

# SERI Forelesning

2014-10-09

Erlend Øverby

@erlendoverby

# Litt bakgrunn

- Senior Rådgiver i Karde AS
- Sektorstyre IKT – Standard Norge
- Leder ISO/IEC JTC 1/SC 36 – Information Technologies, for Learning, Education and Training
- Arbeidet med digital informasjon siden Gopher
- Arbeidet med XML siden det het SGML
- Arbeidet med UU siden 1997

# Disclaimer

Jeg er ikke en jurist

# Mål for forelesning

- Gi dere en forståelse for hva semantiske teknologier er og hvordan det vil påvirke arbeidet med lover
- Overbevise dere om at jurister har en viktigere rolle i utviklingen av offentlig sektor og IT systemer

# Bakgrunnsmateriale

- Tim Berners-Lee: The next web (TED Talk)
  - [http://www.ted.com/talks/  
tim berners lee on the next web](http://www.ted.com/talks/tim_berners_lee_on_the_next_web)
- Semicolon rapport:
  - «Etablering av modell for elektronisk informasjonsutveksling»
    - [http://www.semicolon.no/wp-content/uploads/2014/05/  
eSamhandling\\_DIFI-CASE-rapport\\_versj\\_1\\_0.pdf](http://www.semicolon.no/wp-content/uploads/2014/05/eSamhandling_DIFI-CASE-rapport_versj_1_0.pdf)
- Om semantisk web:  
[http://no.wikipedia.org/wiki/Semantisk\\_web](http://no.wikipedia.org/wiki/Semantisk_web)
- Semicolon «Kokebok om informasjonsutveksling»
  - <http://www.semicolon.no/?p=364>

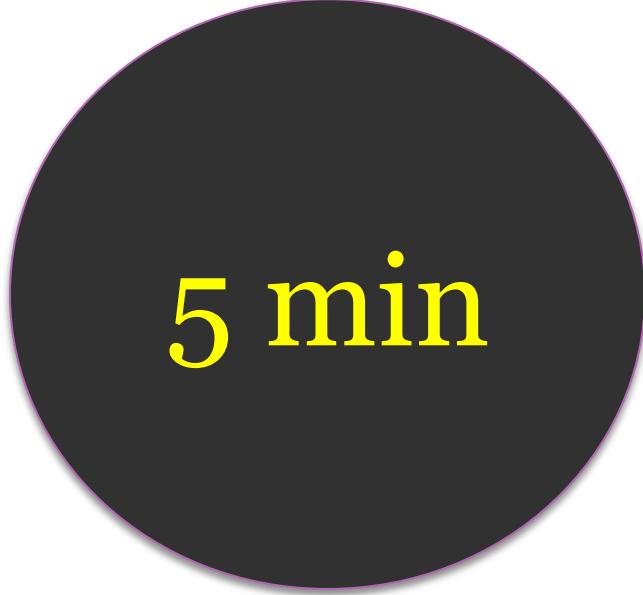
# Denne forelesning

- Basert på foiler fra Per Myrseth (DNV-GL)
- Basert på erfaringer fra Semicolon prosjektet
  - 2008-2013

# **Samhandlingsevne : Collaboration anyone?**

# Oppgave

- Diskuter 2 og 2
- Hva er sammenhengen mellom lover og semantiske teknologier?



5 min

# Hva er sammenhengen?

- Har lovgivning noe med semantiske teknologier å gjøre?
- Skal lovgivning være uavhengig av samfunnet vi lever i?
- Bør lovene speile den virkelighet de skal brukes innenfor?
  - Bør lovene arve prinsipper fra teknologien som skal brukes for å implementere lovene?

# Hva kjennetegner Offentlig forvaltning?

- Inndelt i sektorer (ministerstyre)
- Hver sektor har sitt sett med lover
- Lite samarbeide mellom sektorene
  - Store ytre enheter
- Reflekterer ”dårlig” at vi lever i 21.st århundre
  - Alle er på nett (hele tiden)
  - Forventninger til automatiserte tjenester
  - Forventninger til at informasjon er ”godt” organisert
    - Hvorfor må jeg oppgi informasjon jeg allerede har sendt inn, eller som myndighetene kjenner til?

