



# Data-literacy som en sammensat kompetence

## Francesco Caviglia

*Lektor*

Center for Undervisningsudvikling og Digitale Medier,  
Aarhus Universitet



## Alex Young Pedersen

*Lektor og ph.d.-studerende*

Aarhus Gymnasium, AARHUS TECH og Center for  
Undervisningsudvikling  
og Digitale Medier, Aarhus Universitet



## Abstract

Data-literacy defineres som en sammensat kompetence i spændingsfeltet mellem kompetencer i statistik, datavisualisering og samt mere overordnede dispositioner og kompetencer i at undersøge problemer og skabe løsninger ved hjælp af data. Data-literacy foreslås derfor som et element i den digitale dannelse, der udbygger forestillingen om hvad det vil sige at være 'dannet' i et informationsamfund. Artiklen forsøger at besvare spørgsmålet om hvorvidt idealet om en datakyndig medborger ('citizen data scientist') er en realistisk rollemodel, som med fordel kan inddrages i undervisning, eller om data-literacy udelukkende er et domæne for specialister? Med udgangspunkt i eksempler på ressourcer og praksisser, som forudsætter data-literacy, men som samtidig understøtter tilegnelse af samme, argumenterer artiklen for at data-literacy bedst defineres som en kompetence, der tilskrives et undersøgende praksisfællesskab frem for et enkelt individ. Definitionen kalder derfor på nye samarbejder mellem faglige discipliner samt arbejdsformer, som kan integrere forskellige kompetencer på tværs af fag.

## English abstract

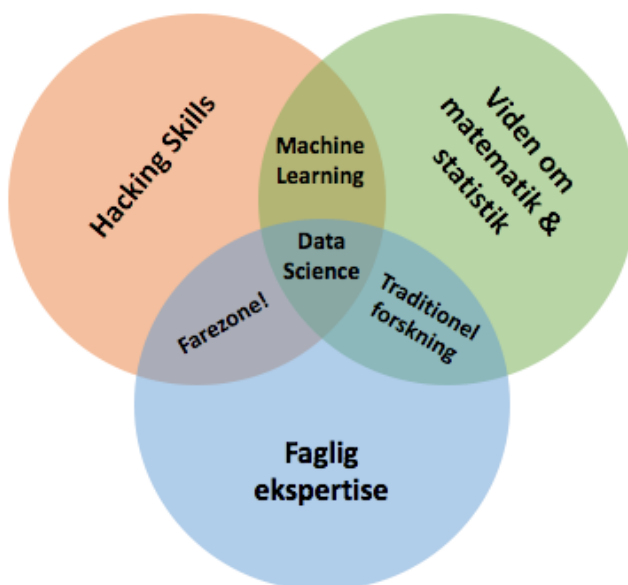
Data literacy is defined as a compound competence at the interplay between statistics, data visualisation and more general dispositions and skills in investigating problems and developing solutions with the help of data. Data literacy is, therefore, a component of digital literacy that plays a crucial role in defining what it means to be an educated citizen in an information society. This article aims at answering the question whether the ideal of a 'citizen data scientist' can be integrated into educational intervention as a useful role model, or whether data literacy ought instead to be left entirely to specialists. By taking as a point of departure examples of resources and practices that both require and foster data literacy, this article argues for understanding data literacy as a competency that is primarily the prerogative of a community of inquiry rather than of a single individual. This definition calls therefore for tighter collaboration among disciplines and educational interventions that integrate different competencies in interdisciplinary contexts.

## Introduktion og problemstilling

Baggrunden for denne artikel er udviklingen mod et mere 'data-drevet samfund' (Pentland, 2013) eller 'datafication' (Mayer-Scöhoenberger & Cukier, 2013), hvor fænomener i naturen og samfundslivet i stigende grad kvantificeres med henblik på at opnå ny viden gennem dataanalyse (Bell, Hey, & Szalay, 2009; Hey, Tansley, & Tolle, 2009). Dataficering og den hastigt øgede mængde af åbent tilgængelig data på internet (Open Data, s.n.) er blandt andet et resultat af vores brug af digitale teknologier som ofte indebærer både positive og negative konsekvenser (Samuels, 2008). På den ene side udgør dataficering et potentiale for en bæredygtig samfundsudvikling (Sharif & Van Schalkwyk, 2016) og økonomisk vækst (Stott, 2014), men der er også en risiko for, at data anvendes mod resourcesvage borgere (Gurstein, 2011) eller som midler til elektronisk overvågning af *alle* borgere (Mau, 2017; Nepper Larsen, 2017).

Disse udviklinger åbner op for en række spørgsmål om politik og økonomi som ligger udenfor dette tidsskriftets fokusområder, men åbner samtidig spørgsmålet om hvilke kompetencer der er nødvendige for et aktivt medborgerskab i et samfund, hvor data er centrale i individuelle og kollektive beslutningsprocesser.

Figur 1 viser samspillet af kompetencer, som er involveret i praksisser af datavidenskab ('Data Science'), ifølge Drew Conway (2013; se kommentarer i eks. Parkkinen, 2015).



Figur 1 – Kompetencesamspil i 'data science' (oversat fra Conway, 2013)

Figuren illustrer ved hjælp af en Venn-diagram fagets sammensatte og tværfaglige karakter, idet datavidenskab opstår i overlappet mellem dyb domænespecifik viden, viden om statistik og ikke mindst It-operationelle kompetencer i at finde, organisere og klargøre data til analyse. Datavidenskab kan desuden anses som en udvidelse af traditionelle forskningsmetoder, som allerede bygger på viden om statistik. Figuren viser også nødvendigheden af at indtænke alle områder samtidig. Ekspertise og adgang til teknologien kan ikke stå alene, da den i fraværet af viden om matematik og grundlæggende metoder udgør en farezone. Problemstillingen tages op igen nedenfor i diskussionen.

Øget efterspørgslen på specialister indenfor it og databehandling samt almindelige, it-kyndige borgere bliver forsøgt adresseret på mange niveauer af uddannelsessystemet i forhåbningen om at kunne imødekomme aktuel og fremtidig efterspørgsel (Breinstrup, 2017). På universitetsniveau er der eks. i 2017 blevet godkendt 6 nye uddannelser ved navn 'datavidenskab', 'data science' eller 'business analytics' (Thomsen, 2017). Allerede nu er der stor opmærksomhed på uddannelsen af specialister. Fokus for denne artikel er ikke specialistkompetencer, men derimod almene kompetencer i at anvende data (Frank, Walker, Attard & Tygel, 2016). 'Data-literacy' er udtryk for denne generelle kompetence, hvor 'literacy' henviser til et udvidet, multimodalt læse- og tekstbegreb samt et flertal af anvendelseskontekster (Lund, 2017), mens 'data' refererer til ekstra viden og kunnen.

Fokus er altså ikke på professionelt arbejde som forsker eller data-analytiker, men på mere almene kompetencer som kan indgå som læringsmål på udskoling- og gymnasieniveau, eller på tværs af videregående uddannelser hvor informatik eller statistik ikke traditionelt har været en del af curriculum. Data-literacy – tilknyttet til idealet om en datakyndig medborger – anses som en forudsætning for såvel almen grunduddannelse samt videregående uddannelse, beskæftigelse og ikke mindst aktivt medborgerskab (se meta-studierne af Ridsdale et al., 2015, Beyond data-literacy, 2015, Wolff, Gooch, Cavero Montaner, Rashid, & Kortuem, 2016).

Det følgende afsnit viser eksempler på, hvordan data-literacy kan og bør spille en rolle i den nuværende proces, hvor it og data i stigende grad integreres i uddannelse.

Med gymnasireformen som trådte i kraft i 2017 blev faget Informatik indført i gymnasiet "som et moderne almindelig og studieforberedende it-fag" (UVM, 2016:15; se eks. Læreplan Informatik B - htx, 2017), og læreplanerne i alle fag i gymnasiet har revideret og i mange tilfælde udvidet afsnittet om 'it'

som en del af fagets tilrettelæggelse (UMV, 2017a). I folkeskolen har man på forsøgsbasis indført faget Teknologiforståelse (Teknologiforståelse valgfag, 2017; Caspersen, 2017). Formålet med faget Teknologiforståelse er “at den enkelte elev opnår en kritisk forståelse af informatikkens muligheder og rolle i det demokratiske digitale samfund, og eleverne skal både analysere og producere digitale produkter” (UVM, 2017b).

Tiltagene i folkeskolen og gymnasieskolen peger dog også på behovet for inkorporering af data-literacy som en almen og tværgående kompetence i et digitalt medieret og datadrevet samfund. Uddannelse med henblik på medborgerskab bør altså integrere data-literacy på meningsfulde måder i de enkelte fag, hvor denne kompetence forbindes til faglig viden i fagene og på tværs af fagene.

Kompetencer i databearbejdelse og -analyse er eksplicit nævnt som faglige mål i flere valgfag på gymnasieniveauet, eks. i Matematik A, Biologi B og Samfundsfag A på stx med fokus på dokumentation, analyse og kommunikation. (eks., “I udarbejdelsen af skriftlige opgaver og projekter skal forskellige datatyper og datarepræsentationer kunne integreres”; Læreplan Samfundsfag A-stx, 2017).

Danskfaget i gymnasiet indgår i fagligt samspil med de øvrige fag, bl.a. i forbindelse med sit fagligt mål om at kunne “anvende forskellige mundtlige og skriftlige fremstillingsformer formålsbestemt og genrebevidst, herunder redegøre, kommentere, argumentere, diskutere, vurdere og reflektere”, mens “arbejde med digitale analyseobjekter” indgår i fagets bidrag til elevernes digitale dannelse (eks. Læreplan Dansk A-htx, 2017), hvilket åbner for at fokusere på det multimodale element i at repræsentere data.

Flere gymnasiale fag giver med andre ord mulighed for og lovhjemmel til at fremme data-literacy. Samtidigt er data-literacy ikke tilstrækkeligt italesat som tværfaglig kompetence i den danske debat om integration af it i undervisning. Denne artikel foreslår, at data-literacy bliver inkluderet som en central komponent af almen og digital dannelse og vil bidrage til formålet med at:

- præcisere behovet for data-literacy i forhold til elever og almene borgere, med udgangspunkt i konkrete eksempler af data, kontekster og problemstillinger
- identificere udfordringer og åbne spørgsmål i forhold til at gøre data-literacy til et eksplicit læringsmål.

Spørgsmålet er om idealet med den datakyndige medborger er en realistisk rollemodel, som med fordel kan inddrages i undervisning eller om det udelukkende er forbeholdt specialister med baggrund i data science?

For at bedre forstå behovet og udfordringer ift. data-literacy er artiklen organiseret i de følgende afsnit:

- et overblik over eksisterende definitioner af data-literacy
- tre cases med data-literacy ressourcer og praksisser, dvs. to websider (Hans Roslings [gapminder.org](http://gapminder.org) og Max Rosers *Our World in Data* [ourworldindata.org](http://ourworldindata.org)) med statistiske data som afspejler forskelle, ændringer og udviklinger i livsvilkår, og en banebrydende data-journalistisk undersøgelse om spilleautomater i Italien (Bruschi, 2017)
- et overblik over tilgange til at fremme data-literacy, med fokus på tværfaglige, problembaserede aktiviteter i undersøgende praksisfællesskaber
- en opsamlende diskussion over pædagogiske og etiske udfordringer i forbindelse med at fremme data-literacy i uddannelse.

