and develop the underlying technology and the scientific theories, methods and tools upon which it is based.

*Subsection 3.* The student must have the ability to independently initiate and carry out collaborative work in professional and multidisciplinary settings. Furthermore, the student must have the ability to engage in global and distributed interaction, drawing on research-based perspectives.

*Subsection 4.* On the background of the student's preceding bachelor's programme, the programme provides the student with the qualifications to define his or her own academic profile within the field of data science and to take independent responsibility for his or her own professional development and specialisation.

*Subsection 5.* Within the framework of the programme, the student can acquire the requisite individual qualifications for specialised posts in business, industry, the public sector, and for research training programmes (PhD programme) in data science.

### *Objectives for Learning Output*

**Section 3** The graduate will develop research-based *knowledge and understanding of, and will be able to reflect on:*

- Theory and practice within data-science specific areas of mathematics (principles of advanced statistical analysis, inference and calculus).
- Theory and practice within data-science specific areas of scalable computing and data analytics (e.g., algorithm design, advanced visualization, data acquisition, learning from

- heterogeneous including unstructured data sources) and its applications to real-world scenarios.
- Principles of ethics and fairness within Data Science.
  - Theory, scientific methodology and scientific issues within data science in the above areas at the highest international research level.

*Subsection 2.* The graduate will develop the following research-based *skills*:

- The graduate can master a modern programming language to implement and develop software for data analysis.
- The graduate can apply, assess and develop fundamental processes and practices to solve problems in data science. This includes the evaluation of theoretical issues of problems in data science to select, apply, implement and design scalable algorithms for fundamental data analysis (e.g., in machine learning and statistical inference) and their adequate empirical validation.
- The graduate is able to communicate, visualise and discuss the acquired data-driven knowledge with both academic peers and non-specialists.

*Subsection 3.* The graduate will develop the following research based *competences*:

- The graduate can design and develop new solutions to enhance existing complex data systems, and combine and select different analysis methods in complex and unpredictable settings.
- The graduate can independently initiate collaboration and work professionally with both data science peers and others in complex and inter-disciplinary contexts.
- The graduate can independently take responsibility for own professional development based on theoretical knowledge and practical experience to advance and adapt own competencies to future needs.

## Referencer

Her er samlet henvisning til de rapporter m.m., der er citeret fra i ansøgningen.  
Henvisning til rapporter, der indgår i Markedsundersøgelsen fra Rambøll findes i (Rambøll 2018, s. 17 ff.)

(Erhvervsministeriet 2018) Strategi for Danmarks digitale vækst. Rapport fra Erhvervsministeriet, 30. januar 2018.

[https://em.dk/media/9244/strategi-for-danmarks-digitale-vaekst\\_online.pdf](https://em.dk/media/9244/strategi-for-danmarks-digitale-vaekst_online.pdf)

(LinkedIn 2018) LinkedIn Workforce Report. 8. oktober 2018.

<https://economicgraph.linkedin.com/resources/linkedin-workforce-report-august-2018>

(IT-branchen 2018) IT barometer 2018. IT-Branchen, 2018.

[www.itb.dk/sites/default/files/it-barometer\\_2018\\_-\\_samlet\\_resultat.pptx](http://www.itb.dk/sites/default/files/it-barometer_2018_-_samlet_resultat.pptx)

(Plank 2018) *Analysis of the Danish Data Science Job Market*. Technical report.

IT-Universitetet, 14. december 2018. Se bilag i upload *Dokumentation af efterspørgsel på uddannelsesprofil*

(Rambøll 2018) *Markedsundersøgelse Kandidatuddannelse i Data Science*. Rambøll

Management Consulting, december 2018. Se bilag i upload *Dokumentation af efterspørgsel på uddannelsesprofil*

(ITU 2019) *Beskæftigelsesgrad for engelske og danske dimittender fra IT-Universitetet i*

*perioden 2009-2014*. IT-Universitetet 11. januar 2019. Se bilag i upload *Dokumentation af efterspørgsel på uddannelsesprofil*

IT-Universitetet

E-mail: [itu@itu.dk](mailto:itu@itu.dk)

### Godkendelse af ny uddannelse

Uddannelses- og forskningsministeren har på baggrund af gennemført prækvalifikation af IT-Universitetet i Københavns (ITU) ansøgning om godkendelse af ny uddannelse truffet følgende afgørelse:

#### Godkendelse af ny kandidatuddannelse i datavidenskab

Afgørelsen er truffet i medfør af § 20 i bekendtgørelse nr. 205 af 13. marts 2018 om akkreditering af videregående uddannelsesinstitutioner og godkendelse af videregående uddannelser

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).

Da IT-Universitetet i København er positivt institutionsakkrediteret gives godkendelsen til umiddelbar oprettelse af uddannelsen.

Styrelsen gør opmærksom på, at godkendelsen af uddannelsen ikke ændrer på den samlede ramme for optag af engelsksprogede studerende, som ITU har aftalt med Uddannelses- og forskningsministeriet.

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

#### Hovedområde:

Uddannelsen hører under det naturvidenskabelige område.

