GENERATION DATA

USING DATA FOR PROFIT

M5: Datastyring, sikkerhed og privatliv

Fuldfør dette modul sammen med modul 5-noter

Dette program er finansieret med støtte fra Europa-Kommissionen

Erasmus+

Præsenter de vigtigste regulatoriske og etiske aspekter af data og deres praktiske anvendelse.

Forstå fordelene og omkostningerne ved software-as-a-service i skyen

en) Vælg passende datateknologi løsninger baseret på omkostning / fordel og langsigtede værdianalyser

Oversigt

1. **Datasikkerhed**
   Typer af trusler
2. **Datasikkerhed**
   Lovgivning om databeskyttelse - GDPR
3. **Dataetik**
4. **Informationsstyring**

Dette program er finansieret med støtte fra Europa-Kommissionen. Forfatteren er udelukkende ansvarlig for denne publikation (meddelelse), og Kommissionen påtager sig intet ansvar for brug af informationen deri.

Med stor magt kommer stort ansvar.

”“

- Voltaire

ressourcenotater 5)

Wannacry

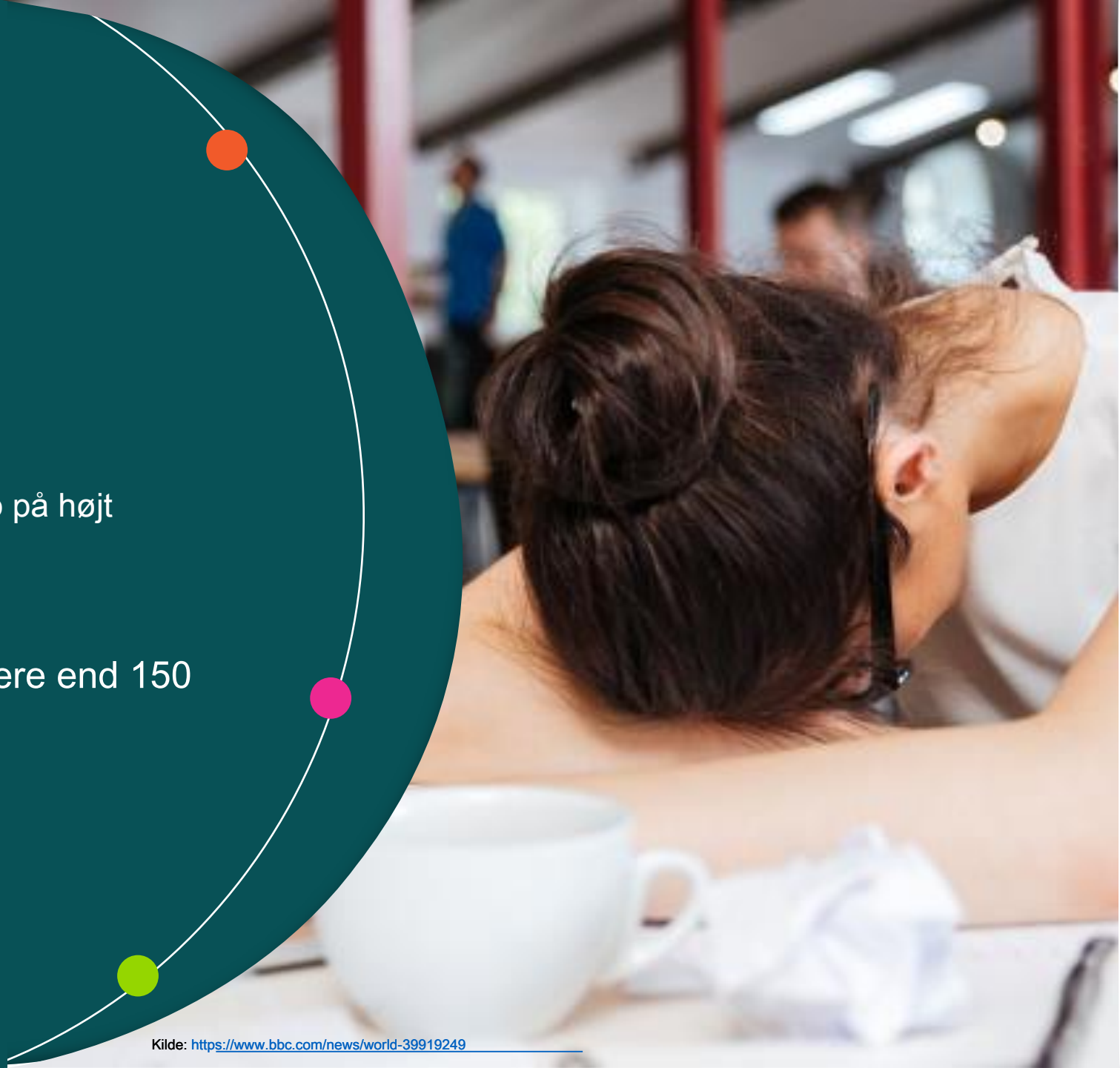
Indstil Scene Case Study (se

I maj 2017 var Wannacry et ransomware-angreb på højt niveau, der påvirkede mere end

200.000 systemer fra virksomheder,

offentlige agenturer og enkeltpersoner i mere end 150 lande.