Behovet for data-literacy er samtidigt et praktisk spørgsmål om forfatterens egen datakyndighed. Med primære uddannelsesbaggrunde indenfor humaniora er ingen af forfatterne formelt kvalificeret indenfor statistik, men begge forfattere har arbejdet og tilegnet sig en lægmands tilgang til at kunne læse og anvende nogle data-analytiske metoder. Vores viden om statistik er derfor utilstrækkelig eksempelvis i at forstå matematikken bag lineær regression og vi er derfor afhængige af at kunne rådføre os med andre i spørgsmål om statistik. For os som undervisere er anvendelse af statistiske begreber og værktøjer til stadighed et spørgsmål om at undervise på grænsen af egen viden. Vi ser dog begge vigtigheden og behovet for at kunne bidrage til at fremme kompetencer indenfor data-literacy. Derfor ønsker vi at pege på relevansen af data-literacy samt at anviser en fælles forståelsesramme af data-literacy som en sammensat kompetence.

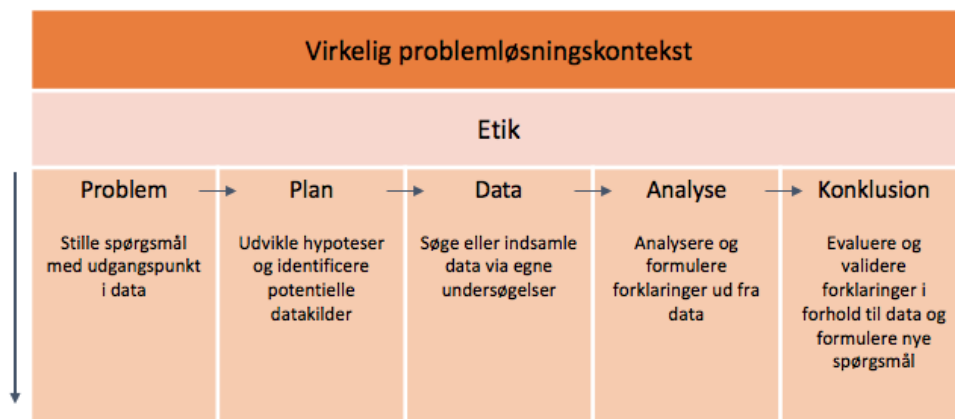
Det primære fokus er samspelet mellem fagligheder, digitale kompetencer og aktivt medborgerskab, og især på processer hvor opbygning af viden sker i samarbejde med andre. Et eksempel på samarbejde er 'fælles faktatjek', der omhandler etableringen af et fælles grundlag for viden, som kan danne udgangspunkt for håndtering af et problem eller udfordring (Innes & Booher, 2010). Det handler om at kunne blive enige om, hvilke data der er til rådighed og hvordan man kritisk vurderer dem. Det er en forudsætning for meningsfuldt at kunne diskutere tolkning af data, at man er enige om datas status (Shapin & Schaffer, 1985).

## At definere data-literacy: kompetencer og kontekst

Flere studier har i de seneste år defineret data-literacy som en af de 'literacies of the digital' der kræver livslang kompetenceudvikling (Littlejohn, Beetham & McGill, 2012). Det er vigtigt at fremhæve at data-literacy ikke er fast defineret, men under stadig udvikling på grund af dets forbindelse til samfundsmæssige og teknologiske forandringer.

Eksempelvis definerer Risdale et al. (2015) data-literacy som evnen til at indsamle, administrere, evaluere og anvende data på en kritisk måde, mens *Beyond data-literacy* (2015) fokuserer på interessen for og evnen til konstruktivt at engagere sig i samfundet gennem brugen af data. I et review af litteraturen har Wolff et al. (2016) vist hvordan data-literacy er defineret og har identificeret to brede kategorier som svarer til et spændingsfelt mellem specifikke faglige kompetencer i statistik samt datavisualisering og mere overordnede dispositioner og kompetencer i at undersøge problemer og søge løsninger ved hjælp af data.

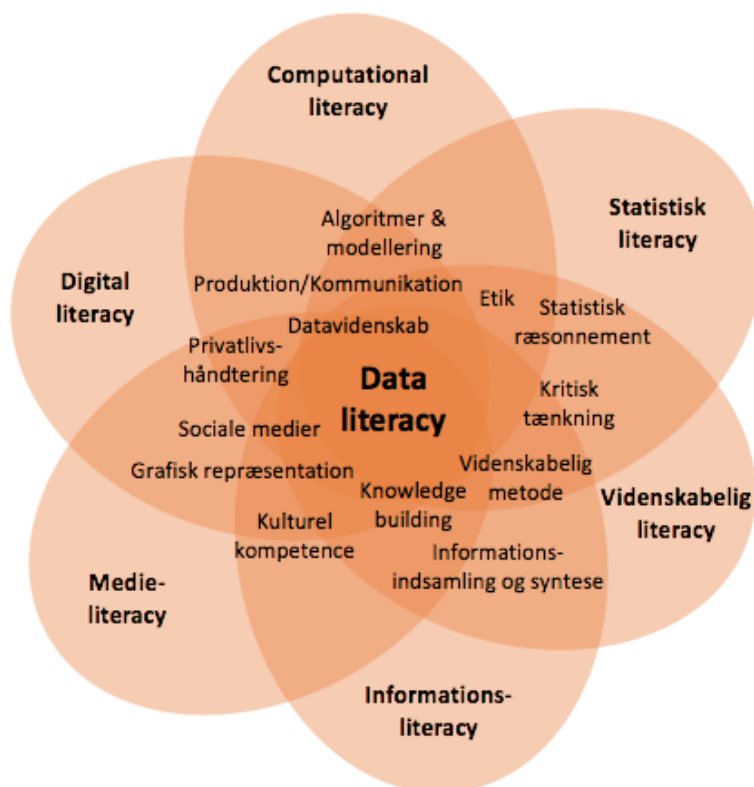
De foreslår på den baggrund en model for data-literacy, hvor specifikke kompetencer i behandling og analyse af data er sammenkoblede med viden og kompetencer til kritisk at undersøge problemstillinger (se figur 2).



Figur 2 – Model for data-literacy oversat fra Wolff et al. (2016), som er inspireret af en *Problem-Plan-Data-Analysis-Conclusion*-model (PPDAC i Wild & Pfannkuch, 1999). Pilen på venstre side peger på mere avancerede kompetencer i databehandling.

Modellen omfatter desuden den etiske anvendelse og behandling af data med henblik på at beskytte folks sikkerhed og privatliv, samt at overveje etiske konsekvenser af brugen af data i forskellige sammenhænge. For eksempel var den studerende, som offentliggjorde oplysninger om 70.000 brugere af online datingservicen 'Ok Cupid' kompetent i databehandling, men ikke tilstrækkelig datakyndig til at forstå at offentliggørelsen af disse data kunne være skadelige for individer, som efterfølgende ville kunne identificeres på baggrund af deres seksuelle præferencer (Allingstrup & Wissing, 2016; en diskussion om mål og midler med at dele netop disse forskningsdata i Resnik, 2016; McCook, 2016).

Angående forståelsen af forholdet mellem data-literacy og andre kompetencer, fremlægger rapporten *Beyond data-literacy* (2015) en figur i form af et Venn-diagram, hvor data-literacy placeres i midten af delvis overlappende kompetencer indenfor andre literacies (Se figur 3)



Figur 3 – Model af 'literacies' oversat fra *Beyond data literacy* (2015).

Modellen forudsætter en bred definition af data-literacy og forklarer, hvordan data-literacy overlapper med flere 'literacies' og har stærk overførelsesværdi på tværs af kontekster og discipliner. Især vil vi i forbindelse med artiklens fokus om almen dannelse fremhæve, hvordan data-literacy delvis er sammenfaldende med elementer fra informations-

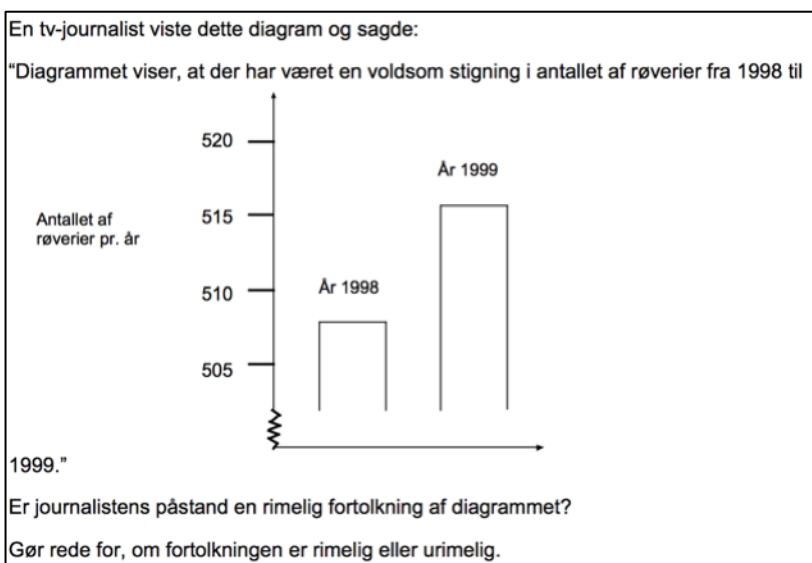


literacy, herunder informationsindsamling samt videnskabelig literacy, herunder kritisk tænkning. For eksempel er det ikke tilstrækkeligt at kunne læse og forstå resultaterne af en spørgeskemaundersøgelse, uden at forholde sig til validiteten af stikprøven og risikoen for 'social desirability bias' (se eks. Tourangeau, Rips & Rasinski, 2000). Artiklens fokus er dog primært på den konstruktive dimension af data-literacy, mens den kritiske dimension kort bliver behandlet i diskussionen.

### **Data-literacy som tværfaglig kompetence**

Data-literacy er både en kompetence som går på tværs af fag og som bedst integreres i de enkelte fag i stedet for at oprette et nyt fag. Stephenson og Schifter Caravello (2007) har påpeget fordelene ved at inkorporere data-literacy i kurser indenfor sociologi, hvor mulighederne for meningsfuldt at anvende det de studerende lærer øges og hvor den kritiske refleksion og evaluering sker i en faglig kontekst. I deres projekt kobles informations-literacy med data-literacy i en sociologi-faglig kontekst. *Beyond data literacy* (2015) viser også at det er vanskeligt at tænke data-literacy afkoblet fra andre literacies, faglige og tværfaglige kontekster, derfor er det også vanskeligt at finde undersøgelser som har forsøgt at måle data-literacy som en selvstændige kompetence. OECD' PISA-test giver dog indblik i, hvordan elementer af 'data-literacy' forekommer i prøverne indenfor matematik ('mathematical literacy'), læsning ('reading literacy') og naturfag ('scientific literacy'), hvilket bekræfter *Beyond data-literacy's* (2015) model med overlappende kompetencer.

Som en anonym reviewer bemærker, handler data-literacy "både om at kunne forholde sig til indholdet, men i meget høj grad også til måden indholdet er repræsenteret". At kunne forstå og ideelt producere multimodale repræsentationer af data er nemlig et centrale element i data-literacy. Eksempelvis beder test item *Røverier* (spørgsmål 9.1 under 'Matematik' i OECD, 2009:110 og 4.7 i Lindenskov & Weng, 2010) eleverne om at vurdere, hvorvidt en udtalelse er en forsvarlig fortolkning på baggrund af et givent diagram (se figur 4).



