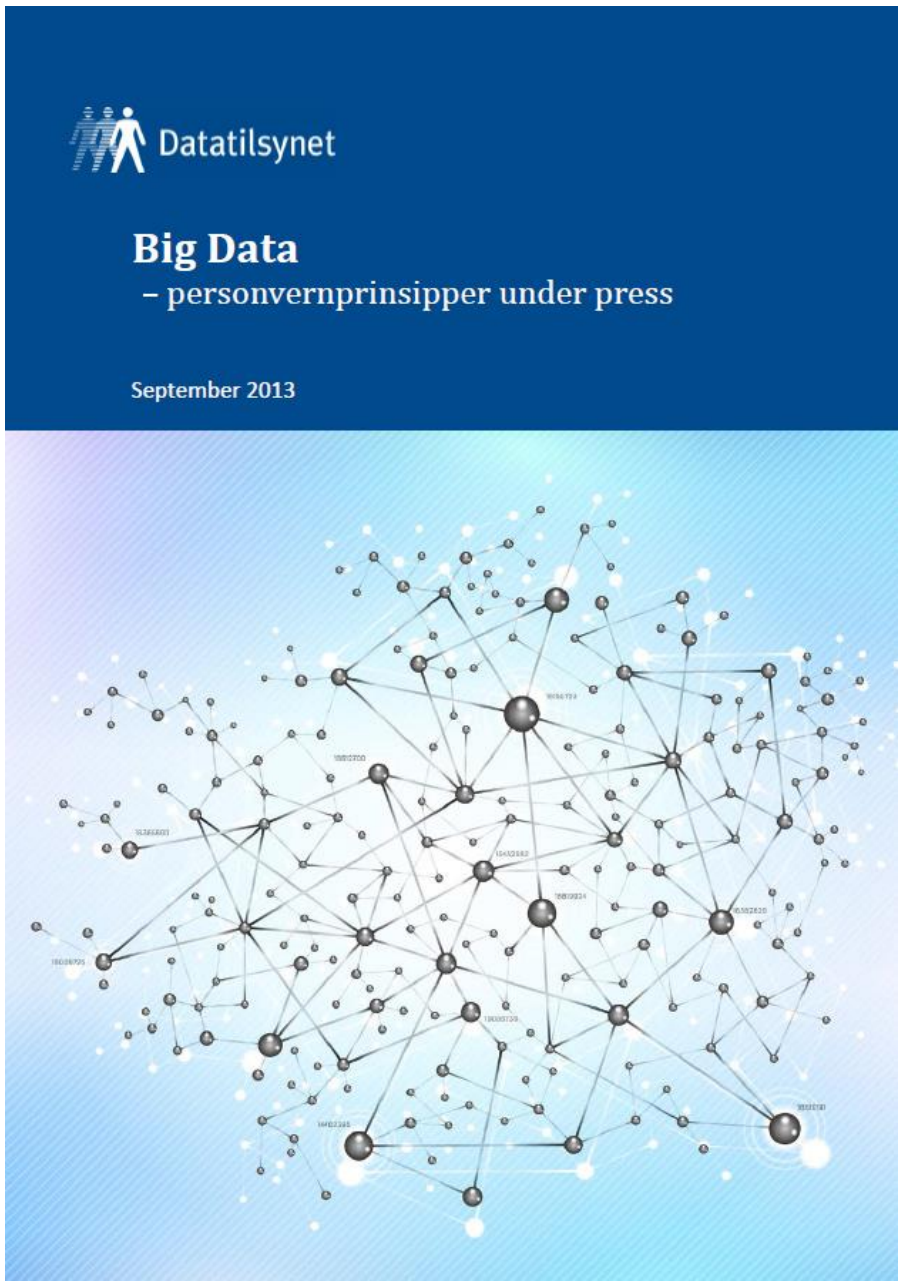
# Big Data –
## personvernprinsipper under press