Hvorfor var angrebet så potent?



Kilde: https://www.bbc.com/news/world-39919249

1. DATA SIKKERHED

Mere end 4.000 ransomware-angreb forekommer dagligt over hele verden.

Mindst 80% af virksomhederne i Europa har oplevet mindst en cybersikkerhedshændelse i 2016.

60% af de små og mellemstore virksomheder, der var ofre for cyberangreb, kom sig ikke og blev lukket inden for 6 måneder.

Kilde: https://www.europeanfiles.eu/industry/education-as-a-tool-against-cybercrime

Typer af trusler

Vira og orm

Kode, der inficerer computere gennem sikkerhedsfejl og gentager sig selv.

Adgang angreb

Udnyttelse kompromitteret digitale certifikater og adgangskoder til adgang til netværk.

Malware

Ondsindet software designet til at beskadige, forstyrre, eller kontrolnetværk, computere eller data. (Ransomware, spyware, adware, bots, trojanere.)

DDos

Distribueret afslag på service: afbrydelse af en tjeneste eller et netværk ved at overvælde mål eller dens omgivende infrastruktur med en oversvømmelse af internet Trafik.

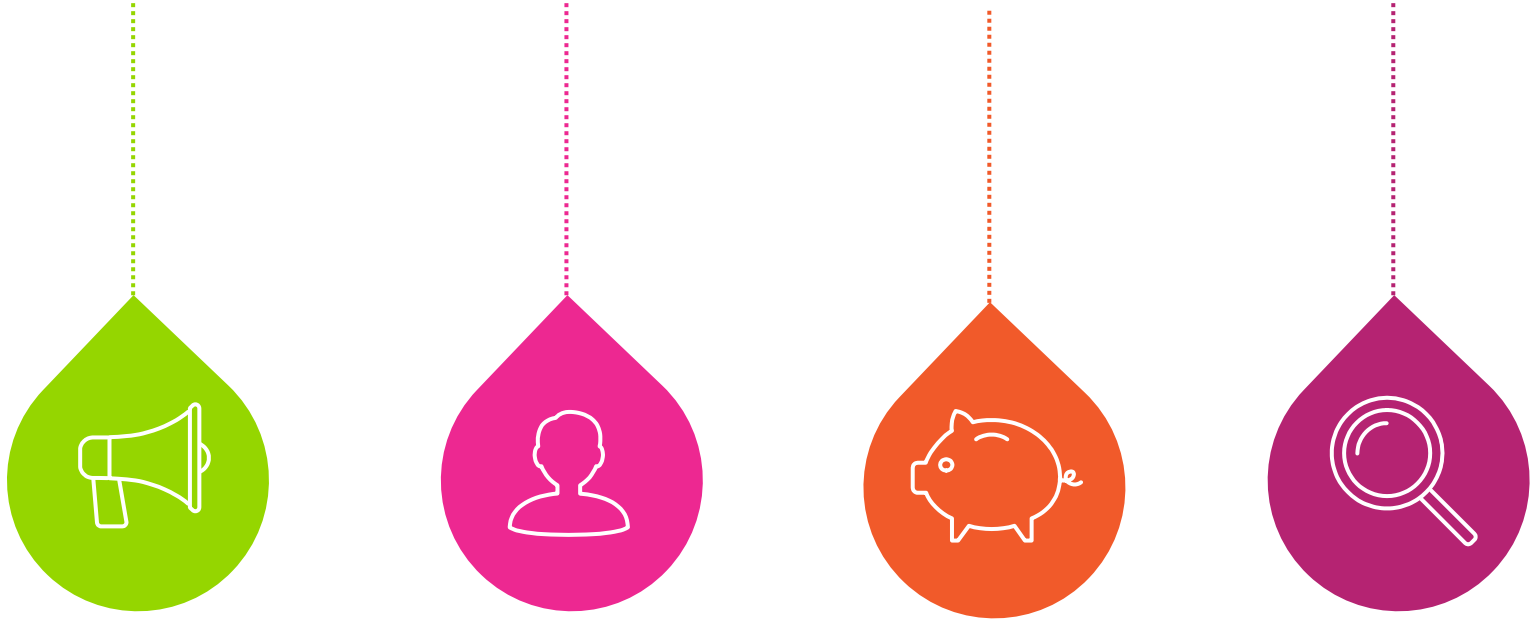
Hacking

Hackere infiltrerer lettere netværk og computere, efterhånden som flere data kobles sammen.