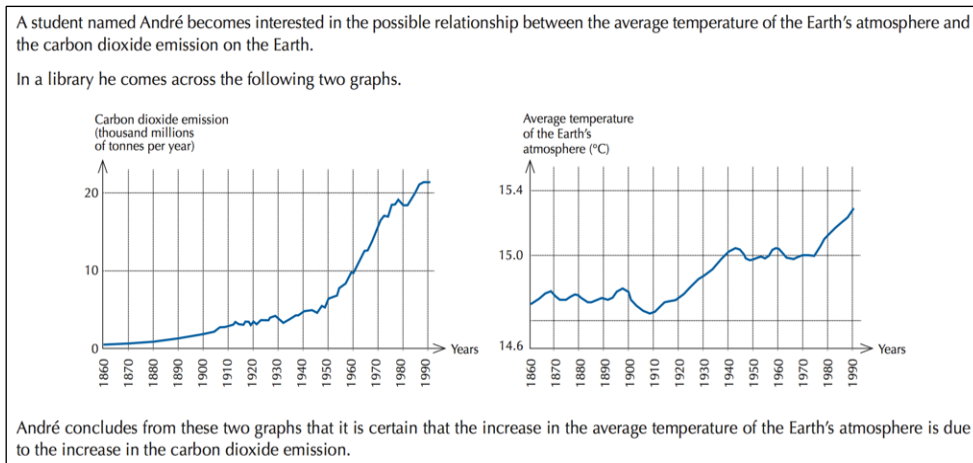
Figur 4 – PISA test item 4.2 Røverier (Lindenskov & Weng, 2010)

I 2000 gav 34% af de adspurgte danske folkeskoleelever et korrekt svar (beregnet som procent for 'fyldstgørende korrekt' plus halvdelen af procent for 'delvis korrekt', mens OECD gennemsnit var 30% og op til 46% i Finland (se Lindenskov & Weng, 2010, s. 40-45 med detaljer og kommentarer over elevernes svar; OECD, 2009, s. 300 viser 23% 'fyldstgørende korrekte svar' for Danmark, 26% som gennemsnit for OECD-lande og 37% for Finland).

Siden 2015 har OECD-PISA primært rubriceret fortolkning af data under videnskabelig literacy ('scientific literacy') der blandt andet omhandler evnen til at

- vurdere og designe naturvidenskabelige undersøgelser [...]
- *tolke naturvidenskabelige data og kendsgerninger, dvs. analysere og vurdere data*, påstande og argumenter i forskellige sammenhænge og repræsentationer samt komme frem til passende naturvidenskabelige konklusioner. (Christensen, 2016, s. 17, vores fremhævning; se den originale formulering og forklaring i OECD, 2016, s. 17-46, hvor den videnskabelige metode præsenteres som 'scientific method' og naturfag udgør anvendelseskonteksten).

Denne evne eksemplificeres i OECD-PISA's ramme for 'scientific literacy' ved hjælp af et test item, som kræver at deltagerne observerer to grafer og argumentere for eller imod en sammenhæng mellem to variabler (figur 5).



Figur 5 – Test item for naturfag ('scientific literacy'; OECD, 2016, s. 29).

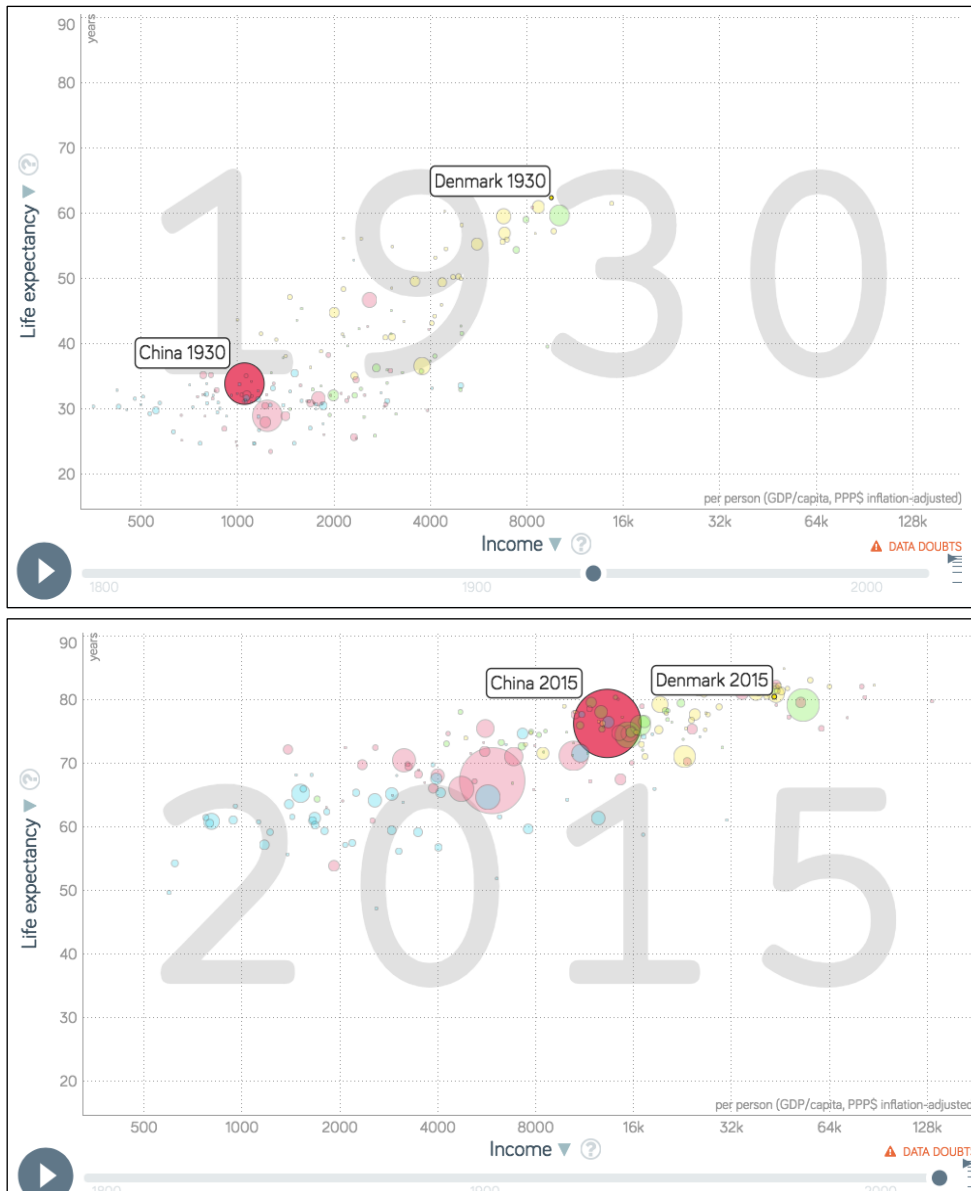
Et korrekt svar kan be- eller afkræfte sammenhængen, men skal som minimum argumentere med udgangspunkt i grafikens data og/eller metoder, som kan bruges til at verificere om der findes en korrelation.

I det følgende afsnit præsenteres eksempler på praksisser, som bygger på teknologier til analyse og visualisering af data. Eksemplerne har til formål at skabe en bedre forståelse for data-literacy som praksis og læringskontekst.

### Cases 1. og 2. To databaser om udviklingen af levevilkår

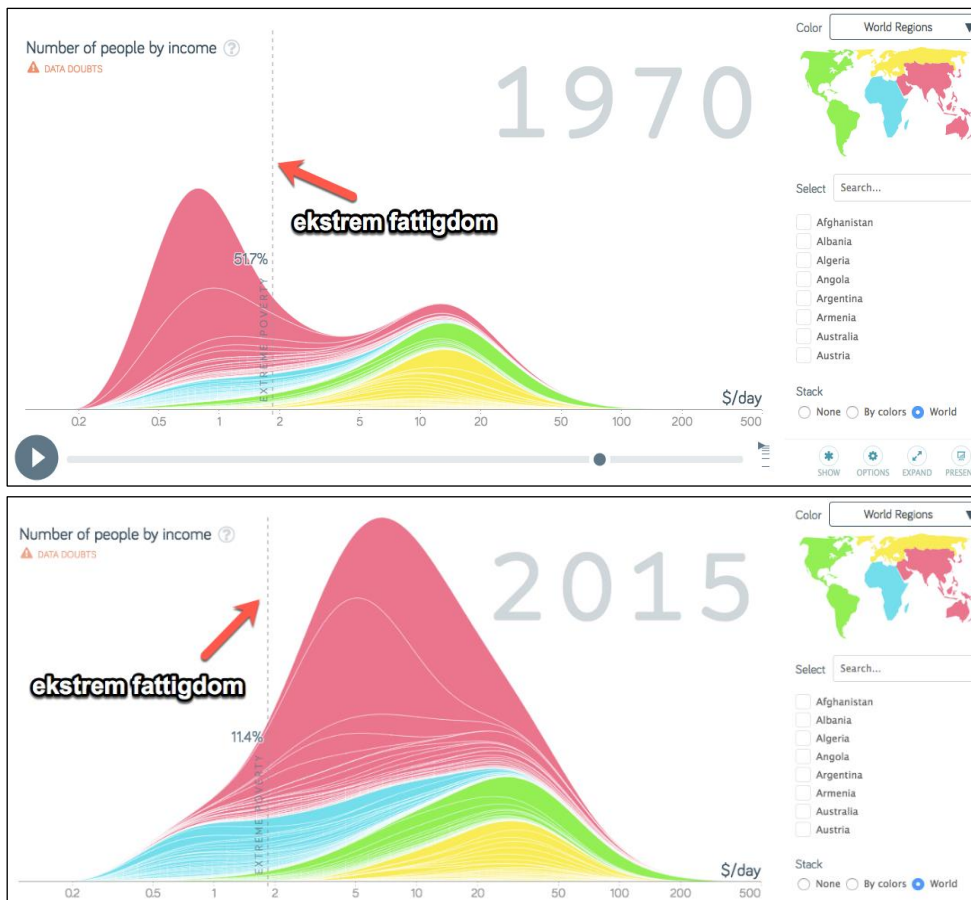
Hans Roslings TED-talk "[The best stats you've ever seen](#)" (Rosling, 2006), som er p.t. blevet set af mere end 12 millioner mennesker, er et berømt eksempel på, hvordan data og visualiseringer kan hjælpe med at fremme forståelsen af levevilkår i verden og de dybe forandringer de har undergået i de seneste 100 år.

Rosling opbygger sine præsentationer på baggrund af tal fra en database som er tilgængelig på websiden [gapminder.org](http://gapminder.org). Databasen indeholder en række indikatorer om demografi, økonomi, uddannelse og miljø, og det er muligt at kortlægge udviklingen for de enkelte lande over tid (se eksemplet figur 6a and 6b).



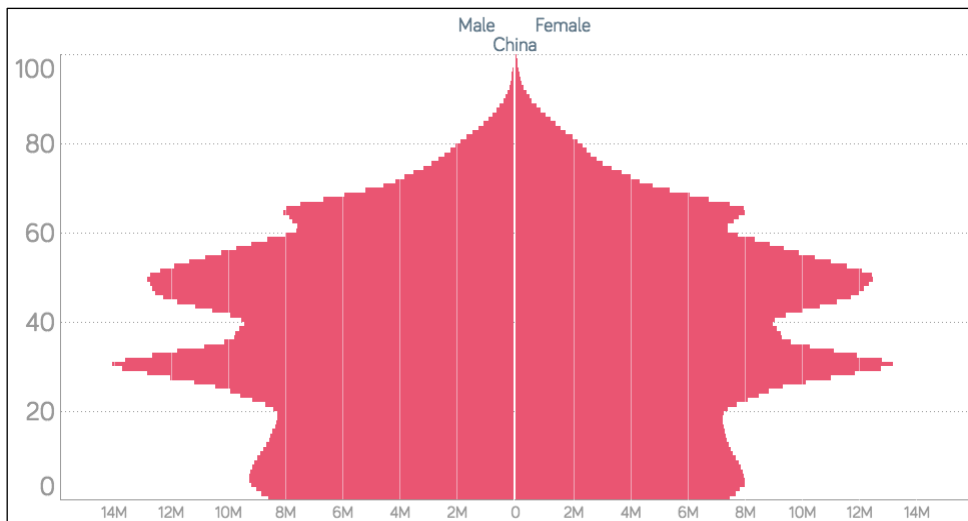
Figur 6a-6b – Forventede levealder og indkomst i 1930 og 2015; Kina og Danmark er fremhævede. Cirkulens størrelse referer til befolkningstallet i det pågældende land. (Gapminder.org).

