



GENERATION DATA

USING DATA FOR PROFIT

Noter 1.0



Erasmus+

This project has been
funded with support from
the European Commission



EXPOSITION

1. EN KORT HISTORIE OVER DATA

Data er ikke nye. Men det bliver så stort og så allestedsnærværende, at det har konsekvenser for alle, uanset om vi vil have det eller ikke. Det transformerer den måde, vi gør forretning på, og påvirker de fleste andre dele af vores liv.

2. HVAD ER STORE DATA?

Så nu ved vi, hvad der har ændret sig, hvad er big data?

- Den officielle definition er *"Data af en meget stor størrelse, typisk i det omfang, at dens manipulation og styring giver væsentlige logistiske udfordringer."* Oxford English Dictionary, 2013. Men denne definition har begrænsninger, fordi der kun fokuseres på tekniske aspekter og præsenterer data som problematiske.
- En alternativ definition: *"Big data er informationsaktiver af høj volumen, høj hastighed og stor variation, der kræver omkostningseffektive, innovative former for informationsbehandling for forbedret indsigt og beslutningstagning."* Gartner, 2014 ¹ Dette giver os et større billede. Den grundlæggende idé bag udtrykket "Big Data" er derfor, at alt, hvad vi gør, i stigende grad efterlader et digitalt spor, som vi kan bruge og analysere. Big Data henviser derfor til selve dataene, men også til vores evne til at gøre brug af stadig stigende datamængder.
- Nogle mennesker går videre: *"Big data er en ny holdning fra virksomheder, nonprofitorganisationer, offentlige agenturer og enkeltpersoner, der ved at kombinere data fra flere kilder kan føre til bedre beslutninger."* Gil Press, bidrager til Forbes. ²

2.1 De 5 Vs Big Data

Hvad der er ændret om data i de sidste 10 år, hvilket gør det til et varmt emne, er 5V'erne: VOLUME - størrelsen af data, der genereres.

VELOCITY - Den hastighed, hvormed data genereres og aggregeres. VARIETY - De typer data, der er tilgængelige for os. VERACITY - Rigtigheden eller pålideligheden af dataene.

VALUE - I hvilket omfang data genererer økonomisk værdifuld indsigt eller fordele.

Bind

Hvor Gigabytes (GB) for bare få år siden var så store, som det blev, er Terabyte (TB) og Petabyte (PB) nu blevet den fælles valuta i Data Managers' liv.

- Flere data krydser internettet hvert sekund end der blev gemt på hele internettet for kun 20 år siden.
- Ifølge IBM blev 90 procent af de data, der findes i dag, genereret i de foregående to år.⁴

De store mængder data er faktisk blevet så store, at vi ikke længere kan lagre og analysere data ved hjælp af traditionel databaseteknologi. Vi bruger nu distribuerede systemer, hvor dele af dataene gemmes forskellige steder og samles af software.

Velocity

- Hvert sekund på hver dag stiger data. Ikke kun skal det analyseres, men transmissionshastigheden og adgangen til dataene skal også forblive øjeblikkelig for at give adgang til data i realtid, ej til websted, verifikation af kreditkort og instant messaging.
- For mange applikationer er hastigheden af oprettelse af data endnu vigtigere end volumen, fordi det ved at analysere information i realtid eller næsten realtid gør det muligt for en virksomhed at være meget mere smidig end sine konkurrenter.

¹ Hentet 18. juli 2019 <http://www.gartner.com/itglossary/big-data/>

² <https://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/#21c1558d13ae>



Bred vifte

Big data har form af meddelelser, opdateringer og billeder, der er sendt til sociale netværk; aflæsninger fra sensorer; GPS-signaler fra mobiltelefoner med mere. Efterhånden som flere og flere forretningsaktiviteter digitaliseres, kombineres nye informationskilder og stadig billigere udstyr for at give os store mængder enormt forskellige typer data. De tilgængelige data er ofte ustrukturerede - ikke organiserede i en database - og u håndterlige, men der er en enorm mængde, der blot venter på at blive frigivet.

Veracity

I nutidens æra skal vi også sætte spørgsmålstegn ved pålideligheden eller pålideligheden af dataene. De fleste data er sandsynligvis ikke "falske", men de kan være unøjagtige eller ufuldstændige. F.eks. Vil en GPS ofte "drive", når du rejser gennem et byområde, hvilket betyder, at det ikke er 100% nøjagtigt, nøjagtigt hvor du er. Dette skyldes, at satellitsignaler går tabt, når de spretter fra høje bygninger eller andre strukturer. Når dette sker, skal lokalitetsdata smeltes sammen med en anden datakilde som vejdata eller data fra et accelerometer for at give nøjagtige data.