Kilde: https://securitytrails.com/blog/top-10-common-network-security-threats-explained

Hvorfor betyder sikkerhed noget?

En hændelse, hvor følsom eller på anden måde er beskyttet information fås uden tilladelse.



omdømmerelaterede SKADE

NEDRE KUNDETILfredshed

FINANSIEL KOSTE

BØDE OG LOVEDRAG

2. DATA PRIVACY Hvorfor betyder det noget?

Begrænsninger for strøm

Privatlivets fred er en grænse for regeringens magt såvel som den private sektors virksomheders magt. I de forkerte hænder kan personlige data bruges til at forårsage os

stor skade.

Væsentlige friheder

Privatliv er nøglen til frihed til at tænke, udtrykke og beskytte vores evne til

omgås andre mennesker. Privatliv er en kritisk komponent i et demokratisk samfund.

Ret til sekund chance

Privatliv giver anledning til evnen til at lære af vores fejltagelser, vokse og forbedre os uden at blive bundet til fejlene

af vores fortid.

Kilde: https://teachprivacy.com/10-reasons-privacy-matters/

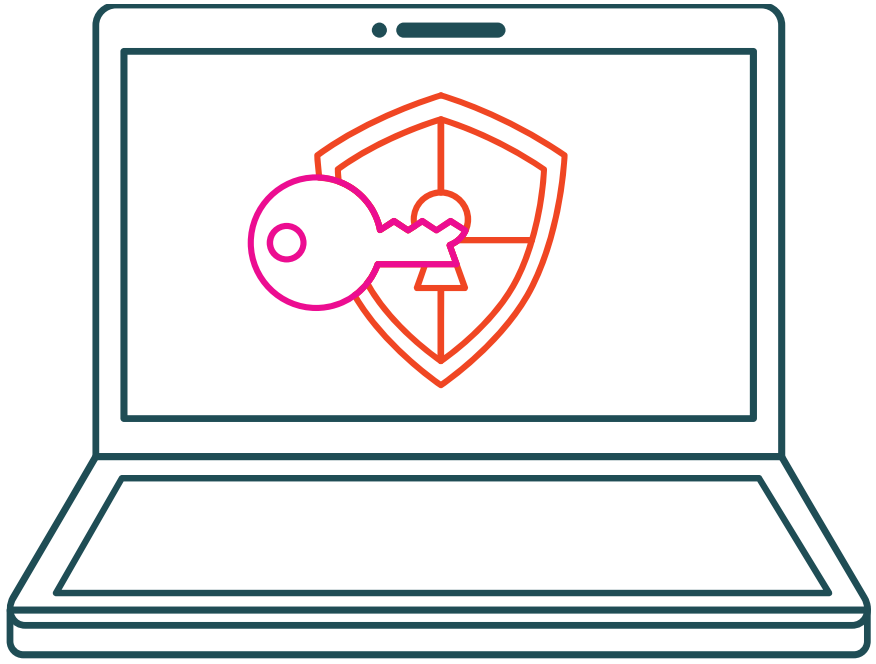
DATA BESKYTTELSE

ER ANSVARLIG TIL

DATA PRIVACY

Handler om at sikre data mod uautoriseret adgang.

Det er et teknisk problem, der kræver sikkerhedsforanstaltninger for at beskytte data mod kompromis fra eksterne angribere og ondsindede insidere.



Forholder sig til rettigheder og friheder, der er nedfældet under lov.

Det er et juridisk spørgsmål, der vedrører, hvordan data indsamles, deles og bruges.

Teknologi alene kan ikke sikre privatlivets fred for personlig data

Datasikkerhed

EU's almindelige databeskyttelsesforordning

I henhold til den almindelige databeskyttelsesforordning (EU) 2016/679

Personlige oplysninger skal være:

- behandles lovligt, retfærdigt og gennemsigtigt
- indsamlet til specificerede, eksplicitte og legitime formål og ikke viderebehandlet på en måde, der er uforenelig med disse formål
- passende, relevant og begrænset til, hvad der er nødvendigt i forhold til de formål, de behandles til
- nøjagtig og om nødvendigt opdateret
- opbevares i en form, der ikke tillader identifikation af registrerede i længere tid end nødvendigt

- behandles på en måde, der sikrer passende sikkerhed for personoplysninger, herunder beskyttelse mod uautoriseret eller ulovlig behandling og mod utilsigtet tab, ødelæggelse eller skade, ved hjælp af passende tekniske eller organisatoriske foranstaltninger.