Brugerne kan starte ved at tage en [Gapminder Test](#), hvor den typiske bruger vil blive overrasket af sin egen mangel på viden om nuværende vilkår og forudsigelser om bl.a. demografisk udvikling. Derefter kan brugeren udforske forskellige data og sammensætte egne grafiske repræsentationer og dermed følge eksempelvis den historiske udvikling i verdens indkomstdistribution (figur 7a and 7b).



Figur 7a and 7b – Fordeling af indkomst i US-dollar pr. dag over 4 verdensregioner i henholdsvis 1970 og 2015. Den stiplede linje markerer grænsen for ekstrem fattigdom. (Gapminder.org)

Gapminder er primært et redskab til datavisualiseringer. Fortolkning af data, eksempelvis med henblik på at forstå markante ændringer eller mønstre i data, overlades til brugeren. Kinas alderspyramide kan eksempelvis indikere at landet gennem historien har været præget af forskellige dramatiske begivenheder (figur 8).



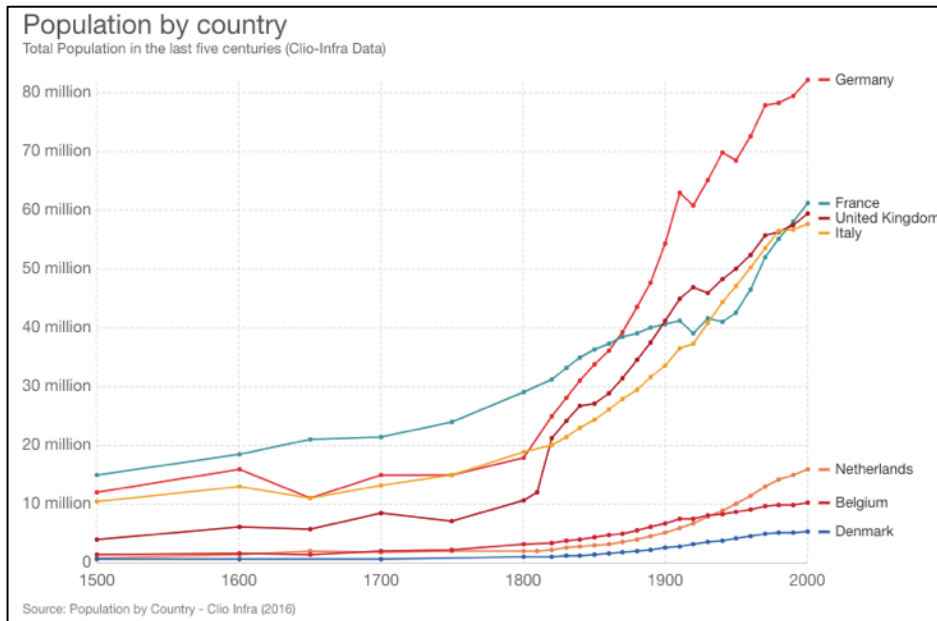
Figur 8 – Kinas alderspyramide fordelt på køn. (Gapminder.org)

Ved nærmere iagttagelse kan læseren også undre sig over, at landet har markant flere drenge end piger i alderen mellem 0 og 20 år.

Alderspyramider kan med andre ord anvendes til at fremme iagttagelsesevne samt dispositioner til at undre sig, og kan dermed danne baggrund til undersøgelsesaktiviteter, som kan udføres ved hjælp af tekstbaserede kilder og søgninger på nettet.

Gapminder er en ressource som har også udviklet sig til et manifest for en data-baseret tilgang til viden eller 'factfulness' (Rosling, Rosling og Rosling Rönnlund, 2018). Desuden har arbejdsgruppen bag Gapminder gjort deres data samt forsknings- og undervisningsmaterialer tilgængelige online på en offentlige mappe (Gapminder Public, s.d.)

Økonomen Max Roser leder et team af forskere og udviklere på Oxford University som står bag projektet *Our World in Data* (Roser, 2018; fremover OWiD), der primært er rettet mod ældre eller mere avancerede brugere. Projektets webside tilbyder data og visualiseringer, samt uddybende artikler som forklarer dataenes relevans og metodologien bag undersøgelsen. Ud over at anvende offentlige databaser tager artiklerne i OWiD ofte udgangspunkt i historisk forskning, eksempelvis demografisk udvikling (figur 9).



Figur 9 – Befolkningstilvækst i udvalgte europæiske lande, 1500-2000.  
Kilde: Roser & Ortiz-Ospina, 2017, baseret på data fra Clio Infra, 2016.

OWiDs ambition er at give et detaljeret billede af nuværende levevilkår samt at kortlægge ændringer og afdække trends på forskellige indikatorer.

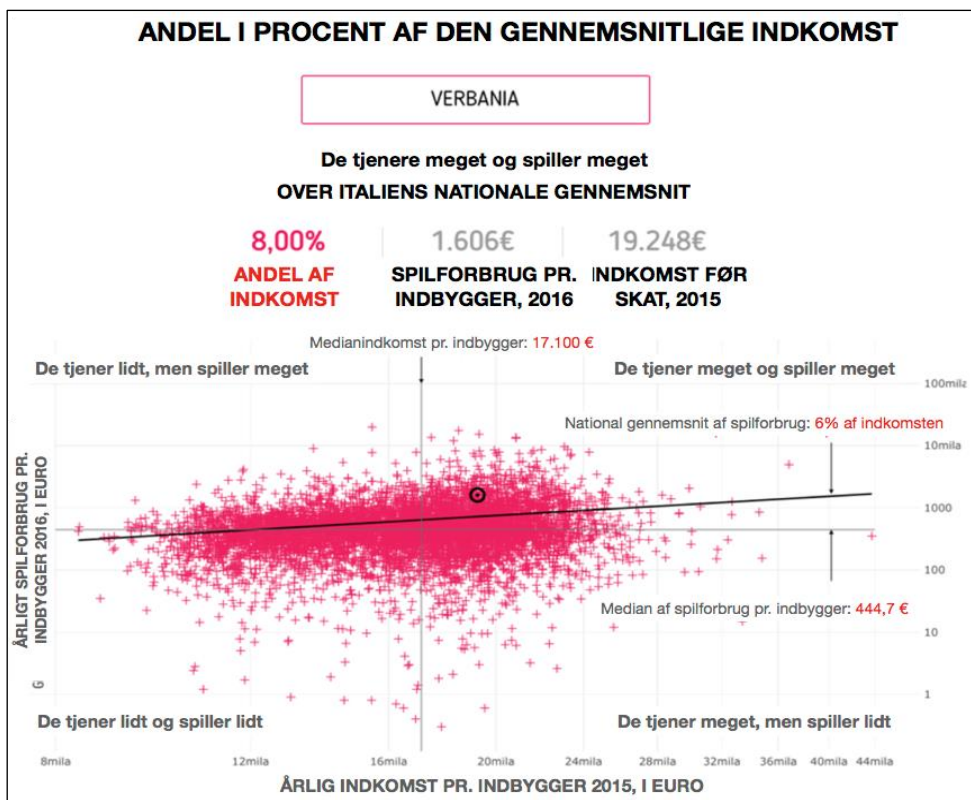
Både Gapminder og OWiD tilbyder mulighed for at downloade datae og selv foretage analyser og visualiseringer ved hjælp af regneark-software (eks. *Google Sheets* eller *Microsoft Excel*) eller mere specialiserede værktøjer til dataanalyse (eks. *Tableau*). Der er imidlertid stor forskel på at kunne læse datavisualiseringer og selv at fremstille dem. Projekter som Gapminder eller OWiD udmærker sig blandt andet ved at reducere kompleksitet i arbejdet med data og som et redskab til læring ved hjælp af visualiseringer. Samtidigt er det vigtigt at være bevidste om, at vi befinder os på et indledende stadium i opbygningen af data-literacy.

### Case 3. En undersøgelse om spillemaskiner i Italien

Spilleautomater (slot machines) udgør mere end 50% af omsætningen i den lovlige spilindustri i Italien (Talamo & Manuguerra, 2016), hvor politisk regulering er et tilbagevendende tema i den offentlige debat. Lokale græsrodsbevægelser og lokale myndigheder forsøger rutinemæssig at begrænse omfanget og spredningen af spilleautomater ved hjælp af administrativ regulering og boykot-kampagner. Debatten foregår ofte langs en principiel konfliktlinje mellem frihed og økonomisk interesse på den ene

side overfor hensynet til folkesundheden på den anden side. Kun sjældent beror debatten på faktuelle data om fænomenet. De nationale statistiske institutioner i Italien indsamler data om den lovlige spilindustri (Talamo & Manuguerra, 2016), men detaljer om hvor meget der bruges på spilleautomater i det enkelte lokalsamfund var ikke tilgængeligt. Det ændrede sig imidlertid da det ved hjælp af loven 'Freedom of Information Act' udstedt i 2016 (Bruschi, 2017), lykkedes et konsortium af journalister at tvinge de nationale skattemyndigheder til at frigive disse data. Efter at have overvundet en række tekniske forhindringer – de 10.000 sider af data blev udleveret i pdf-format – lykkedes det endelig konsortiet at lancere en interaktiv webside som giver brugerne mulighed for at finde antallet af spilleautomater og det gennemsnitlige spillebeløb for hver kommune (it. *comuni*), som udgør den mindste administrative enhed i Italien med et gennemsnitligt indbyggertal på 7.604 (2017). I Danmark er det gennemsnitlige indbyggertal i landets kommuner til sammenligning 58.660 (2017). Brugeren kan via websiden se lokale data om spilforbrugets omfang pr. indbygger både som en enkel monetær størrelse (€) eller som andel i procent af den gennemsnitlige indkomst i lokalområdet. Endvidere kan man sammenligne de lokale data med andre kommuner eller i forhold til landsgennemsnittet. Eksempelvis kan man finde 2015-data fra Verbania – en bykommune i regionen Piemonte, hvor lokale myndigheder forgæves har forsøgt at indskrænke tidsrummet, hvor man kunne spille på spilleautomaterne (Stella, 2012) – der viser at man i snit brugte 1.600 € pr. indbygger svarende til 8% af den gennemsnitlige indkomst (se figur 10).





Figur 10 – Data fra byen Verbania fra databasen om spilleautomater i Italien (GEDI, 2017, delvis oversat)

Websiden sammenligner figurer for hver kommune i forhold til nationale indikatorer og tilbyder yderligere information om penge-flow'et, hvoraf ca. 70% går tilbage til spillerne, 17% til staten, 6% til forretningsindehaveren og 5% til det selskab som ejer spilleautomaterne (dette er data for én type spilleautomat, andre typer sender op til 88% tilbage til spillerne). Også disse data om de enkelte typer af spilleautomater var ikke særlig udbredt i offentligheden.

I kølvandet på denne undersøgelse bragte mange lokalaviser artikler som skulle hjælpe med at anskueliggøre og nuancere data. Eksempelvis var en høj værdi ikke ensbetydende med et højt forbrug blandt lokalbefolkningen i turistområder, mens et område med et meget lavere forbrug stadig kunne afspejle store sociale problemer. Fortolkningen af data forudsætter kendskab til den lokale kontekst. Overordnet har undersøgelsen været med til at øge kvaliteten af den offentlige debat ved at inkludere data som alle parter er tvunget til at forholde sig til og være enige om, dog uden at dette dikterer en løsning.