# Hva kjennetegner myndighetenes ambisjoner for offentlig sektor

- Digitalt førstevalg
- Forenklinger
- Mer effektiv – behov for varme hender
- Større grad av automatisering

# Hva betyr dette?

- Lovene i hver sektor må støtte deling av informasjon mellom sektorene
- Lovene må støtte samhandling
- Lovene må være utformet slik at det er mulig å lage ”regler” som sikrer at loven følges automatisk
  - Det må være klart skille mellom saker som kan behandles automatisk og saker som krever skjønn og manuell behandling (xx-yy regel)

# Hva kjennetegner semantiske teknologier?

- Distribuert
  - Data er ikke samlet i store databaser, data kopieres ikke mellom enheter, men refereres til
- Basert på tillit?
  - Hvis vi skal ta semantiske teknologier i bruk i offentlig sektor, må vi ha tillit til de som forvalter informasjonen som legges ut
  - Dataene er oppdatert og riktig
- Åpenhet
  - Om dataene som ligger ute, hvordan de forvaltes, hva dataene omfatter, hvilket domene de dekker, gode beskrivelser av begrepene som brukes og omfattes av domenet

# RDF – den enkleste datamodellen

- Består av tre – 3 – kolonner
  - Subjekt
  - Predikat
  - Objekt

## **Subject**

<http://erlend.com/>  
<http://erlend.com/>  
<http://erlend.com/>

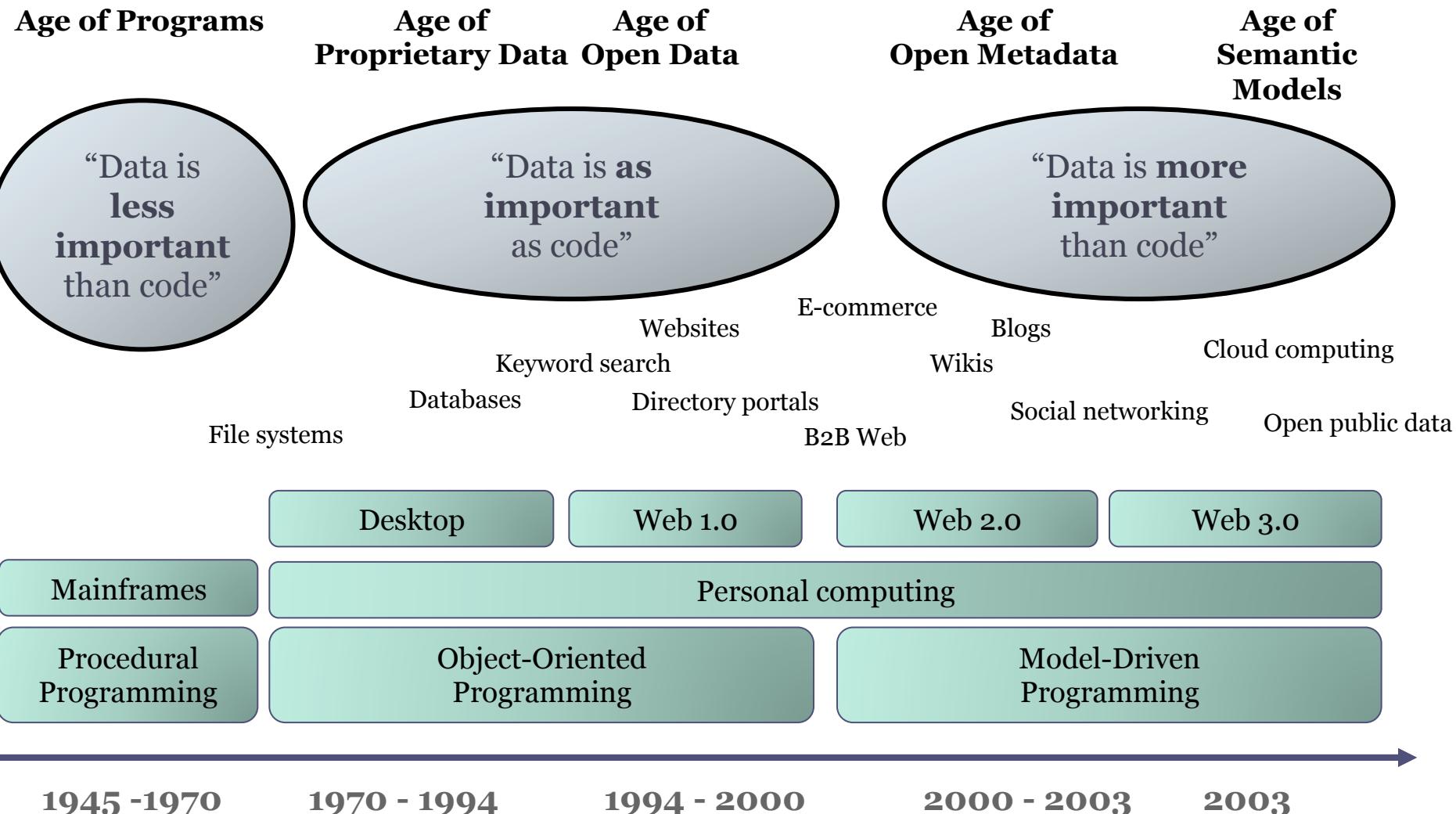
## **Predikat**

Foaf:name  
foaf:knows  
Foaf:email

## **Objekt**

"Erlend Øverby"  
<http://arild.com/>  
erlend@overby.com

# Semantic Technologies: From code to meaning



Inspired by: Nova Spivack (2007): How the WebOS evolves?  
See [http://novaspivack.typepad.com/nova\\_spivacks\\_weblog/2007/02/steps\\_towards\\_a.html](http://novaspivack.typepad.com/nova_spivacks_weblog/2007/02/steps_towards_a.html)