BNPR

Kunst. 4 (1). Personlige data er enhver information, der er relateret til en identificeret eller identificerbar fysisk person. Forskellige informationer, der indsamles sammen, kan føre til identifikation af en bestemt person, udgør også personlige data.

Økonomisk information

- Kreditkort
- Kontodata
- Kundenummer

Kontakt information

- Hjemme adresse
- Email adresse
- Telefon nummer
- IP-adresse
- Placeringsdata (GPS)

Andet

- Sundhedsregistre
- Kreditbedømmelse
- Eksamensbesvarelser
- Udseende

I tilfælde af overtrædelse:

Underret tilsynsmyndigheden senest 72 timer efter, at den er blevet opmærksom på overtrædelsen.

Informer alle de berørte personer, hvis dataovertrædelsen udgør en høj risiko.

British Airways

- En halv million passagerrekorder blev adgang i et cyberangreb. Virksomheden løste overtrædelsen og underrettede politiet.
- Informationskommissærens kontor (ICO) undersøgte og beskyldte dårlig sikkerhedsordning.

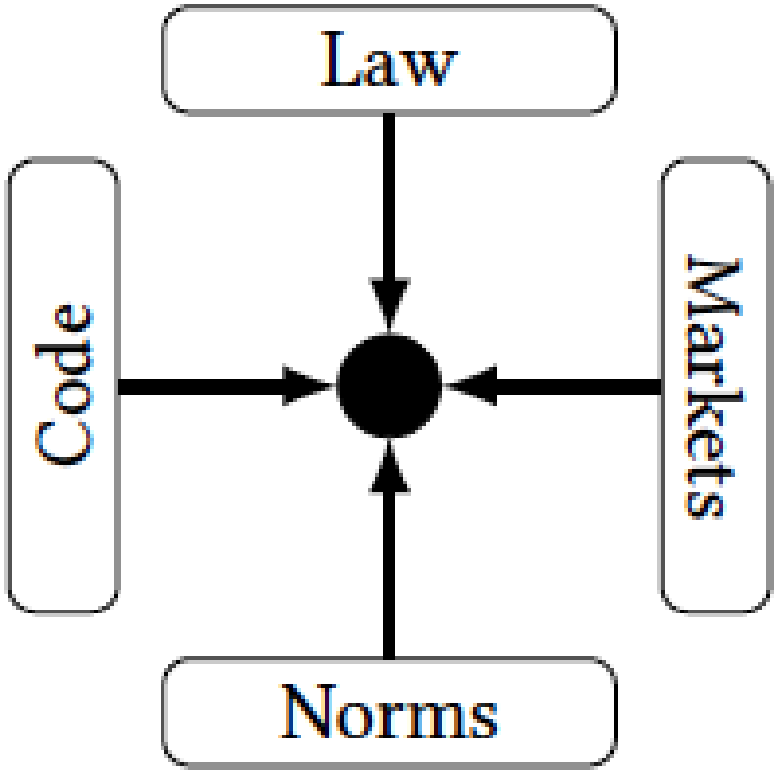
- Flyselskabet skal betale 183,39 millioner pund (230 millioner dollars) til ICO for ikke at beskytte sine kunders data.

Kilde: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/07/ico-announcesintention-to-fine-british-airways/

Overvågning er forretningsmodellen for internettet.

Bruce Schneider

Den patetiske dotteori



Kilde: Lessig. *Kode og andre love inden for cyberspace,* 1999.

3. DATAETIK

Samtykke

Samtykke er ikke længere en engangsbeslutning, men en igangværende problem, der skal overvåges nøje. Mekanismerne for tilbagetrækning af samtykke bør

være så nem som for giver samtykke.

Datadeling

Stærk etisk praksis udelukker deling af data uden samtykke, men effektiv brug af data kræver, at de skal deles. At tackle problemerne og mindske risikoen forbundet med deling af data er

vital.

Algoritmisk bias

Data kan ofte indeholde skjulte forspændinger, der er produceret af systematiske og gentagne fejl i et computersystem. Disse kan skabe upålidelige eller urimelige udfald.

Kilde: https://www.bpe.co.uk/services/need/data-protection-the-gdpr/brilliantly-simple-guide-to-the-gdpr/explaining-what-is-meant-by-consent/

Data Ethics Canvas

What are your data sources?