Datatilsynet

SINTEF, Oslo 24.10.2013

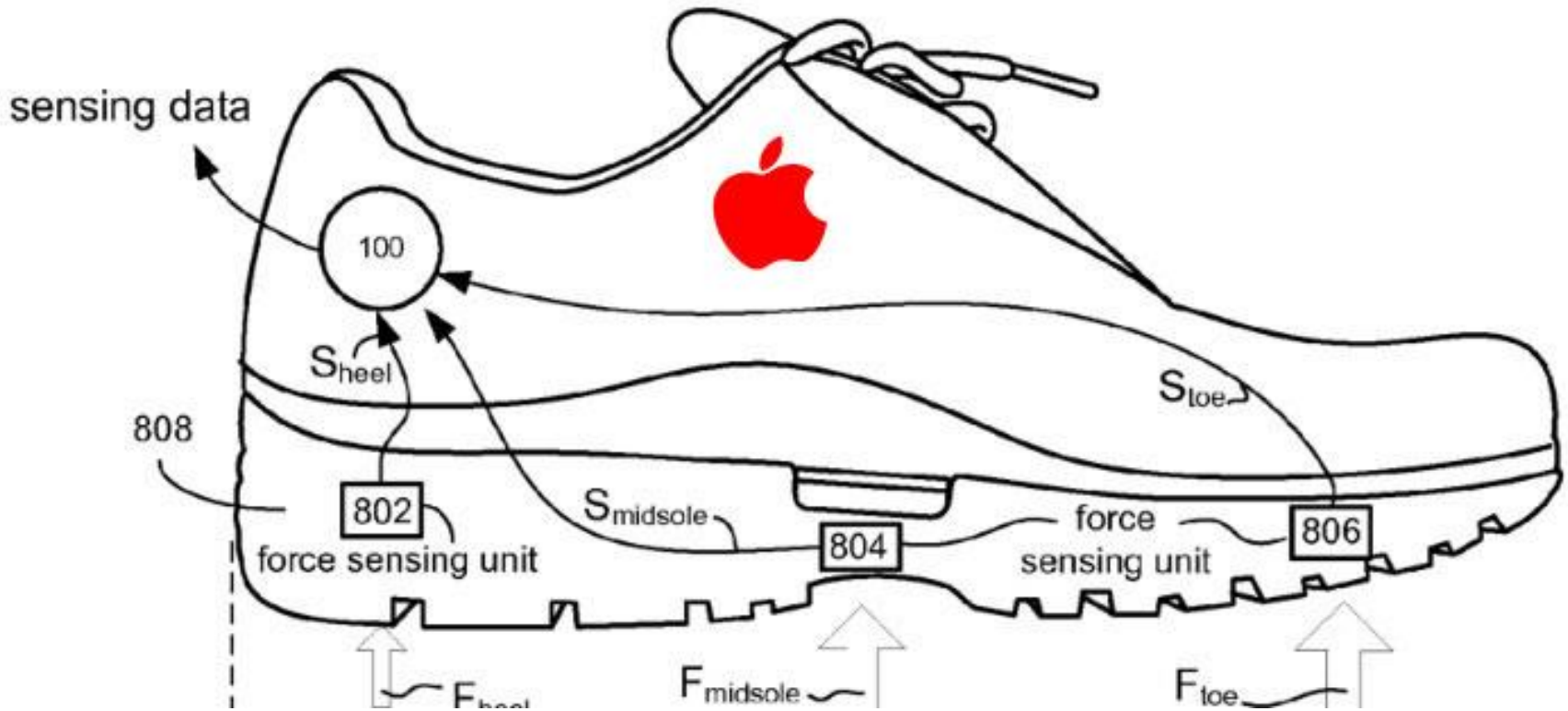
Catharina Nes, seniorrådgiver


Datatilsynet
Big Data
– personvernprinsipper under press
September 2013

# Big Data from Cheap Phones

**Collecting and analyzing information from simple cell phones can provide surprising insights into how people move about and behave—and even help us understand the spread of diseases.**

