 En verden fuld af data

Kilde: <https://digitaldannelse.systime.dk/?id=166&L=1>

En masse nye datafænomener er begyndt at dukke op, en konsekvens af den digitale tidsalder, vi lever i. Begreber som big data, big data analytics, algoritmer, åbne data og mange andre er ved at blive almindelige ord og vendinger i det danske sprog. I vores samfund, skal vi altså blive bedre til at anvende data i en værdiskabende kontekst, som kan hjælpe os med at træffe mere intelligente beslutninger. Datafænomenerne kan altså være med til at skabe mærkbare fremskridt for det enkelte menneske og samfundet som helhed, hvis vi forstår at gøre brug af det.

[](https://digitaldannelse.systime.dk/fileadmin/_processed_/0/4/csm_14_smart_city_c79577a19b.jpg)

iStockphoto.com/chombosan / Systime

I forhold til at anvende data til forskellige formål skal man være bevidst om, hvorvidt det giver mening at anvende de givne data, om det skaber værdi, og om man overskrider en grænse etisk og moralsk ved anvende disse data?

Situationen er den, at muligheden for at overvåge og følge alle, er til stede. Vi kan modtage alle verdens nyheder, men giver det mening? Vi kan udstyre alle med en GPS-tracker, der informerer alle i verden om, hvor de er og tracke deres bevægelsesmønstre, men giver det mening? Vi kan løse langt flere kriminalsager, hvis vi overvåger og indsamler relevant data om borgerne og deres adfærd, men overskrider det en grænse i forhold til vores privatliv? Der er et ubegrænset antal anvendelsesmuligheder, når det gælder opsamling af data og anvendelsen heraf, men hvilke data skal vi vælge, og hvornår overskrider vi en grænse, som ikke er acceptabel?

# DEFINITION AF DATA KEY

Data er enhver repræsentation af fakta eller ideer på en formaliseret måde, som kan kommunikeres eller manipuleres ved en eller anden proces. En ting eller begivenhed er ikke i sig selv data, men bliver det først, når den indgår i en proces, hvori dens repræsentation af fakta og ideer er det afgørende.

Kilde: Datalogisk Institut (2010)

Datatyper kan deles op i to hovedkategorier: struktureret data og ustruktureret data.

**Struktureret data** kan eksempelvis være produktinformationer, butiksinformationer, kontaktinformationer og mange andre egenskaber vedrørende personer, beløb, virksomheder, steder, anmeldelser, opskrifter og meget mere, som kan sættes direkte ind i et excel-ark.

**Ustruktureret data** omfatter billeder, video, tale mm., hvilket alt sammen er data, som først skal bearbejdes, før det kan omsættes til værdiskabende viden.

Der findes fire typer af gennemgående data, som hører til ovennævnte kategorier: aktivitetsdata, samtaledata, billede- og videodata samt sensordata.

# Aktivitetsdata KEY

Aktivitetsdata omfatter simple aktiviteter som at lytte til musik, læse en bog eller se serier på Netflix, da det i dag genererer data, fordi det ofte foregår via nettet. Digitale musikafspillere (Spotify, iTunes osv) indsamler således dine aktiviteter i forhold til musiksmag og frekvens og anvender dem til at give dig de bedste bud på "Denne playliste er noget for dig" osv.

Mange ejer i dag en smartphone, og den er en af de største kilder til indsamling af enorme mængder information. Den indsamler og gemmer data om, hvordan du anvender den. Hvilke apps du åbner på daglig basis, og hvor du befinder dig, når du åbner dine apps osv. Den ved fx også hvor langt du har til skole, og at du hver dag kl. 7.45 bevæger dig ad den samme rute.

Din webbrowser indsamler også information om, hvad du søger på, hvilke ord du hyppigst anvender, hvor lang tid du bruger på hvert website, med hvilken frekvens du besøger siderne og hvilke ligheder, der er blandt de sider du besøger.

Eksempelvis indsamler Danske Banks Mobilepay data om, hvor du handler, hvad du køber, og hvor høje dine beløb er, når du anvender deres betalingsform.

Det er faktisk svært at forestille sig en aktivitet, som i dag ikke generer data, når det involverer en smartphone. Det er selvfølgelig ikke al data, som må anvendes, og en del af det er beskyttet, men mulighederne for at bruge det kommercielt er enorme.