Name and describe key data sources used in your project, whether you're collecting them yourself or getting access from third parties.

Who has rights over your data sources?

Where did you get the data from? e.g. is it data produced by an organisation or data collected directly from individuals?

Do you have permission or another basis on which you're allowed to use this data? What ongoing rights will the data source have?

What's your core purpose for using this data?

What is your primary use case, your business model?

Are you collecting more data than is needed for your purpose?

Who could be negatively affected?

Could the manner in which this data is collected, shared, used cause harm?

- be used to target, profile, prejudice people
- unfairly restrict access (eg exclusive arrangements)

Could people *perceive* it to be harmful?

Are you communicating potential risks/issues, if any?

How are limitations and risks being communicated to people affected by your project, and organisations using data?

What channels are you using?

Are there any limitations in your data sources?

Which might influence the outcomes of your project, like:

- bias in data collection, inclusion, algorithm
- gaps, omissions
- other sensitivities

What policies/laws shape your use of this data?

Data protection legislation, IP and database rights legislation, sector specific data sharing policies/regulation (e.g. health, employment, taxation)

Sector specific ethics legislation?

Do people understand your purpose?

If this is a project/use that could impact on people or more broadly shape/impact society, do people understand your purpose?

Has this been clearly communicated to them?

How are you minimising negative impact?

What steps can you take to minimise harm? Are there measures you could take to reduce limitations in your data sources? Could you monitor potential negative impact to support mitigating activities? What benefits will these actions add to your project?

When is your next review?

When will this Data Ethics Canvas be reviewed?

How will ongoing issues be monitored?

Are you going to be sharing this data with other organisations?

If so, who?

Who will be positively affected by this project?

What individuals, demographics, organisations?

How will they be positively affected?

Do they know and understand how they are positively affected?

How can people engage with you?

Can people affected appeal or request changes to the service? To what extent?

Are the appeal mechanisms reasonable?

What are your actions?

What steps are you going to take prior to moving forward with this project?

Download og brug dette værktøj:

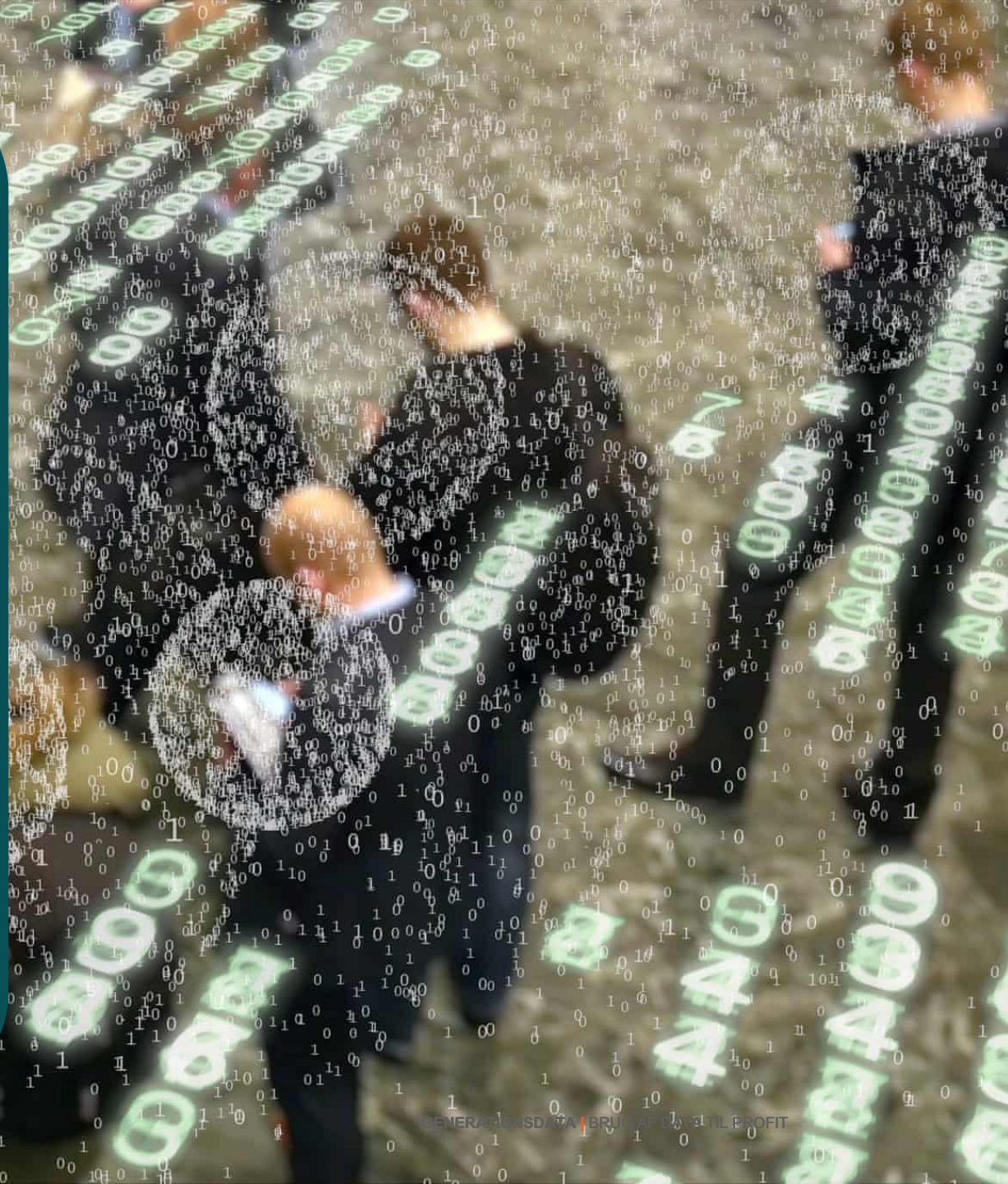
https://theodi.org/article/data-ethics-canvas/