Undersøgelsen har kun været mulig ved den samlede indsats af personer med forskellig baggrund og kompetencer indenfor jura, dataindsamling, dataanalyse, grafisk datarepræsentation, webredigering og traditionel journalistisk formidling (undersøgelsesprocessen er mere detaljeret beskrevet i Bruschi, 2017).

Desuden fremmer projektet data-literacy i samfundet, idet læserne er nødt til at forstå statistiske begreber og data-repræsentationer for at kunne tale om, hvordan en kommune kan sammenlignes med en anden eller med nationale data. Flere italienske lærere – både i matematik og samfundsfag – har påpeget casens relevans som et oplagt udgangspunkt for tværfaglige læringsforløb på skole- og gymnasieniveau, hvor elever og lærere kan gøre hinanden klogere på emnet ved at forbinde almen viden (om eks. matematik eller psykologien bag gambling) med viden om lokale forhold og problemer (Manuela Delfino, personal communication).

## At tilegne sig data-literacy

Dette afsnit samt det følgende fremhæver nogle principper og åbne spørgsmål ift. at indramme data-literacy i curricula og lærerplaner.

Wolff et al. (2016) har identificeret fire typer af borgere ift. deres rolle i data-analytiske praksisser:

- *læsere*, som kan læse og fortolke data, som andre har forberedt for dem (eks. læsere/brugere af Gapminder, OWiD or websiden om spilleautomater i Italien)
- *formidlere*, som kan forstå data og formidle data i forskellige formater (eks. dem som står bag de præsentationer og brugeroverflade, som man kan læse/bruge i Gapminder, OWiD eller undersøgelsen om spilleautomater)
- *forskere*, som skaber indsigt i data i kraft af stor domænespecifik viden samt avancerede kompetencer i databehandling og statistik (eks. de data-analytikere, som rensede og forberede data til undersøgelsen om spilleautomater, eller de forskere, som i første omgang lavede de undersøgelser som nu er tilgængelige i OWiD)
- *beslutningstagere* ('makers' i Wolff et al., 2016), som integrerer data i problembaserede processer i forhold til at identificere og løse problemer (eks. ved at organisere indsamling af data med henblik på at bedre fordele ressourcer, forbedre processer, anbefale eller tage beslutninger, som i ramme af smart-city projekter, se Wolff, Kortuem & Caverio, 2015).

Disse idealtyper har et fællestræk som består af basiskompetencer i at læse data og datavisualiseringer, men typerne adskiller sig i forhold til ansvarsområder og specialer. *Formidlere* og *beslutningstagere* har behov for dybere forståelse af statistik end *læsere*, men ikke nødvendigvis samme niveau som *forskere*. *Beslutningstagere* har til gengæld behov for større indsigt i lovgivnings- og forvaltningsprocesser end andre. Samtidigt skal alle disse typer kunne forholde sig til pålideligheden af de data, som de har til rådighed. Eller sagt på en anden måde skal alle i fællesskab have viden nok til at kunne vurdere reliabiliteten af data og samtidig have nok kontekstspecifik viden til at vurdere validiteten af data.

En første udfordring for didaktisk design er derfor at skabe en kontekst, hvor deltagerne har eller tilegner sig et fælles sprog for at kunne samarbejde, samtidigt med at der gives plads til at udfolde individuelle og specifikke kompetencer i relation til den fælles problemstilling. Wolff et al. (2016) foreslår, at det muligvis er mere produktivt at anskue data-literacy som en distribueret kompetence tilhørende en arbejdsgruppe eller fællesskab, hvor alle kan give deres bidrag, i stedet for at se data-literacy som individuel kompetence (Wolff et al., 2016).

Som eksempel på sådanne kontekster beskriver vi kort et format - 'civic data hackathon' - som er blevet brugt til at forbinde personer med et behov for at undersøge et problem ved hjælp af data, med personer med kompetencer i databehandling.

'Hackathons' (hacking-maratons) er events hvori en gruppe af computerprogrammører går sammen i en koordineret indsats for at producere kode til et bestemt projekt som relaterer sig til en teknologi eller et formål, eksempelvis at udvikle apps til at hjælpe personer med specifikke handicaps.

Medborger-hackathons ('Civic data hackathons', eks. [Data Dive](#) and [Open Data Day](#)) er typer af hackathons der oprindeligt blev organiseret af data-journalister (Vermanen, 2012) og satte non-profit organisations (NPOs) i forbindelse med frivillige med datakompetencer ('data volunteers') med det formål at samarbejde om dataanalyse for at skærpe forståelsen af samfundsmæssige problemstillinger og understøtte datadrevne beslutningsprocesser (Chou, Li & Sridharan, 2014; Choy, 2017). Repræsentanter for NPOs har typisk stor områdespecifik viden og praktisk kendskab, men mangler i visse tilfælde den fornødne ekspertise i dataanalyse og ressourcerne til at kompensere for denne mangel ved at ansætte eksperter.

Studerende fra Michigan University har eksempelvis organiseret to sådanne events med i alt 9 NPOs and 40 data-frivillige, med forskellige specialer og kompetenceniveauer indenfor dataanalyse og datavisualisering til viden om 'åbne data' som kan inddrages i analyserne (Hou & Wang, 2017). De organisationsansvarlige studerende havde sørget for, at der i hver arbejdsgruppe mindst var en 'knowledge broker' (Wenger, 1999, Panofsky & Robery, 2004) der var i stand til at facilitere oversættelse, koordination og alignment mellem de to praksisfællesskaber af NPOs og de data-frivillige. En 'knowledge brokers' opgave er tosidet indenfor disse medborger-hackathons: På den ene side skal der udvikles brugbar handlingsviden ('actionable knowledge', se Markauskaite & Goodyear, 2016) til NPO'erne og på den anden side skal den samlede data-literacy styrkes indenfor praksisfællesskabet (Hou & Wang, 2017). Det er vigtigt at fremhæve, at begge mål var lige vigtige for deltagerne i en sådan event.

Spørgsmålet er om det høje niveau af engagement som ofte kendetegner en medborger-hackathon kan overføres til formelle læringskontekster i eks. gymnasier eller universitetskurser for ikke-specialister, hvor manglende motivation kan være en barriere for samarbejde og læring? Data-literacy synes indtil videre at være et felt der er præget af ildsjæles og uddannelsesforskeres pioner-projekter. Den første konference om 'Data Science Education Technology' (fremfor DSET; <http://codap.concord.org/dset/>) er fra 2017, men feltet er i en rivende udvikling ift. såvel metoder og teknologier samt de fag og tværfaglige kontekster som er involveret.

Matematik og naturvidenskab spiller en nøglerolle i denne udvikling og de initiativer, som er præsenteret som eksempler på data-literacy i skolesammenhæng af Wolff et al., tager alle afsæt i STEM-fagene.

Matematik er en væsentlige komponent i 'data-literacy', men vi foreslår at sammenligne data-literacy med læsning og skrivning som er kompetencer, som skal dyrkes og gradvis udvikles i flere discipliner og på tværs af discipliner (se eks. Horning, 2007).

Dette synspunkt kommer klart til udtryk i et oplæg fra DSET-konferencen, hvor matematik- og naturfaglærer Tim Eriksson foreslår, at skolen burde sigte mod at bringe "a whiff of data science" (hvilket vi fortolker som 'data-literacy') til undervisning:

*[...] students in school aren't ready to do "real" data science. At least not in 2017. [But] we can design lessons and activities in which regular high-school students can do what amounts to proto-*

*data-science. The situations and data might be simplified, and they might not require coding expertise, but students can actually do what they will later see as parts of sophisticated data science investigation. (Eriksson, 2017)*

Eriksson beskriver derefter som eksempel på et 'proto-data-science'-forløb i fysikundervisning. Bioingeniør Kevin Karplus kommenterer indlægget som:

*It's a nice start, but intro physics is probably not the best place to do data science. Physics is a model-driven field, where clean models do a ridiculously good job of describing real-world phenomena. Data science is more appropriate for data-driven fields (like biology and sociology), where there are no clean simple models to discover, and a big chunk of the work is coming up with approximate models that capture some important aspects of the messy data. (Eriksson, 2017)*

Denne udveksling afspejler den udviklingsfase, hvor undervisere og forskere i forskellige discipliner arbejder på at identificere opgaver og aktiviteter, som forbinder data-literacy (eller 'proto-data-science') med aktiviteter der både er uafklarede og minder om virkelighedsnære problemløsningsaktiviteter der går på tværs af discipliner.

En mulig model for didaktisk design kunne være projektet 'Thinking with Data' (TWD), der netop omhandler data-literacy på tværs af curriculum (Vahey et al., 2012; Swan et al., 2013) og som eksplicit bygger på det didaktiske design *Preparation for Future Learning* (PFL) (Bransford & Schwartz, 1999). I stedet for en traditionel tilgang, hvor eleverne først bliver instrueret i en metode og derefter anvender den på et givent problem, begynder en PFL-aktivitet med at præsentere eleverne for et problem, som de endnu ikke er i stand til at finde en 'korrekt' løsning på. Eleverne er nødt til at arbejde kreativt for at nærme sig en løsning og i denne proces opnår de en bedre forståelse af problemfeltets kontekst og udfordringer.

Denne udforskningsfase forbereder eleverne på en 'time for telling' (Schwartz & Bransford, 1998), hvor eleverne får præsenteret en generaliserbar, eksemplarisk løsning. TWD-projektet har desuden et stærkt fokus på kommunikation af løsningen ved hjælp af både tekst og visualiseringer.

Som opsamlende principper for didaktisk design til data-literacy foreslår vi at data-literacy ideelt set kan fremmes ved at bygge på

- konkrete virkelighedsnære og tværfaglige problemstillinger
- fokus på alle dimensioner af problemløsningsprocessen, ikke kun med fokus på dataanalysen men også på det domæne-specifikke, på kommunikation mm.
- 'guided discovery' af problemet og konteksten i starten, efterfulgt af undervisning med 'time for telling' (Schwartz & Bransford, 1998), sådan at nye begreber i eksempelvis statistik opleves som generaliserbare løsninger.
- 'collaborative inquiry' (e.g., Garrison & Anderson, 2003; Garrison, 2016), sådan at alle deltagerne kan forstå en kompleks sammenhæng, men også udvikle mere specialiserede færdigheder indenfor eksempelvis kodning eller kommunikation.

Vi foreslår desuden, at disse principper så vidt som muligt indrammes i et 'kontaktsøgende undervisningsrum' (Tække & Paulsen, 2016; Fung, 2017 som 'connected learning'), hvor elever eller studerende kan se sig selv som del af en større fællesskab.

## Åbne spørgsmål og udfordringer i at fremme data-literacy

At forbinde data-literacy med kompetencer til aktivt medborgerskab er et vigtigt greb for at fremhæve, at vi står foran en vanskelig opgave: vi (som undervisere, forskere, administratorer) skal selv opnå et bedre niveau af data-literacy samtidig med at vi skal være med til at fremme data-literacy hos andre.

Dette afsnit vil, for det første, problematisere medborgerskab i forbindelse med data-literacy og derefter diskutere hvilke ansvar opgaven med at fremme data-literacy tildeler undervisere i STEM- og i humanistiske discipliner.