# Valg av utvekslingsformatvalg i nye prosjekter

veldig grov skisse

Formattyper

Formatet linker til semantikk

RDF

XML

SGML

EDIFACT

Records som ren text / ren ASCII

1970

1980

1990

2000

2010

2020

\$ 700.000,-

\$ 53.000,-

Eksempler på integrasjonsteknologi

\$ 0.01,-

„ekte“ søk, semantisk ”lus“ / Sparql

Linked open data

Web services

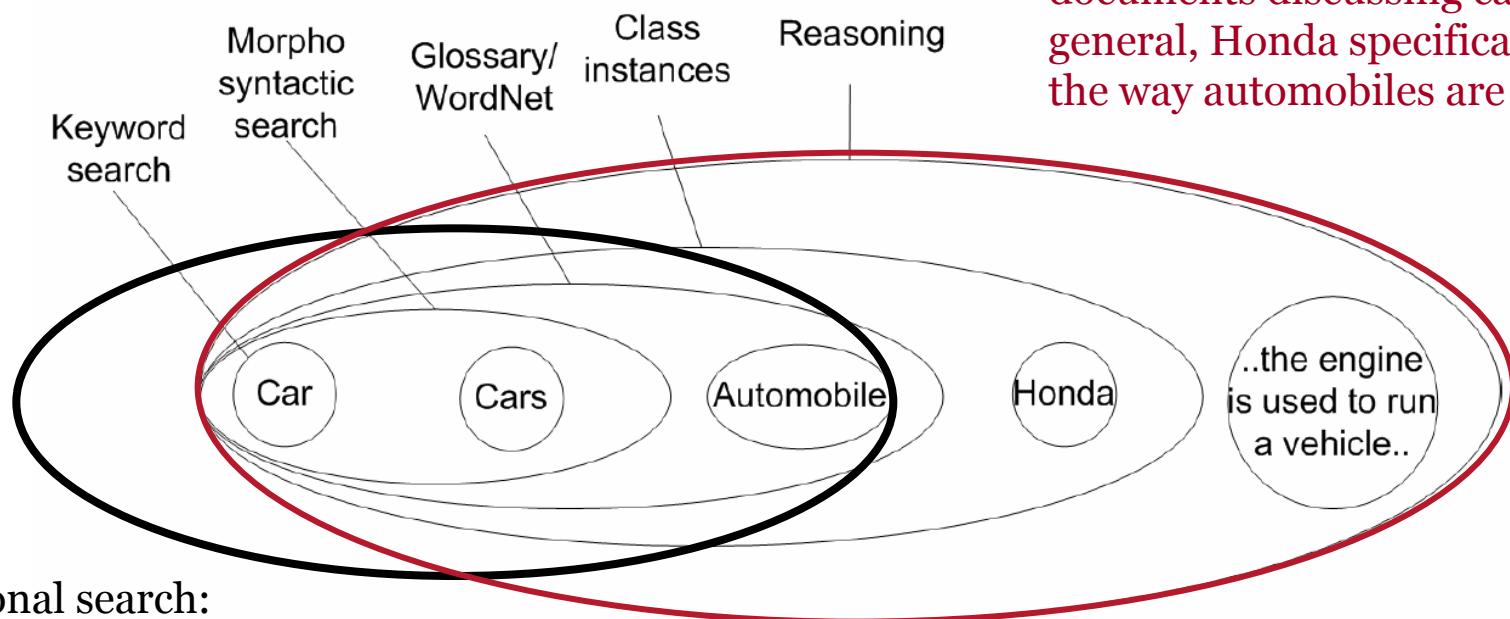
Integrerte systemer, felles masterdata

Kopi av masterdatakilder/registre

Stand alone systemer

Vanlige egenskaper ved systemer

# Semantic vs Traditional Search: Concepts instead of Keywords

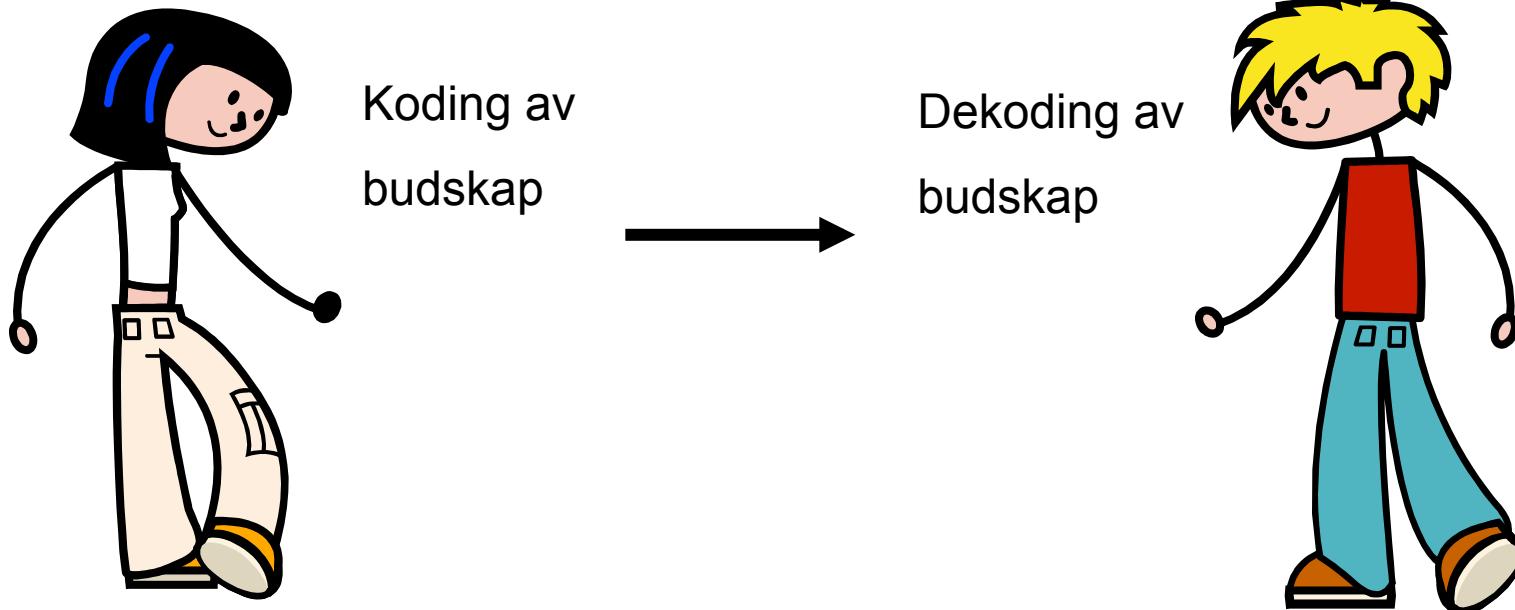


Traditional search:  
“car” returns documents  
containing the keywords *car*,  
*cars*, *automobile*