#### Titel

Efter reglerne i uddannelsesbekendtgørelsens § 21, stk. 1 og nr. 4.7. i bilag 1, fastlægges uddannelsens titel til:

**Dansk:** Cand.scient. i datavidenskab

**Engelsk:** Master of Science (MSc) in Data Science

12. april 2019

#### Styrelsen for Forskning og Uddannelse

Professions- og Erhvervsrettede  
Videregående Uddannelser

Bredgade 40  
1260 København K  
Tel. 3544 6200

[www.ufm.dk](http://www.ufm.dk)

CVR-nr. 1991 8440

Sagsbehandler  
Jørgen Prosper Sørensen  
Tel. 72 31 90 01  
[jso@ufm.dk](mailto:jso@ufm.dk)

Ref.-nr.  
19/007249-17

Udbudssted:

Uddannelsen udbydes i København.

Sprog:

Ministeriet har noteret sig, at uddannelsen udbydes på engelsk.

Ministeriet bemærker hertil, at det fremgår af § 7, stk. 1, i bekendtgørelse nr. 106 af 12. februar 2018 om adgang til kandidatuddannelser ved universiteterne (kandidatadgangsbekendtgørelsen) som ændret ved bekendtgørelse nr. 257 af 19. marts 2019, 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 § 20 fastlægges uddannelsens normering til 120 ECTS-point.

Takstindplacering:

Uddannelsen indplaceres til takst 3.

Aktivitetsgruppekode: 6255.

Koder Danmarks Statistik:

UDD: 8027.

AUDD: 8027.

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.

Adgangskrav:

Efter det oplyste er følgende uddannelser direkte adgangsgivende til kandidatuddannelsen, jf. § 11, stk. 2, i uddannelsesbekendtgørelsen:

- Bachelor i datavidenskab (IT-Universitetet) (vil få retskrav til ny kandidatuddannelse)
- Bachelor i datavidenskab (Århus Universitet)
- Bachelor i datavidenskab (Aalborg Universitet)
- Bachelor i machine learning og data science (Københavns Universitet)
- Bachelor i kunstig intelligens og data (DTU)

Ministeriet bemærker hertil, at det af hensyn til de studerendes retssikkerhed tydeligt skal fremgå af uddannelsens studieordning samt universitetets hjemmeside, såfremt der er andre uddannelser end de ovenfor nævnte, der anses som adgangsgivende til uddannelsen.

Herudover skal den studerende have sproglige færdigheder i engelsk svarende til gymnasialt B-niveau.

Dimensionering/maksimumramme/kvote:

Ministeriet har ikke fastsat en maksimumsramme for tilgangen til uddannelsen.

Med venlig hilsen

Jørgen Prosper Sørensen  
Chefkonsulent

## Bilag: RUVU's vurdering

Nr. A6 - Ny uddannelse – prækvalifikation (forår 2019)		Status på ansøgningen: Godkendelse	
<b>Ansøger og udbudssted:</b>	ITU		
<b>Uddannelsestype:</b>	Kandidat		
<b>Uddannelsens navn (fagbetegnelse):</b>	Datavidenskab (dansk) Data science (engelsk)		
<b>Den uddannedes titler på hhv. da/eng:</b>	Cand.scient. i datavidenskab (dansk) Master of Science (MSc) in Data Science		
<b>Hovedområde:</b>	Naturvidenskab	<b>Genansøgning:</b> (ja/nej)	Ne j
<b>Sprog:</b>	Engelsk	<b>Antal ECTS:</b>	120 ECTS
<b>Link til ansøgning på http://pkf.ufm.dk:</b>	<a href="http://pkf.ufm.dk/flows/3704d145882a4305254cb3e2d1479013">http://pkf.ufm.dk/flows/3704d145882a4305254cb3e2d1479013</a>		
<b>Om uddannelsen: indhold og erhvervsigte</b>	<b>Beskrivelse af den nye uddannelse, dens konstituerende elementer/struktur, erhvervsigte og adgangskrav</b>		
<b>Beskrivelse af uddannelsen:</b>	<p>IT-Universitetet i København (ITU) ansøger om at oprette en kandidatuddannelse (cand.scient.) i Data Science (datavidenskab), som bygger på kernekompetencer lig dem, der kan erhverves på ITU's bacheloruddannelse i Data Science.</p> <p>Således har den nye uddannelse jf. ansøgningen fokus på stærke tekniske kompetencer med et dybt teoretisk fundament.</p>		
<b>RUVU's vurdering på møde d. 7. marts 2019</b>	<p>RUVU vurderer, at ansøgningen opfylder kriterierne, som fastsat i bekendtgørelse nr. 205 af 13. marts 2018, bilag 4.</p> <p>RUVU har ved vurderingen lagt vægt på, at behovsafdækningen og dialogen med potentielle aftagere understøtter og dokumenterer den aktuelle efterspørgsel efter kandidater inden for datavidenskab fra den ansøgte uddannelse.</p> <p>Det noteres endvidere, at der er tale om en engelsksproget kandidatuddannelse, hvilket i forhold til det pågældende fagområde forekommer relevant i det konkrete tilfælde. Også set i lyset af, at ITU tidligere har fået godkendt en engelsksproget bacheloruddannelse i Data Science.</p>		