Bias i AI

Tre af de seneste AI'er for kønsgenkendelse, fra IBM Microsoft og det kinesiske firma Megvii, kunne identificere en persons køn fra et fotografi 99 procent af tiden, men kun for hvide mænd. For mørkhudede kvinder faldt nøjagtigheden til kun 35 procent.

Det øger risikoen for falsk identifikation af kvinder og mindretal.

Kilde: Ny videnskabsmand



Bias i implementering

Et firma byggede en smartphone-app, der overvåger efter huller i vejen ved passivt at indsamle accelerometerdata.

De første byer, der implementerede denne teknologi til at prioritere vejvedligeholdelse, så velhavende samfund få mest opmærksomhed.

De havde ikke de værste veje, de var folk med flest smartphones .

Kilde: New Scientist

4. INFORMATIONSSTYRING

Et sæt regler og vejledninger til, hvordan data eller information håndteres i en organisation. En informationsstyringsstrategi skal specificere beslutningsrettighederne og en ansvarlighedsramme, der tilskynder til ønskelig adfærd ved værdiansættelse, oprettelse, opbevaring, brug, arkivering og sletning af information. ”

Målene er at:

1. Maksimer værdien af information til organisationen ved at sikre, at informationen er pålidelig, sikker og tilgængelig til beslutningstagning
2. Beskyt information, så dens værdi for organisationen ikke formindskes ved hjælp af teknologi eller menneskelig fejl, tab af rettidig adgang, upassende brug eller fejlagtig fejl.

Kilde: Logan 2010, citeret i Rising, Kristensen, Tjerrild-Hansen, 2014

Informationsstyring

Hvordan ser en strategi for informationsstyring ud?

Antecedents (Aktiverere)

Organisation / IT-strategi
Organisation / IT-struktur
IT-standardisering
IT-kultur (fremmer it-brug)
Brancheregler Information
Væksthastighed

Antecedents (hæmmere)

Produktkompleksitet / Mix
Legacy IT-systemer IT-kultur
(Packrat mentalitet)

Sammensætning af informationsstyring

Strukturelle fremgangsmåder

- Procedurer for indstilling af politik
- Tilsynsmekanismer
- Ansvar for dataejerskab

Procedurepraksis (livscyklus)

- Håndhæv tilbageholdelse / arkivering
- Anvend sikkerhedskopieringspraksis
- Etablere og overvåge adgang
- Klassificer oplysninger efter værdi
- Serviceniveau for databeskyttelse
- Overvåg omkostninger via tilbageførsler
- Migration mellem oplagringslag

Relationspraksis

- Brugernes uddannelse
- Kommunikation / idéudveksling

Konsekvenser

- Filmydelse - Mellemniveau
- Risikobegrænsning

Figur 7: Oversigt over sammensætning og kontekst for konceptet Information Governance (Tallon etal. 2013b)

Informationsstyri

Opbygning af en informationsstyringsstrategi

SENIOR LEDERSKAB

- Udvikle en klar, skriftlig strategi, der deles med alt personale.
- Dette bør være baseret på en risikovurdering og dataetisk politik.
- Tilbyde regelmæssig vejledning om databeskyttelse og etik er ajour og implementere politikker korrekt.