At fremme data-literacy, forstået som et kontinuum af kompetencer, hvor elever og studerende "nærmer sig til data science" (Eriksson, 2017), er et ambitiøs projekt, der som minimum kræver at man re-orienterer og gen-reflekterer over egen undervisningspraksis og faglighed med henblik på både indholdsdimensionen såvel som metoderne. Ligeledes bør uddannelsesinstitutioner understøtte et kompetenceløft der muliggør inkorporeringen af de forskellige kompetencer der tilsammen udgør et praksisfællesskabs data-literacy.

Selve idealet om en datakyndig medborger ('citizen data scientist') som et overordnet formål med at fremme data-literacy er ambitiøs, og der er ikke

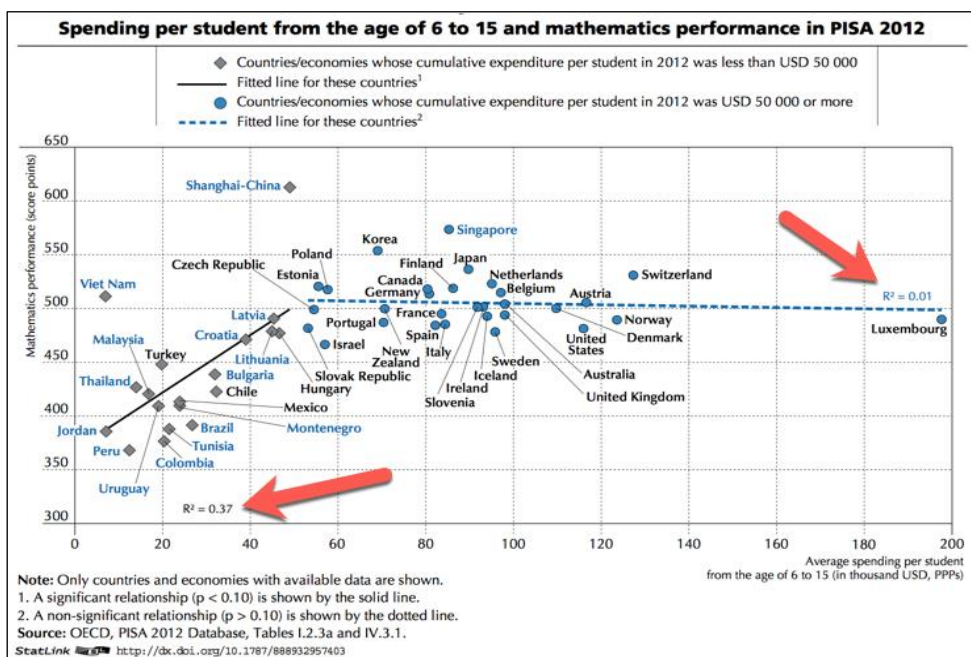
enighed om, hvorvidt det er praktisk realiserbart. Frank, Walker, Attard, & Tygel (2016) observerer eksempelvis at “The concept of the ‘armchair auditor’, a citizen who browses government data to monitor its activities and hold it to account, has not materialised”, og tilføjer at langt de fleste fortsat vil have behov for en formidler i form af journalister eller stakeholders, som udvælger, hvilke data de kan få adgang til.

Gregory Piatetsky-Shapiro er blandt skeptikerne overfor ‘citizen data scientist’-begrebet. I kølvandet af en forudsigelse om at delvis automatiseret software vil medvirke til at et øget antal af data-analyser vil blive udført af ‘citizen data scientists’ (Gartner, 2016), påpegede Piatetsky-Shapiro, at dataanalyse i hænderne på utilstrækkeligt uddannede analytikere ville være en farlig kombination. En kombination han sammenlignede med en autopilot der fungerer 90% af tiden, mens den i resten af tiden overlader kontrollen til en uøvede pilot (Piatetsky-Shapiro, 2016). Denne indvending bør tages seriøst, fordi det stadig er et åbent spørgsmål om hvorvidt det er ønskværdigt at ignorere kompleksiteten i de statistiske udregninger ved at uddelegere vigtige elementer i analysen og fortolkningen af data til softwaren. Teknologien har på sin vis medvirket til at sænke barren for at kunne påbegynde dataanalyse, fordi den simpelthen har gjort det nemmere at indsamle, sammenligne og behandle data ved populært sagt at lade softwaren og computerkraften gøre arbejdet. Men det fjerner ikke behovet for at kunne evaluere datas validitet og reliabilitet som stadig kræver indsigt i viden om metode og domæner. Faktisk er det sandsynligt at behovet for stærkere metodiske kompetencer øges i takt med omfanget af tilgængelige data.

Problemet bliver yderligere kompliceret af at viden om statistik og videnskabelige metoder er koncentreret om en forholdsvis lille gruppe af personer og at den brede befolkning, store andele af personer indenfor medier og politik og måske også en del forskere ofte er usikre om basale statistiske metoder. Psykologen Gerd Gigerenzer, økonomen Thomas Bauer og statistikeren Walter Krämer har hver måned siden 2012 analyseret en case af ‘unstatistik’ (<http://www.rwi-essen.de/unstatistik/>; se også Bauer, Gigerenzer & Krämer, 2014), hvor journalister eller forskere har begået grove fejl i indsamling, analyse og kommunikation af data, hvilket i flere tilfælde har haft en negativ indvirkning på den offentlige debat og i beslutningsprocesser. Statistik er kompliceret og selv en vigtig tekst som Mayer-Schönberger & Cukier (2013) om big-data, med over 3000 citater i Google Scholar, er blevet beskyldt af to epidemiologer for at misforstå nogle begreber om statistiske metoder (Naimi & Westreich, 2014).

At arbejde med data-literacy er med andre ord en aktivitet forbundet med en høj risiko for at tage fejl, men at undlade at arbejde med data vil være et svigt givet at data udgør en vigtig del af den offentlige debat og lægger til grund for vigtige beslutninger indenfor stort set alle politiske områder, såsom miljø, socialpolitik og sundhed.

En af forfattere af denne artikel har i flere tilfælde vist en graf som det i figur 11 til kandidatstuderende i 'Learning and Teaching for Digital Media' fra humanistiske fag, som havde valgt en 'gymnasielærer uddannelsesprofil'.



Figur 11 - Forholdet mellem udgiften pr. elev (i US dollars) og resultater fra OECD-PISA's matematiktest. Den fuldt optrukne grå tendenslinje repræsenterer lande som i 2012 brugte mindre end 50.000 USD pr. elev, mens den punkterede blå tendenslinje repræsenterer lande som bruger mere end 50.000 USD pr. elev. Markeret med røde pile er de tilhørende korrelationsværdier for de respektive tendenslinjer. (OECD, 2013, p. 41).

Grafikken fra OECD-PISA undersøgelse viser forbindelsen mellem gennemsnitlige udgifter pr. elev og resultater i matematik. Af de 20 studerende som blev præsenteret for grafen kunne kun en til to af dem forklare, hvad " $R^2 = 0.37$ " eller " $R^2 = 0.01$ " betød (markeret med røde pile i figur 11). Det var dog ikke ensbetydende med, at de øvrige studerende ikke kunne forstå grafen, men de kunne ikke forklare den ved hjælp af



statistiske begreber. Vi synes at det ville være et rimeligt læringsmål, at gymnasieelever og førsteårsstuderende fra alle fag på de videregående uddannelser skulle være i stand til at forklare en sådan graf og vide hvordan man tolker en korrelationsværdi som  $R^2$ . Det kan lyde ambitiøst, men risikoen ved ikke at besidde en sådan viden vil være en forringelse af egen dømmekraft og myndighed som medborgere, på lige fod med at mangle læsefærdighed på engelsk. Problemet vil antagelig formindskes eller løses i takt med, at nye studerende tilegner sig statistiske og data-analytiske kompetencer i gymnasiet, men i en overgangsfase bør det overvejes, hvordan de enkelte uddannelser kan sikre et fælles basisniveau i data-literacy.

Samtidigt er det også vigtigt, at disse elever og studerende kan modstå overfortolkninger af sådan en grafik. I takt med at data bliver byggesten af hegemonisk diskurs ('dataisme' ifølge Harari, 2016) kan man kun forvente, at data også i stigende grad vil blive anvendt til at lyve og manipulere.

Allerede fra 2012 blev det klart, hvor megen viden et firma potentielt har om hver enkelt kunde - bl.a. om hvorvidt en kvindelig kunde var gravid - ved at analysere hendes købshistorik med henblik på at målrette tilbud og reklamer (Duhigg, 2012). I marts 2018 blev det afsløret, hvordan det britiske firma Cambridge Analytica Ltd har indsamlet data fra ca. 50 millioner Facebook-brugere med flere tusinde indikatorer på den enkelte bruger med henblik på at sende målrettede politiske budskaber i valgkampen forud for det amerikanske præsidentvalg i 2016 (Cadwalladr & Graham-Harrison, 2018, marts d. 17); det samme firma har desuden spillet en rolle ved valgkampagner i Europa, Indien og Afrika (Channel 4, 2018). Men spørgsmålet er om data-literacy virkelig kan hjælpe ved at modstå sådanne forsøg på datadrevet manipulation?

For at forstå teknikken bag brugerprofilering og at være i stand til at beskytte egne personlige data bedst muligt kan ses som elementer af data-literacy, men det er vigtigt at understrege at det er viden om politik og kritisk dømmekraft, som primært kan hjælpe med at afsløre løgne og manipulation i den offentlige debat. En studie om spredning af sande vs. falske nyheder på nettet konkluderer at "false news spreads more than the truth because humans, not robots, are more likely to spread it" (Vosoughi, Roy, & Aral, 2018). Med andre ord fungerer teknologien til at sprede information (herunder misinformation) som en 'forstærker' (Toyama, 2015) af individuelle og sociale tilbøjeligheder. Rose Luckin har overbevisende argumenteret for, at uddannelsen til et gennemteknologiseret samfund i højere grad bør fokusere på livslang kompetenceudvikling af 'epistemisk viden' og 'bevidstheden om egen bias',

så menneskene ikke bliver overhalet og overflødiggjort af computere (Luckin, 2018). Data-literacy kan og bør integreres i en uddannelsesmæssig sammenhæng med disse overordnede læringsmål. Men uddannelsesinstitutioner kan ikke løfte opgaven alene, medierne spiller også en væsentlig rolle.

Det kan ikke forventes at et flertal af medierne vil finde incitament i at fremme læsernes data-literacy (Kayser-Bril, 2016), men samtidigt kan vi mere optimistisk registrere en stigende interesse for at forstå omverden med udgangspunkt i fakta og data, hvilket bl.a. afspejles i initiativer som den *European Investigative Journalism Conference and Dataharvest* (journalismfund.eu, s.n., etableret i 2011), den 2017-etableret *Nordic Data Journalism Conference* ([noda2017.dk](http://noda2017.dk)), eller TV-succesen af faktatjek-programmet "Detektor" (Andreassen, 2011).

I takt med at datavidenskab bliver etableret på landets universiteter er det værd at overveje, hvordan de studerende i de nye fag kan inddrages i 'hackathon-like' initiativer, hvor de udfordres af andre discipliner og hvor de bringes i forbindelse andre domæner, der vil styrke deres forståelse af og behovet for samarbejde med personer med stor praktisk og teoretisk områdespecifik viden. Det ville være en oplagt mulighed for at tænke 'connected curricula' hvor man udnytter de datavidenskabelige studerendes specialviden til at understøtte udviklingen af mere generelle kompetencer hos studerende fra andre fag som ikke uddannes til specialister indenfor feltet (Fung, 2017).