# Big Data er deg



HOME NEWS WORLD SPORT FINANCE COMMENT BLOGS CULTURE TRAVEL LI

Technology News | Technology Companies | Technology Reviews | Video Games | Technology V

HOME » TECHNOLOGY » FACEBOOK

## 'Like' curly fries on Facebook? Then you're clever

'Like' curly fries? Then there's a good chance you've got a high IQ, according to a Cambridge University project to discover what we unwittingly reveal about ourselves on Facebook.

DinSide

15 ÅR PÅ NETT

Har du et

DEN | DATA | MOBIL | MOTOR | ØKONOMI | BOLIG | REISE | BÅT | TESTER | VIDEO

Bank & forsikring | Skatt | Jobb | Kjøp & salg | Lov & rett | Verktøy

**HVA PUTTER DU I POSEN?** Er du Coop-medlem må du regne med at butikken har full kontroll på hva akkurat du handler. Informasjon om ditt handlemønster bruker kjeden for å lage tilbud som de tror at akkurat du vil bite på. (Foto: Kim Jansson)

## Coop sender handlekurven din til analyse

**Går gjennom alt medlemmene handler for å gi «relevante tilbud».**

# Personvernprinsipper under press

- Big Data utfordrer særlig to personvernprinsipper:
  - **Formålsbestemthet:** Personopplysninger må benyttes kun for bestemte formål
  - **Dataminimalisering:** Ikke lagre mer data enn nødvendig. Data som ikke lenger er nødvendige for det angitte formål må slettes eller anonymiseres

# Bruk av data til nye formål

- Big Data = *gjenbruk* av data
- Utfordrer prinsippet om **formålsbestemthet**
- Vil vi få en utvikling mot svært brede samtykker?