DATA LEDER

- Test dine sikkerhedsforanstaltninger ofte i henhold til en forud fastlagt kalender
- Vær forberedt på muligheden for ransomware-angreb med en skriftlig politik for at vejlede responsen.

DET PERSON

- Hold dine systemer og software opdateret og bortskaffes udstyret sikkert.
- Fastlæg en politik for sikkerhedskopier, herunder hvor ofte og hvor de skal opbevares.
- Sørg for, at personalet bruger stærke adgangskoder.
- Brug en firewall som en beskyttelsesforanstaltning.

Informationsstyring

Risikovurdering

FORSTÅ DIN RISIKO

Gennemgå de data og oplysninger, der er mest værdifulde; spørg, hvordan det gemmes, hvem der har adgang, og hvordan det er beskyttet af både teknologi og processer.

MINIMISER DIN RISIKO

Gør noget for at afbøde de svage punkter i dit nuværende system.

Udforsk cyberforsikring, hvis det er relevant.

Opret et svarplan

Udarbejd en handlingsplan i tilfælde af dataovertrædelse eller hændelse.

Det skal omfatte:

Juridisk reaktion

Undersøgelse og berigtigelse af situationen

Informere kunder Håndtering

af mediespørgsmål

Informationsstyring

Principper for oprettelse af en kodeks for datahetik

1. Den højeste prioritet er at respektere personerne bag dataene

Hvor indsigt fra data kan påvirke den menneskelige tilstand, bør den potentielle skade på enkeltpersoner og lokalsamfund være de største overvejelser. Store data kan producere overbevisende indsigt i populationer, men den samme indsigt kan bruges til at uretfærdigt begrænse en persons muligheder.

2. Redegør for downstream-brug af datasæt

Datafagfolk skal stræbe efter at bruge data på måder, der er i overensstemmelse med den afslørende parts intentioner og forståelse. Mange regler regulerer datasæt på grundlag af status for dataene: “offentlig”, “privat” eller “proprietær”, for eksempel. Men hvad der gøres med datasæt, er i sidste ende mere konsekvens for emner / brugere end den type data eller den kontekst, hvori de er samlet. Korrelativ brug af repurposed data i forskning og industri repræsenterer det største løfte og den største risiko for dataanalyse.

3. Konsekvenserne af at bruge data og analyseværktøjer i dag er formet af, hvordan de er blevet brugt i fortiden

Der er ikke noget som rå dat. Alle datasæt og ledsagende analytiske værktøjer har en historie med menneskelig beslutningstagning. Historien skal så vidt muligt være auditerbar. Dette bør omfatte mekanismer til sporing af indsamlingssammenhæng, metoder til samtykke, ansvarskæder og vurderinger af datakvalitet og nøjagtighed.

Figur 2: Universelle principper for datetetik - Retningslinjer for oprettelse af en kode for dataetik

Informationsstyring

Principper for oprettelse af en kodeks for datahetik

4. Forsøg at matche beskyttelse af personlige oplysninger og sikkerhed med privatlivets fred og sikkerhedsforventninger.

Den registrerede har en række forventninger til privatlivets fred og sikkerhed for deres data. Disse forventninger er ofte kontekstafhængige. Designere og data-fagfolk skal tage behørigt hensyn til disse forventninger og tilpasse sikkerhedsforanstaltninger og forventninger så meget som muligt.

5. Følg altid loven, men forstå, at loven ofte er en mindste bar.

Digital transformationer har været en standard evolutionær vej for virksomheder og regeringer. Men fordi love stort set ikke har holdt trit med den tid, hvor digital innovation og ændring er gået, er eksisterende regler ofte forkert kalibreret til aktuelle risici. I denne sammenhæng betyder overholdelse selvtilfredshed. For at udmærke sig i dataetik skal ledere definere deres egne overholdelsesrammer for at overgå lovgivningsmæssige krav

6. Vær forsigtig med at indsamle data bare for at have flere data

Kraften og faren ved dataanalyse er, at data indsamlet i dag vil være nyttige til uforudsigelige formål i fremtiden. Overvej grundigt muligheden for, at mindre data kan resultere i både bedre analyse og mindre risiko.

Figur 2: Universelle principper for datetetik - Retningslinjer for oprettelse af en kode for dataetik

Informationsstyring

Principper for oprettelse af en kodeks for datahetik

7. Data kan være et værktøj til både inklusion og ekskludering

Mens alle skal have adgang til de sociale og økonomiske fordele ved data, påvirkes ikke alle lige så meget af processerne med dataindsamling, korrelationer og forudsigelse. Datafagfolk skal bestræbe sig på at afbøde de forskellige konsekvenser af deres produkter og lytte til bekymringerne fra de berørte samfund.