## Foreløbig konklusion

Vi befinder os i en fase, hvor kompetencer til 'avancerede ikke-specialister' (som ville få "a whiff of data science") i første omgang udbygges lokalt hos interesserede ildsjæle ofte med assistance fra kyndige kollegaer med baggrund i eksempelvis matematik eller samfundsfag. Hvordan man løfter opgaven med at fremme data-literacy på nationalt plan ligger udenfor denne artikels fokus, men det er en politiske proces som ofte tager mange år. Vi har valgt at fokusere på en bottom-up tilgang, hvor det er underviserne – individuelt eller i fællesskab – der ser behovet for at inddrage data-literacy som en del af deres undervisning.

At inddrage data-literacy som læringsmål i egen undervisning vil for mange undervisere betyde at man arbejder på grænsen af sin egen faglige kompetence og i visse tilfælde overskride sin egen faglighed. Derfor er det vigtigt at man i fællesskab understøtter hinanden. Man kan eksempelvis gøre brug af gensidig supervision eller peer learning (Boud et al., 1999; Pedersen & Caviglia, 2019). Som nævnt tidligere kan det være produktivt at

anskue data-literacy som en sammensat kompetence, ikke blot med hensyn til de enkelte vidensområder, men også til den sociale sammensætning af en gruppe af personer der tilsammen kan siges at besidde data-literacy. De cases vi har præsenteret viser den praktiske nødvendighed af en heterogen gruppe med supplerende og overlappende kompetencer der ved fælles indsats kan løse komplekse problemstillinger. I en læringsammenhæng bør man derfor også tilstræbe denne sammensathed og udvælge problemstillinger der på en substantiel måde går på tværs af fag og som derfor kalder på kompetencer man i et givent praksisfællesskab ikke altid har til rådighed, men må supplere andetsteds fra. Derfor vil der ofte være behov for at åbne uddannelsesinstitutionerne op i forhold til omverden.

For at understøtte en brede bevægelse mod inklusionen af data-literacy på tværs af fag kan opbygningen af et netværk, hvor underviserne kan diskutere problemer og dele materialer, blive en nødvendighed.

Denne artikel kan være et bidrag i denne retning ved at påpege vigtigheden af at forstå data-literacy som sammensat kompetence der kalder på nye samarbejder mellem faglige områder indenfor matematik, naturvidenskab, humaniora og samfundsvidenskab.

Til lærere i matematik vil vi videreføre matematikprofessors Arthur Benjamins anbefaling i TED-talk format om at undervise statistik fremfor regning (Benjamins, 2009), samt en anbefaling om at inddrage eksempler og problemer fra både naturfag og samfundsfag, geografi, historie og endda sprogfag (se eks. Leńko-Szymańska & Boulton, 2015), ideelt i tværfaglige forløb.

Til kollegaer med humanistisk baggrund anbefaler vi at udvise 'epistemologisk ydmyghed' og medvirke til at styrke egen data-literacy gennem tværfaglige samarbejder og muligvis efteruddannelse. Det er vanskeligt og det tager tid. Det handler ikke bare om at tilegne sig nogle overordnede principper. For at kunne komme i dialog med andre om konkrete problemstillinger, hvor data er en del af billedet, er det ikke nok at besidde solid områdespecifik viden: der er også brug for metodologiske og praktiske kompetencer for meningsfuldt at kunne diskutere en løsning. Til gengæld har humanister den rette baggrund for at kunne bidrage til at "forstå og mobilisere sprogets magt i formidlingen af data" (*Beyond data-literacy*, 2015), hvilket også er en væsentlig udfordring i tilrettelæggelsen af programmer og kurser indenfor data-literacy (id.).

Vi konkluderer ved at vende tilbage til Wolff et al.'s (2016) forslag om at se data-literacy som en kompetence, der først og fremmest kan tilskrives til et undersøgende fællesskab frem for et enkelt individ, og til de afsluttende

henstillinger i *Beyond data-literacy* (2015) om hvordan man fremmer data-literacy. At fremme data-literacy involverer nemlig at bestyrke menneskers evne til at navigere i data med udgangspunkt i deres nuværende situation og position i samfundet på måder der føles meningsfuld og udslagsgivende for dem. Til formålet kan man skabe og tilbyde en flerhed af adgange og muligheder for personer med forskellige data-literacy behov og forskellige kompetencemæssige udgangspunkter til meningsfuldt at kunne samarbejde.

## Referencer

- Allingstrup, M., & Wissing, M. (2016). Dansk studerende offentliggør 70.000 brugere af datingsite. *DR Nyheder*, d. 16. maj.  
<https://www.dr.dk/nyheder/penge/dansk-studerende-offentliggør-70000-brugere-af-datingsite>
- Andreassen, Andreas Marckmann (2011). Detektor - nu også på DR2. *Journalisten*, d. 13. oktober. <http://journalisten.dk/detektor-nu-også-pa-dr2>
- Bauer, T., Gigerenzer, G., & Krämer, W. (2014). *Warum dick nicht doof macht und Genmais nicht tötet: Über Risiken und Nebenwirkungen der Unstatistik*. Frankfurt: Campus.
- Benjamins, A., (2009). Teach statistics before calculus! *TED*, februar.  
[https://www.ted.com/talks/arthur\\_benjamin\\_s\\_formula\\_for\\_changing\\_math\\_education](https://www.ted.com/talks/arthur_benjamin_s_formula_for_changing_math_education)
- Bereiter, C., & Scardamalia, M. (2005). Technology and literacies: from print literacy to dialogic literacy. In N. Bascia, A. Cumming, A. Datnow, K. Leithwood, & D. Livingstone (Eds.), *International handbook of educational policy* (pp. 749–762). Dordrecht, The Netherlands: Springer.  
[https://doi.org/10.1007/1-4020-3201-3\\_39](https://doi.org/10.1007/1-4020-3201-3_39)
- Bell, G., Hey, T., & Szalay, A. (2009). Beyond the Data Deluge. *Science*, 323(5919), 1297–1298. <https://doi.org/10.1126/science.1170411>
- Beyond data-literacy (2015). *Beyond data-literacy: Reinventing Community Engagement and Empowerment in the Age of Data*. Data-Pop Alliance White Paper Series. Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute) and Internews. September 2015.
- Boud, D., Cohen, R., & Sampson, J. (1999). Peer Learning and Assessment. *Assessment & Evaluation in Higher Education*, 24(4), 413–426.  
<https://doi.org/10.1080/0260293990240405>

- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24, 61–100.
- Breinstrup, T. (2017). Trods større optag skriger erhvervslivet stadig på IT-specialister. *Berlingske Business*, d. 28. juli. Hentet 01.02.2018 fra <https://www.business.dk/content/item/432023>
- Bruschi, M. (2017). L'Italia delle slot. Ovvero: di Foia, pdf, database e inchieste collettive [Slot-machines in Italy: about Freedom-of-information-act, pdf, database and collective inquiry]. *Medium/Visuallab*, d. 29. december. Hentet 01.02.2018 fra <https://medium.com/visuallab/litalia-delle-slot-ovvero-di-foia-pdf-database-e-inchieste-collettive-562686eb2f02>
- Cadwalladr, C. & Graham-Harrison, E. (2018, marts d. 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. Retrieved from <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- Caspersen, M.E. (2017). Computational Thinking. I J. Dolin, G. Holten Ingerslev & Hanne Sparholt Jørgensen (Red.), *Gymnasiepædagogik. En grundbog*. København: Hans Reitzels, s. 470-478.
- Channel 4 (2018). Data, Democracy and Dirty Tricks. *4 News*, d. 19 marts. Website. <https://www.channel4.com/news/data-democracy-and-dirty-tricks-cambridge-analytica-uncovered-investigation-expose>
- Citizen-Statistician (u.d.). What's a Citizen Statistician?. Blog post. Hentet 01.02.2018 fra <http://citizen-statistician.org/whats-a-citizen-statistician/>
- Clio Infra (2016). Website. <https://www.clio-infra.eu/>
- Christensen, V. Tornhøj (red.) (2016). *PISA 2015 – Danske unge i en international sammenligning*. København: KORA. Hentet 01.02.2018 fra <https://uvm.dk/-/media/filer/uvm/udd/folke/pdf16/dec/161206-pisa-2015-danske-unge-i-en-international-sammenligning.pdf>
- Conway, Drew (2013). The data science Venn diagram. Blog post, March 26th. <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>
- Duhigg, C. (2012). How Companies Learn Your Secrets. *The New York Times*, d. 16. februar. <https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>

- Eriksson, Tim (2017). Smelling like data science. Blog post  
<https://bestcase.wordpress.com/2017/02/21/smelling-like-data-science/>
- Frank, M., Walker, J., Attard, J., Tygel, A. (2016). Data-literacy: what is it and how can we make it happen? Editorial. *The Journal of Community Informatics*, 12(3), 4-8.
- Fung, D. (2017). *A Connected Curriculum for Higher Education*. London: UCL Pres. <https://doi.org/10.14324/111.9781911576358>
- Gapminder (s.d.). Website. <https://www.gapminder.org/>
- Gapminder Public (s.d.). Website.  
<https://drive.google.com/drive/folders/0B3DxAuXLkRM6flhmbgl2WkhFSnVLQ1dTdFNkellTNOZTaORhTXA3dFUtcDhZd2ZQUU1HazQ>
- Garrison, D. R. (2016). *Thinking collaboratively: Learning in a community of inquiry*. London, UK: Routledge.
- Garrison, D. R., & Anderson, T. (2003). *E-learning in the 21st century: A framework for research and practice*. London, UK: RoutledgeFalmer.  
<https://doi.org/10.4324/9780203166093>
- Gartner (2016). Gartner Says More Than 40 Percent of Data Science Tasks Will Be Automated by 2020. Blogpost, hentet d. 01-02-2018 fra  
<https://www.gartner.com/newsroom/id/3570917>
- GEDI (2017). L'Italia delle slot [Spilleautomaternes Italien]. Website. GEDI Gruppo Editoriale. Hentet 31.01.2018 fra  
<http://lab.gruppoespresso.it/finegil/2017/italia-delle-slot/>
- Gray, J., Bounegru, L. & Chambers, L. (2012). *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*. Open Knowledge Foundation. Hentet 01.02.2018 fra  
<http://datajournalismhandbook.net/1.0/en/>
- Harari, Y. N. (2016). *Homo Deus: A Brief History of Tomorrow*. New York: Random House.
- Hey, T., Tansley, S., & Tolle, K. (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Redmond, WA: Microsoft Research. Hentet 01.02.2018 fra <https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery/>
- Horning, A. S. (2007). Reading across the curriculum as the key to student success. *Across the Disciplines*, 4, d. 14. maj. Hentet 01.02.2018 fra  
<https://wac.colostate.edu/atd/articles/horning2007.cfm>

- Hou, Y., & Wang, D. (2017). Hacking with NPOs: Collaborative Analytics and Broker Roles in Civic Data Hackathons. *Proceedings of the ACM on Human-Computer Interaction*, 1 (CSCW), 53:1-53:16.  
<https://doi.org/10.1145/3134688>
- Innes, J. E., & Booher, D. E. (2010). *Planning with complexity: An introduction to collaborative rationality for public policy*. London, UK: Routledge.
- Kayser-Bril, N. (2016). Don't ask much from data-literacy. *The Journal of Community Informatics*, 12(3), 217—222.
- Leńko-Szymańska, A. & Boulton, A. (Eds) (2015). *Multiple Affordances of Language Corpora for Data-driven Learning*. Studies in Corpus Linguistics (SCL), 69.
- Lindeskov, L. & Weng, P. (2010). *15 matematikopgaver i PISA: et materiale for matematiklærere og andre med interesse i grundskolens matematikundervisning*. København: Danmarks Pædagogiske Universitetssskole, Aarhus Universitet, 2010.
- Luckin, R. (2018). *Machine Learning and Human Intelligence: The Future of Education for the 21<sup>st</sup> Century*. London: UCL Institute of Education Press.
- Lund, H. Romme (2017). Literacy på dansk. *Folkeskolen.dk*, d. 27. januar.  
<https://www.folkeskolen.dk/601653/literacy-paa-dansk>
- Markauskaite, L., & Goodyear, P. (2016). *Epistemic fluency and professional education: innovation, knowledgeable action and actionable knowledge*. Dordrecht, The Netherlands: Springer. <https://doi.org/10.1007/978-94-007-4369-4>
- Mau, Steffen (2017). *Das metrische Wir - Über die Quantifizierung des Sozialen*. Berlin: Suhrkamp.
- Mayer-Schönberger, V. & Cukier, K. (2013) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York: Houghton Mifflin.
- McCook, A. (2016). Publicly available data on thousands of OKCupid users pulled over copyright claim. *Retraction Watch*, d. 16. maj.  
<http://retractionwatch.com/2016/05/16/publicly-available-data-on-thousands-of-okcupid-users-pulled-over-copyright-claim/>
- Naimi, A. I., & Westreich, D. J. (2014). Big Data: A Revolution That Will Transform How We Live, Work, and Think. *American Journal of Epidemiology*, 179(9), 1143–1144.  
<https://doi.org/10.1093/aje/kwu085>
- Nepper Larsen, Steen (2017). Internettets 'stille giganter' er på jagt efter de gode tal. *Information*, d. 11. november.