# Datamaksimalisering

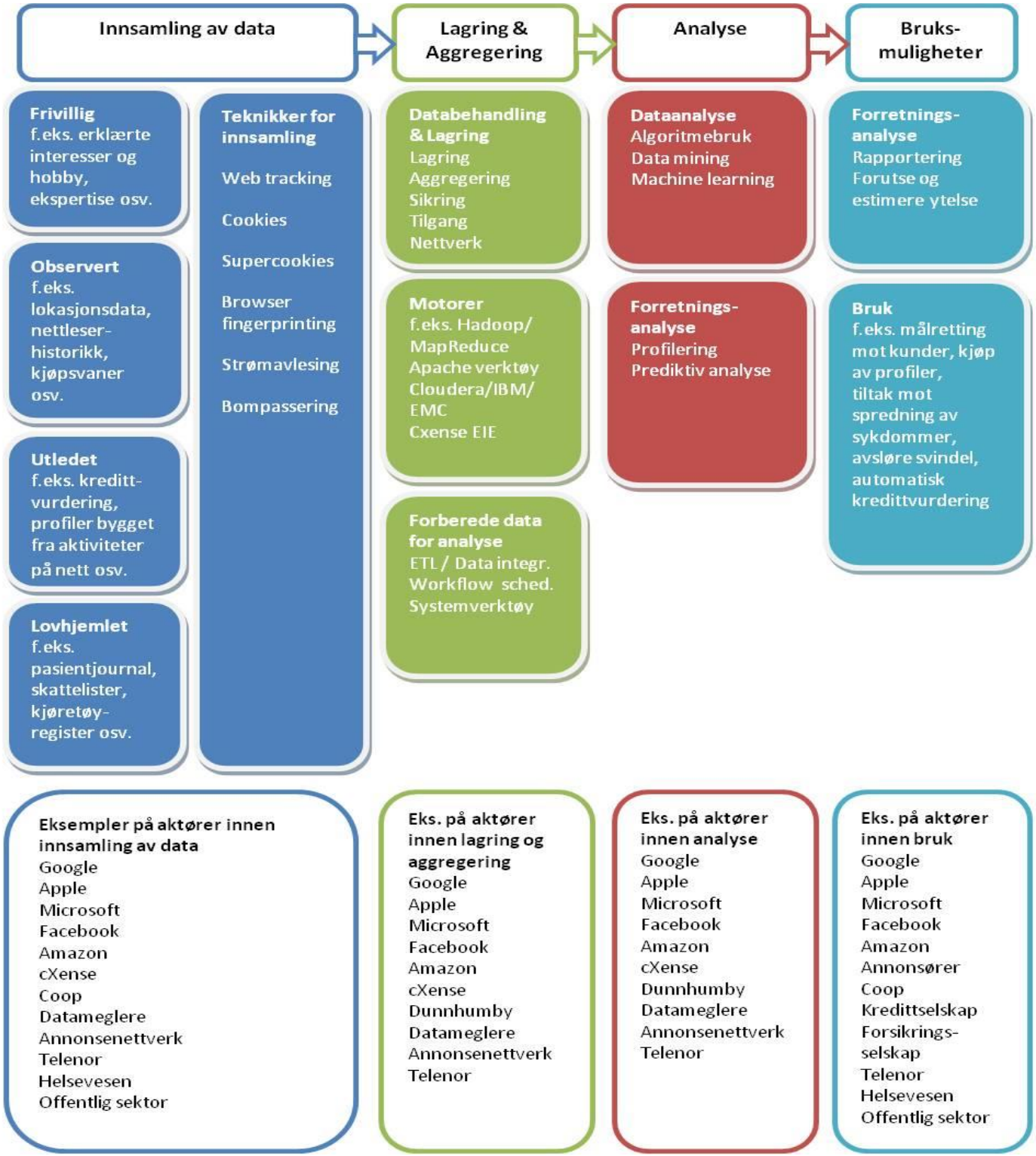
- Big Data = nytt syn på data
- Verdien ligger i dataenes *fremtidige* bruksmuligheter
- Dette utfordrer prinsippet om **dataminimalisering**
- Men hvem vil ønske å slette penger?

# Utfordring 1: Tap av kontroll

- Hvilke opplysninger samles inn om meg, av hvem og til hvilket bruk?
- Vi blir gjenstand for beslutninger vi ikke forstår og som vi ikke har kontroll over
- Marked preget av lite gjennomsiktighet: retten til innsyn blir vanskelig å praktisere når vi ikke vet hvem som samler inn hva.

*" the time has come for businesses to move their data collection and use practices out of the shadows and into the sunlight"*

FTC Charwoman Edith Ramirez, aug. 2013


Innsamling av data
Lagring & Aggregering
Analyse
Bruks-muligheter
Frivillig
f.eks. erklærte interesser og hobby, ekspertise osv.
Observert
f.eks. lokasjonsdata, nettleser-historikk, kjøpsvaner osv.
Utledet
f.eks. kreditt-vurdering, profiler bygget fra aktiviteter på nett osv.
Lovhjemlet
f.eks. pasientjournal, skattelister, kjøretøy-register osv.
Teknikker for innsamling
Web tracking
Cookies
Supercookies
Browser fingerprinting
Strømavlesing
Bompassering
Databehandling & Lagring
Lagring
Aggregering
Sikring
Tilgang
Nettverk
Motorer
f.eks. Hadoop/ MapReduce
Apache verktøy
Cloudera/IBM/ EMC
Cxense EIE
Forberede data for analyse
ETL / Data integr.
Workflow sched.
Systemverktøy
Dataanalyse
Algoritmebruk
Data mining
Machine learning
Forretnings-analyse
Profilering
Prediktiv analyse
Forretnings-analyse
Rapportering
Forutse og estimere ytelse
Bruk
f.eks. målretting mot kunder, kjøp av profiler, tiltak mot spredning av sykdommer, avsløre svindel, automatisk kredittvurdering
Eksempler på aktører innen innsamling av data
Google
Apple
Microsoft
Facebook
Amazon
cXense
Coop
Datameglere
Annonsenettverk
Telenor
Helsevesen
Offentlig sektor
Eks. på aktører innen lagring og aggregering
Google
Apple
Microsoft
Facebook
Amazon
cXense
Dunnhumby
Datameglere
Annonsenettverk
Telenor
Eks. på aktører innen analyse
Google
Apple
Microsoft
Facebook
Amazon
cXense
Dunnhumby
Datameglere
Annonsenettverk
Telenor
Eks. på aktører innen bruk
Google
Apple
Microsoft
Facebook
Amazon
Annonsører
Coop
Kredittselskap
Forsikrings-selskap
Telenor
Helsevesen
Offentlig sektor