Semantic search:  
A search for “car” returns  
documents discussing cars in  
general, Honda specifically, or  
the way automobiles are built

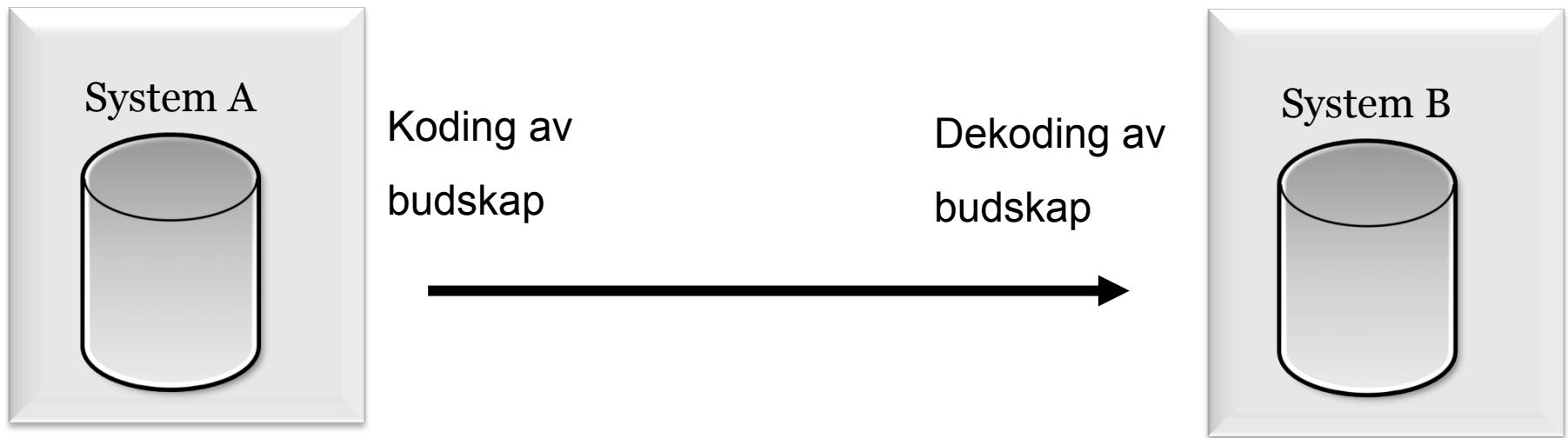
Samhandlingsevne og felles begrepsbruk i offentlig forvaltning