<https://www.information.dk/moti/anmeldelse/2017/11/internettets-stille-giganter-paa-jagt-gode-tal>

OECD (2009). *Take the test. Sample questions from OECD's PISA assessments*. Paris: OECD Publishing.

<https://www.oecd.org/pisa/pisaproducts/Take%20the%20test%20e%20book.pdf> (p. 297)

OECD (2013). *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices (Volume IV)*. Paris: OECD Publishing.

<http://dx.doi.org/10.1787/9789264201156-en>

OECD (2016), *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*. Paris: OECD Publishing.

<http://dx.doi.org/10.1787/9789264255425-en>

Open Data (n.d.). *Wikipedia*. Hentet d. 01.02.2018 fra

[https://en.wikipedia.org/wiki/Open\\_data](https://en.wikipedia.org/wiki/Open_data)

Pawlowski, S. D. & Robey D. (2004). Bridging user organizations: Knowledge brokering and the work of information technology professionals. *MIS Quarterly* 28 (4), pp. 645-672.

Parkkinen, J. (2015). Data scientist - statistician, programmer, consultant and visualizer? My personal attempt to define a data scientist. Blog post.

<http://ouzor.github.io/blog/2015/02/02/data-science-definition.html>

Pedersen, A. Y., & Caviglia, F. (2019). Data Literacy as a Compound Competence. In T. Antipova & A. Rocha (Eds.), *Digital Science* (Vol. 850, pp. 166–173). Springer International Publishing.

[https://doi.org/10.1007/978-3-030-02351-5\\_21](https://doi.org/10.1007/978-3-030-02351-5_21)

Pentland, A. (2013). The Data-Driven Society. *Scientific American*, 309(4), 78–83. <https://doi.org/10.1038/scientificamerican1013-78>

Piatetsky-Shapiro, Gregory (2016). The Mirage of a Citizen Data Scientist.

Blog post. <https://www.kdnuggets.com/2016/03/mirage-citizen-data-scientist.html>

Presser, S. & Stinson, L. (1998). Data Collection Mode and Social Desirability Bias in Self-Reported Religious Attendance. *American Sociological Review*, 63:137-45.

Resnik, B. (2016). Researchers just released profile data on 70,000 OkCupid users without permission. *Vox*, d. 12. maj.

<https://www.vox.com/2016/5/12/11666116/70000-okcupid-users-data-release>

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., ... Wuetherick, B. (2015). Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report. Retrieved from



[http://www.mikesmit.com/wp-content/papercite-data/pdf/data\\_literacy.pdf](http://www.mikesmit.com/wp-content/papercite-data/pdf/data_literacy.pdf)

Roser, Max (2018). Our World in Data. Web publication.  
<https://ourworldindata.org/>

Roser, Max & Ortiz-Ospina, Esteban (2017). World Population Growth. Published online at *OurWorldInData.org*. Retrieved 31.01.2018 fra <https://ourworldindata.org/world-population-growth>

Rosling, Hans (2006). The best stats you've ever seen. *TED Talks*, february 2006.  
[https://www.ted.com/talks/hans\\_rosling\\_shows\\_the\\_best\\_stats\\_you\\_ve\\_ever\\_seen](https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen)

Rosling, H., Rosling, O. & Rosling Rönnlund, A. (2018). *Factfulness: Ten Reasons We're Wrong About the World — and Why Things are Better Than You Think*. New York: Flatiron Books.

Samuels, R. (2008). Auto-Modernity after Postmodernism: Autonomy and Automation in Culture, Technology, and Education. In M. Tara (Ed.), *Digital Youth, Innovation, and the Unexpected* (pp. 219-240). Cambridge, MA: The MIT Press.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, 16, 475–523.

Shapin, S. & Schaffer, S. (1985). *Leviathan and the Air-Pump: Hobbes, Boyle and the Experimental Life*. Princeton, NJ: Princeton University Press

Sharif, Raed M. & and Van Schalkwyk, F. (2016). Special Issue on Open Data for Social Change and Sustainable Development. *The Journal of Community Informatics* 12(2).

Stephenson, E. og Schifter Caravello, P. (2007). Incorporating data-literacy into undergraduate information literacy programs in the social sciences: A pilot project. *Reference Services Review*, 53(4), 525-540.  
<https://doi.org/10.1108/00907320710838354>.

Stott, Andrew. 2014. Open Data for Economic Growth. World Bank. Hented 02.02.2018 fra.  
<http://www.worldbank.org/content/dam/Worldbank/document/Open-Data-for-Economic-Growth.pdf>.

Swan, K., Vahey, P., van 't Hooft, M., Kratcoski, A., Rafanan, K., Stanford, T., Yarnall, L., & Cook, D. (2013). Problem-based Learning Across the Curriculum: Exploring the Efficacy of a Cross-curricular Application of Preparation for Future Learning. *Interdisciplinary Journal of Problem-Based Learning*, 7(1). <https://doi.org/10.7771/1541-5015.1307>

- Sweller, J., & Cooper, G. A. (1985). The Use of Worked Examples as a Substitute for Problem Solving in Learning Algebra. *Cognition and Instruction*, 2(1), 59–89. [https://doi.org/10.1207/s1532690xci0201\\_3](https://doi.org/10.1207/s1532690xci0201_3)
- Stella, G. A. (2012). Il Tar multa il sindaco anti slot-machine [En bode til borgmesteren, der kæmper mod spillemaskinerne]. *Corriere della Sera*, d. 12. marts. Hentet 01.02.2018 fra [http://www.corriere.it/cronache/12\\_marzo\\_23/stella-il-sindaco-anti-slot-machine\\_c5b92a22-74af-11e1-9cbf-6c08e5424a86.shtml](http://www.corriere.it/cronache/12_marzo_23/stella-il-sindaco-anti-slot-machine_c5b92a22-74af-11e1-9cbf-6c08e5424a86.shtml)
- Talamo, G. & Manuguerra, G. (2016). The gambling sector: A socio-economic analysis of the case of Italy. *Eastern European Business and Economics Journal*, 2(4), 315-330.
- Teknologiforståelse valgfag (2017). Teknologiforståelse valgfag (forsøg) – Fælles Mål og læseplan. Undervisningsministeriet. Hentet 01.02.2018 fra <https://www.emu.dk/modul/teknologiforst%C3%A5else-valgfag-fors%C3%B8g-%E2%80%93-f%C3%A6lles-m%C3%A5l-og-l%C3%A6seplan>
- Thomsen, Anders Topp (2017). Danmark får 19 nye uddannelser. *DR Nyheder*, d. 12. december. Hentet 01.02.2018 fra <https://www.dr.dk/ligetil/danmark-faar-19-nye-uddannelser>
- Tourangeau, R., Rips Lance J. & Rasinski Kenneth (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Toyama, K. (2015) *Geek Heresy. Rescuing Social Change from the Cult of Technology*. Philadelphia (PA): Foreign Affairs.
- Tække, J. & Paulsen, M. (2016). *Undervisningsfællesskaber og læringsnetværk i den digitale tidsalder*. København: Unge Pædagoger.
- UVM (2016). Aftale mellem regeringen, Socialdemokraterne, Dansk Folkeparti, Liberal Alliance, Det Radikale Venstre, Socialistisk Folkeparti og Det Konservative Folkeparti om styrkede gymnasiale uddannelser. Undervisningsministeriet, d. 3. juni. Hentet 01.02.2018 fra <https://www.uvm.dk/-/media/filer/uvm/udd/gym/pdf16/jun/160603-styrkede-gymnasiale-uddannelser.pdf>
- UMV (2017a). Læreplaner - Informatik B - htx2017. Undervisningsministeriet. Hentet 01.02.2018 fra <https://www.uvm.dk/-/media/filer/uvm/gym-laereplaner-2017/hhx/informatik-b-hhx-og-htx-august-2017.pdf>.
- UVM (2017a). Generelt om læreplaner. Undervisningsministeriet. Hentet 01.02.2018 fra <https://uvm.dk/gymnasiale-uddannelser/fag-og-laereplaner/om-laereplaner>

- UVM (2017b). Forsøg: Nyt valgfag skal ruste eleverne til mødet med den digitale udvikling. Undervisningsministeriet, d. 14. august. Hentet 01.02.2018 fra <https://uvm.dk/aktuelt/nyheder/uvm/2017/august/170811-forsog--nyt-valgfag-skal-ruste-eleverne-til-moedet-med-den-digitale-udvikling>
- UVM (2017c). Teknologiforståelse valgfag (forsøg) – Fælles Mål og læseplan. Undervisningsministeriet. Hentet 01.02.2018 fra <https://www.emu.dk/modul/teknologiforst%C3%A5else-valgfag-fors%C3%B8g-%E2%80%93-f%C3%A6lles-m%C3%A5l-og-l%C3%A6seplan>
- Vahey, P., Rafanan, K., Patton, C., Swan, K., Hooft, M. van 't, Kratcoski, A., & Stanford, T. (2012). A cross-disciplinary approach to teaching data-literacy and proportionality. *Educational Studies in Mathematics*, 81(2), 179–205. <https://doi.org/10.1007/s10649-012-9392-z>
- Vermanen, J. (2012). Harnessing External Expertise Through Hackthons. In Gray, Bounegru & Chambers (2012). [http://datajournalismhandbook.net/1.0/en/in\\_the\\_newsroom\\_6.html](http://datajournalismhandbook.net/1.0/en/in_the_newsroom_6.html)
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Wenger, E., (1999). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge: Cambridge University Press.
- Wegerif, R. (2016). Applying dialogic theory to illuminate the relationship between literacy education and teaching thinking in the context of the Internet Age. *L1-Educational Studies in Language and Literature*, 16, 1–21. <https://doi.org/10.17239/L1ESLL-2016.16.02.07>
- Wild, C. J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review/Revue Internationale de Statistique*, 223–248.
- Wolff, A., Kortuem, G., & Cavero, J. (2015). Towards smart city education. *2015 Sustainable Internet and ICT for Sustainability (SustainIT)*, 14-15 april, pp. 1–3. <https://doi.org/10.1109/SustainIT.2015.7101381>
- Wolff, A., Gooch, D., Cavero Montaner, J. J., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of data-literacy for a Data-driven Society. *The Journal of Community Informatics*, 12(3). <http://ci-journal.net/index.php/ciej/article/view/1286>