# Utfordring 2: Sammenstilling kan avsløre sensitive data

- Sammenstilling av ikke-sensitive data kan gi sensitivt resultat
- Kan frembringe informasjon om:
  - seksuell legning
  - sykdommer
  - politisk overbevisning
- Sensitive data krav på ekstra vern. Krav om samtykke hvis analyse av de innsamlede opplysningene brukes til å predikere f.eks helserisiko

The Creepiness Factor: How Obama and Romney Are Getting to Know You

APR 10 2012, 12:45 PM ET 29 Share 30 Tweet 233 +1 28

*The presidential campaigns have the technology to know more about voters than any other bids in history.*



TECH | 2/16/2012 @ 11:02AM | 1,885,680 views

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

# Utfordring 3: Profileringssamfunnet

- ”Dataenes diktatur”
- Kan befeste fordommer og sosial ekskludering
- Beslutninger treffes på feil og lite etterprøvbare data
- Mer data ikke nødvendigvis lik mer kunnskap, kan også bety mer forvirring og flere falske positiver

# Du blir summen av dine spor på nett

# Utfordring 4: Farvel anonymitet?

- Gjennom sammenstilling av data kan det oppstå risiko for reidentifisering
- Er anonyme data mulig i Big Data-alderen?
- Anonymisering som metode for å redusere personvernulemper blir mindre virkningsfull.

THREAT LEVEL | privacy

Netflix Spilled Your *Brokeback Mountain* Secret, Lawsuit Claims

BY RYAN SINGEL 12.17.09 4:29 PM

Follow

Like 177 Tweet 18 +1 0 5

# Anbefalinger

## for personvernvennlig bruk av Big Data

# skaff samtykke

## eller

# anonymiser

# Husk risikovurdering

Gode rutiner for anonymisering viktig

Kartlegg sannsynligheten for at dataene kan reidentifiseres

Bygg løsninger etter prinsippene for innebygd personvern

# Åpenhet skaper tillitt

Hvilke data samles inn, fra hvor og til hvilket bruk?

Hvordan funker algoritmen?

## OUR PRIVACY & DATA PROTECTION PRINCIPLES

We do not access data containing Personally Identifiable Information on any individual, without a freely given, unambiguous and informed consent

We never access the content of private communications

We never attempt to re-identify de-identified data

We ensure appropriate technical and administrative safeguards are in place to prevent unauthorized disclosure or breach of data

We access, analyze, store, transmit or otherwise use only data that has been lawfully and properly obtained from partners

We design, carry out, report and document our activities accurately, transparently and objectively

We employ even stricter standards of care while conducting research among vulnerable populations and persons at risk, children and young people

We perform due diligence when selecting data or service provider partners and ensure their activities comply with the United Nations' global mandate

We ensure that our research partners are acting in compliance with relevant law and privacy and data protection standards

OUR PRIVACY AND DATA PROTECTION PRINCIPLES

*”With Big Data comes big responsibilities”*

danah boyd, Microsoft Research

Takk for meg!

Catharina Nes
cane@datatilsynet.no