# Kommunikasjon mellom personer



Samhandlingsevne og felles begrepsbruk i offentlig forvaltning

# Kommunikasjon mellom systemer

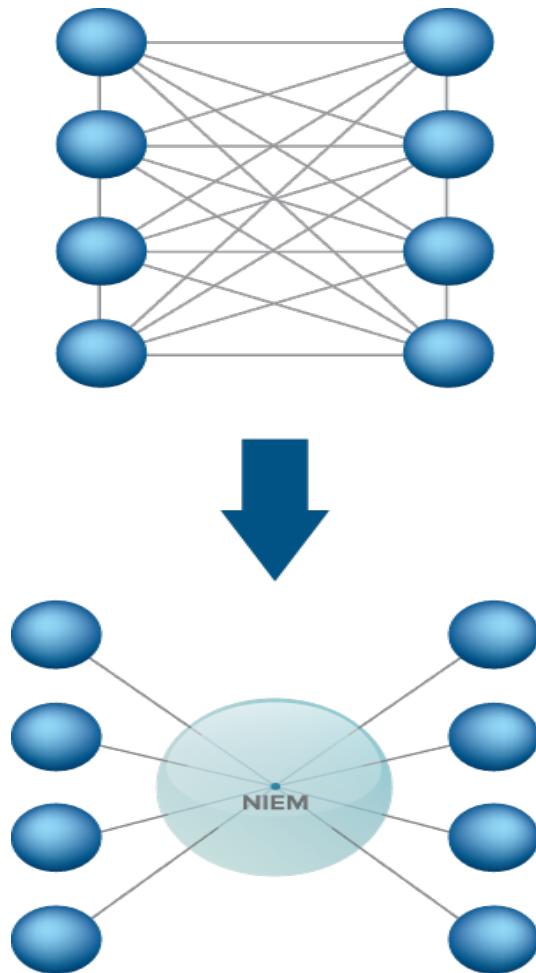


## Samhandlingsevne og felles begrepsbruk i offentlig forvaltning

# Kommunikasjon

	Mellom personer	Mellom systemer
Utveksles ved	Tale	Sekvenser av bits i en henholde til en predefinert teknisk struktur
Hva utveksles	Ord og uttrykk (begrep) i et felles språk	data/verdier som tolkes i henhold til felles språk/terminologi/begrep
Regler for bruk av ord	Grammatikk	I henhold til teknisk format og struktur på meldingen, samhandlingsscenario og informasjonsmodell
Kontekst	Kjent for sender og mottaker	Kjent for sender og mottaker
Intensjon ved kommunikasjon	Varierer veldig	Kun i henhold til predefinert samhandlingsscenario, suksess og feilsituasjoner for scenariet må være forhåndsdefinert og maskinelt håndterbare.

# Peer to Peer alle med sitt språk, versus Peer to peer med felles språk



Figuren kan være gyldig på minst tre måter:

- Juridisk interoperabilitet
- Semantisk interoperabilitet
  - Begrep
- Teknisk interoperabilitet
  - XML bruk
  - Bæreprotokoll
  - Sikkerhet

# Data and metadata kvalitet

- Bør du benytte / gjøre deg avhengig av andres data om :
  - Du er usikker på om de er korrekt eller fullstendig?
  - De ikke er tilgjengelig når du trenger dem?
  - De ikke er etablert og forvaltet på en lovlige og fagmessig god måte?
  - Betydning og definisjoner for data ikke er åpent tilgjengelig?
  - Du ikke forstår data fullt ut?
  - Betydningen endres hyppig, dvs betydningen av data er uforutsigbart for dine systemer og rutiner?
  - Koblingen mellom tolkning av data og jus-kontekst data skal benyttes i er uklar.

# Semantikk

- Hvordan overføre mening til datamaskiner, slik at vi kan dele data

Vi kan anta at det er et telefonnummer:

+47 90129642

Vi kan anta at det er et tidspunkt:

22:43

Litt vanskeligere:

Vi har behov for mer informasjon for å sikre riktig tolkning.

42

# Interoperabilitet:

- Politisk
  - Rammer, forutsigbarhet, internasjonalt
- Juridisk
  - Lover, sammenheng mellom lover, begreper brukt i lover
- Organisatorisk
  - Hvordan organisasjoner arbeider, tjenester ytes, arbeidsprosesser, tjenestedesign osv.
- Semantisk
  - Begreper, definisjoner
  - Tillit, kvalitet, sikkerhet
- Teknisk
  - Protokoller og formater for deling av data

# Norsk offentlig forvaltning

Setter lover og regler for hvordan en etat arbeider

Etterfølger lover og regler

Arbeidsrutiner

En modell av virksomheten

IT systemene støtter

Realiseres av



# Koble lov og programkode

- Dette er ikke et stabilt system
  - Lover endres
  - Prosedyrer endres
  - IKT infrastruktur endres
- Behov for at det henger sammen
- Påvirkes av:
  - EU lovgivning (EØS)
  - FN resolusjoner
  - Lov praksis
  - Klager på saksbehandling
  - Nye paradigmer (smarttlf)

Lover

Organisering

Prosedyrer

Virksomhetsarkitektur

Informasjonsmodell

Datamodell

Kjørende kode

# Samarbeid / samhandling mellom etater

- For å kunne lage gode innbygger tjenester må systemer kunne dele data
- For å gjøre offentlig sektor mer effektiv må data deles
- Digitalt førstevalg forutsetter stor del av datadeling
- Situasjonsbaserte tjenester forutsetter datadeling mellom:
  - Stat
  - Fylke
  - Kommune
  - Private aktører