[](https://digitaldannelse.systime.dk/fileadmin/_processed_/4/c/csm_iStock_83171141_red_e515248d50.jpg)

iStockphoto.com/SolStock

# Samtaledata KEY

Vores samtaler kan i dag risikere at blive digitalt optaget og opbevaret. Det startede med e-mails, men i dag efterlader næsten alle vores samtaler en eller anden form for digitalt spor. Tænk på alle de samtaler, du har haft på Facebook via Messenger eller på Instagram, hvor du kommenterer alverdens billeder og på Twitter, hvor du poster små Tweets.

Når Facebook ændrer sine retningslinjer og dine rettigheder, bliver det ofte diskuteret, hvilke vilkår der egentlig ændres – om det nu er fordi, de opbevarer og optager alle Facetime opkald osv.?

Afslutningsvis skal nævnes, at al taledata nemt kan transformeres til tekst, således det bliver lettere at analysere og identificere mønstre i disse ustrukturerede data.

# Billede- og videodata KEY

Tænk på alle de billeder, som hver dag tages med smartphones og digitale kameraer.

Billeder gemmer vi ikke længere i vores fotoalbum, vi bruger i stedet sociale platforme til at dele vores minder med omverdenen. Statistikker viser, at vi uploader og deler omkring 100.000 billeder på sociale medier hvert sekund (Instagram, Facebook, Flickr osv). Det samme gælder video, hvor der uploades uanede mængder på Youtube hvert minut.

Al dette er med til at generere enorme mængder data, som kan gøre virksomheder klogere på forbrugsmønstre, trends og tendenser. Disse data er ustrukturerede data, da billeder og video ikke kan sættes direkte ind i et excel ark, og det kræver derfor en større indsats at trække information ud af disse data.

[](https://digitaldannelse.systime.dk/fileadmin/_processed_/1/b/csm_iStock_44237030_red_da974020f0.jpg)

iStockphoto.com/Maxiphoto

# Sensordata KEY

Der findes sensorer rigtig mange steder i omgivelserne omkring os. Disse sensorer indsamler og deler relevante data. Det kan være alt fra trafiksensorer, lyssensorer, affaldssensorer osv.

Men sensorer placeres ikke kun på fysiske eller statiske ting. I dag indeholder en iPhone en global positionssensor, som hele tiden måler, hvor du befinder dig. En iPhone har også et accelerometer, som måler hastighed og retning, når du rejser, og den har også en sensor, som måler din puls, og som kan lagre dit fingeraftryk. Vi kommer derfor til at opleve en stigende brug af sensordata, som kan skabe stor værdi i mange forskellige sammenhænge.

Det, som har ændret sig markant fra tidligere, er, at sensorer i dag kan købes til en meget lav pris, og derfor er det let at implementere og anvende dem.

# Big data

I takt med at vi efterlader os et øget antal digitale spor, er big data blevet et kendt begreb. Den grundlæggende tanke bag big data er, at alt, hvad vi foretager os i dag, efterlader en form for digitalt spor (data), som vi (og andre) kan anvende og analysere og derved blive klogere på mønstre, adfærd osv. Eftersom alt hvad vi foretager os ofte er bundet op på digitale løsninger, indsamles der enorme mængder af data, som benævnes med big data.

**DEFINITION AF BIG DATA**KEY

Big data er ekstremt store datasæt, som kan analyseres for at identificere mønstre, trends og associationer - specielt relateret til menneskelig adfærd og interaktioner.

Data kan komme mange steder fra, og vi ved, at der bliver genereret helt ufattelige mængder data hver evigt eneste dag.

Data kan have mange forskellige kilder, som eksempelvis:

* Data opsamlet af målere i en produktion
* Data skabt via sociale medier som Twitter, Facebook og LinkedIn
* Data skabt ved brug af apps på smartphones og tablets
* Data fra store mængder af transaktioner
* Data opsamlet via bevægelse, som trackes på din smartphone

Når big data analyseres og bliver til værdiskabende data, anvender man begrebet big data analytics. Det dækker altså bredt over indsamling, opbevaring, analyse, processering og fortolkning af big data. Big data analytics refererer til vores evne til at skabe værdi ud af de stor mængder data, vi i dag kan få adgang til.