Værdi

At have adgang til big data er ikke godt, medmindre vi kan omdanne dem til værdi. Eksempler på værdi stammer fra data inkluderer:

- Data, der hjælper virksomheder med bedre at forstå og betjene kunder. Ej. De data, som Netflix har indsamlet om vores seervaner, hjælper dem med at anbefale flere programmer til os.
- Data, der hjælper med at optimere processer. Ej. Uber er i stand til at forudsige efterspørgsel, dynamisk prisrejser og sende den nærmeste driver til kunderne takket være de data, den indsamler.
- Data, der muliggør innovation. Ej Sportsfirmaer, der placerer sensorer i bolde og GPS-trackere på tøj for at øge værdien for atleter, som derefter kan analysere deres egne datastrømme og forbedre det, de gør.

2.2 FORSTÅENDE DATA

Data kan forstås og klassificeres efter, hvordan (og af hvem) de genereres og / eller efter, hvor organiseret eller struktureret de er.

Der er uendelige typer data tilgængelige for os, men bare det at have flere data er ikke altid bedre. Kvaliteten af dataene, og hvor relevante eller værdifulde de er til dine formål, er lige så vigtig som hvor meget data du har.

3. FRA STORE DATA TIL SMART DATA.

Vi er nødt til at være realistiske, pragmatiske og endda skeptiske over, hvad der kan opnås, og vide, hvilken værdi der kan udledes, og hvordan man maksimerer den opnåede værdi.

- Data refererer til enhver rå og ustruktureret information (tekst, antal, medier).
- Store data refererer til inflationen i datamængden, som undertiden også forringer kvaliteten.
- Smarte data handler om at skabe skabe omkostningseffektive og innovative måder til at behandle store mængder data for bedre indsigt, beslutningstagning og procesautomation.³
- Det kan også henvise til data, der kræver lidt menneskelig indgriben / koordinering for at være nyttige.

(Tidligere blev de fleste analyser udført med batchbehandling. Data blev indsamlet i henhold til skemaet, konverteret til en ønsket tilstand, anbragt i en database og behandlet på time, natten over eller ugentligt. En ulempe ved denne tilgang er, at af Når dataene analyseres, er de allerede gamle. I modsætning hertil overvåger smarte data data ved kilden, vurderer dem, træffer en beslutning og deler output - alt inden for et bestemt tidsvindue, der består af sekunder eller brøkdele af et sekund.)

Hvordan kan du måle, om dine data vil være smarte?

Stil spørgsmål for at sikre, at det hurtigt og effektivt fører til en bedre og mere rentabel service eller produkt.

Hvordan bliver man smart? Stol ikke på HIPPOS!

³ <https://securities.bnpparibas.com/insights/what-is-smart-data-2.html>



Et af de mest kritiske aspekter ved big data er dens indflydelse på, hvordan beslutninger tages, og hvem der skal tage dem. Når data er knappe, dyre at få, eller ikke er tilgængelige i digital form, er det fornuftigt at lade velplacerede mennesker tage beslutninger, hvilket de gør på grundlag af de erfaringer, de har opbygget, og mønstre og forhold, de har observeret og internaliseret.

Imidlertid hævder mange i big data-samfundet, at virksomheder ofte træffer de fleste af deres vigtige beslutninger ved at stole på "HiPPO" - den højst betalte persons mening. ⁴ For at være sikker, er et antal seniorledere virkelig datadrevet og villige til at tilsidesætte deres egen intuition, når dataene ikke er enige i det. Men i hele erhvervslivet i dag stoler folk for meget på erfaring og intuition og ikke nok på data.

Er din virksomhed smart?

Datadrevne virksomheder viser sig at være mere produktive og rentable. NESTA-undersøgelsen afslører, at Datavores og Data Builders er over 10% mere produktive end Dataphobes efter kontrol af andre faktorer. ⁵

4. FORDELE AF DATA TIL FORRETNING

Betydningen af big data drejer sig ikke om, hvor meget data en virksomhed har, men hvordan en virksomhed bruger de indsamlede data. Hver virksomhed bruger data på sin egen måde; jo mere effektivt en virksomhed bruger sine data, desto mere potentiale har den til at vokse.

Vi ser dette manifesteres i innovative datadrevne produkter og tjenester, forbedringer i processer og mere informeret beslutningstagning på tværs af økonomien og samfundet.

Fordelene kan opdeles i to hovedkategorier:

- Behandle. Virksomheden kan tage data fra enhver kilde og analysere dem for at finde svar, der gør det muligt for virksomheden at foretage forbedringer i interne processer.
- Produkt / innovation. Data skaber ikke blot værdi ved at give os mulighed for at gøre de samme ting bedre, men også ved at give os mulighed for at gøre helt nye ting.

⁴ <https://hbr.org/2012/10/big-data-the-management-revolution>

⁵ <https://www.nesta.org.uk/report/skills-of-the-datavores-talent-and-the-data-revolution/>