8. Forklar så vidt muligt metoder til analyse og markedsføring til data-afslørere

Maksimering af gennemsigtighed på tidspunktet for dataindsamling kan minimere de mere markante risici, der opstår, når data rejser gennem dataforsyningskæden.

9. Datavidenskabsmænd og praktikere skal nøjagtigt repræsentere deres kvalifikationer (og grænser for deres ekspertise), overholde faglige standarder og stræbe efter peer-ansvarlighed

Denne disciplin på lang sigt afhænger af tillid fra offentligheden og klienter. Datafagfolk skal udvikle praksis for at holde sig selv og deres kammerater ansvarlige for fælles standarder.

Figur 2: Universelle principper for datetetik - Retningslinjer for oprettelse af en kode for dataetik

Informationsstyring

Principper for oprettelse af en kodeks for datahetik

10. Stræb efter designpraksis, der indeholder gennemsigtighed, konfigurerbarhed, ansvarlighed og revisionsevne

Ikke alle etiske dilemmaer har designløsninger. Men at være nøje opmærksom på designpraksis kan nedbryde mange af de praktiske barrierer, der står i vejen for fælles, robuste etiske standarder. Dataetik er en ingeniørudfordring, der er værdig for de bedste sind i området.

11. Produkter og forskningspraksis skal underkastes intern (og potentielt ekstern) etisk gennemgang

Organisationer bør præstatisere etablering af ensartet, effektiv og handlingsmæssig etisk gennemgang for de nye produkter, tjenester og forskningsprogrammer. Intern peep-review-praksis hjælper med at mindske risikoen, og et eksternt evalueringsudvalg kan bidrage væsentligt til offentlighedens tillid.

12. Regeringsførelsespraksis skal være robuste, kende til alle holdmedlemmer og regelmæssigt gennemgås.

Dataetik udgør organisatoriske udfordringer, der ikke kan løses ved overholdelsesordninger alene. Da de lovgivningsmæssige sociale og tekniske terræner er i flux, har organisationer, der beskæftiger sig med dataanalyse, brug for samarbejde, rutine og gennemsigtig praksis for etisk regeringsførelse.

Figur 2: Universelle principper for datetetik - Retningslinjer for oprettelse af en kode for dataetik

AKTIVITET 1

Diskussionsspørgsmål

Overvågningskapitalisme hævder ensidig menneskelig oplevelse som gratis råstof til oversættelse til adfærdsdata. Selvom nogle af disse data anvendes til forbedring af tjenester, erklæres resten som et proprietært adfærdsoverskud, fodret med maskinens intelligens og fremstillet i forudsigelsesprodukter. Disse forudsigelsesprodukter handles på adfærdsmarkedet futures markeder. Overvågningskapitalister er vokset enormt velhavende fra disse handelsoperationer.

Shoshana Zuboff

Arbejdet med Shosana Zuboff, en fremtrædende amerikansk akademiker, er blevet en primær ramme for forståelse af big data og det større felt med kommerciel overvågning. Hun argumenterer for det
hverken privatlivets fred eller antitrustlove giver tilstrækkelig beskyttelse mod den hidtil usete praksis med overvågningskapitalisme.

Tror du, at virksomhedspraksis truer individuel autonomi og demokrati? Hvor mener du, at ansvaret ligger for at beskytte individuelle rettigheder: hos virksomheder eller med regeringer?

Kilde: https://www.theguardian.com/books/2019/oct/04/shoshana-zuboff-surveillance-capitalism-assault-human-automomy-digital-privacy

AKTIVITET 2

Foretag en konsekvensanalyse af databeskyttelse (DPIA)

DPIA er en proces, der hjælper dig med at identificere og minimere risikoen for databeskyttelse i et projekt. Du skal lave en DPIA til behandling, der sandsynligvis vil resultere i en høj risiko for enkeltpersoner. Dette inkluderer nogle specificerede typer behandling. Du kan bruge vores screeningchecklister til at hjælpe dig med at beslutte, hvornår du skal foretage en DPIA. Det er også god praksis at lave en DPIA til ethvert andet større projekt, der kræver behandling af personoplysninger.

Opret en DPIA-skabelon til et projekt efter eget valg, skal du:

- beskrive arten, omfanget, konteksten og formålet med behandlingen;
- vurdere nødvendighed, proportionalitets- og overholdelsesforanstaltninger
- identificere og vurdere risici for enkeltpersoner og
- identificere eventuelle yderligere foranstaltninger til at afbøde disse risici.

Kilde: https://www.dataprotection.ie/en/organisations/know-your-obligations/data-protection-impact-assessments