# Samhandling mellom etater

Lover

Organisering

Prosedyrer

Virksomhetsarkitektur

Informasjonsmodell

Datamodell

Kjørende kode

Legale definisjoner

Harmoniserte arbeidsprosesser

Felles modeller

Fysisk lag for deling

Mellom begrep og definisjon

Mellom begrep og definisjon

ing

r

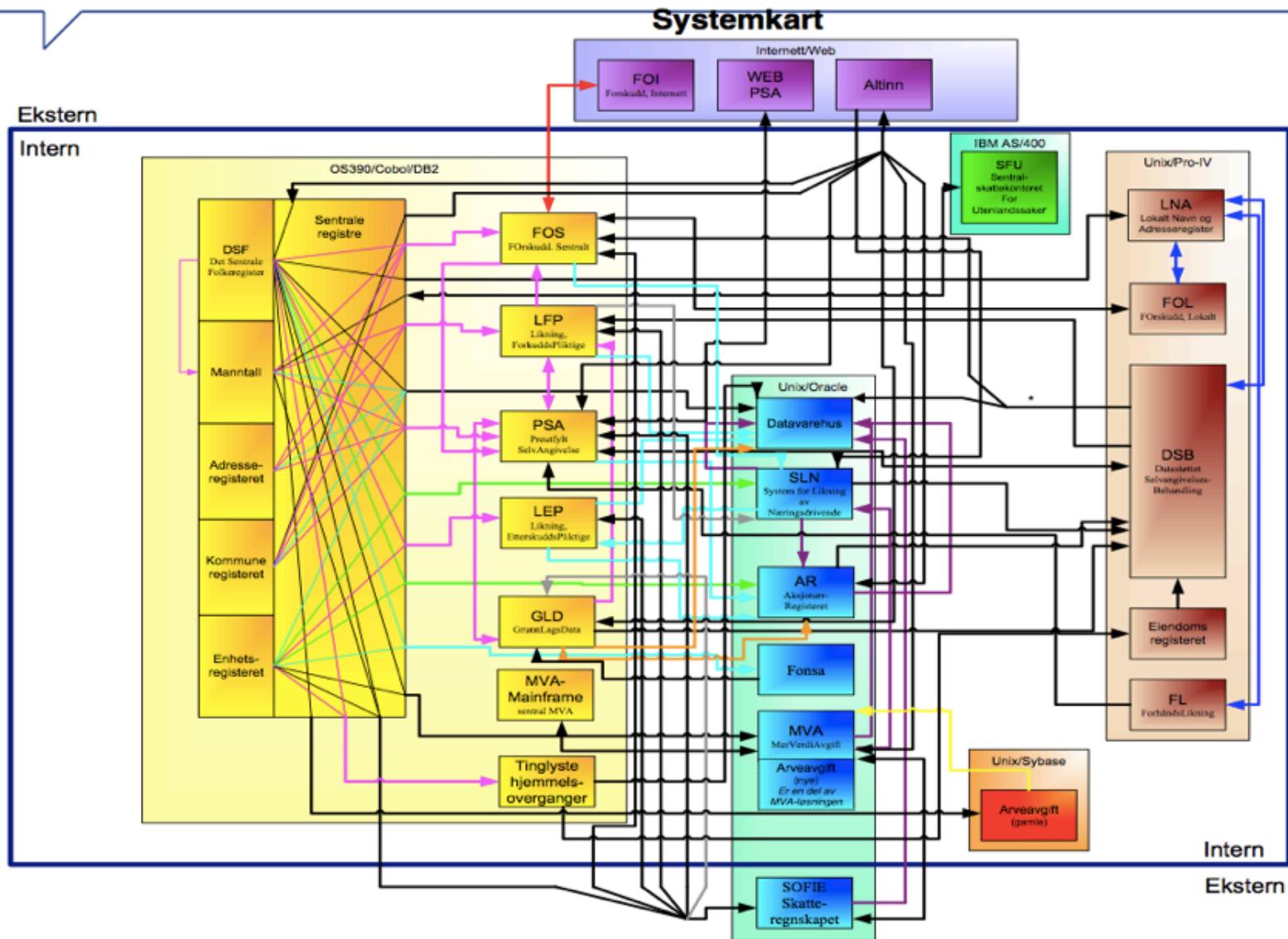
etsarkitektur

onsmodell

ll

kode

# Forenklet modell av en del av IT-systemene...



# Oppgave:

- Video:

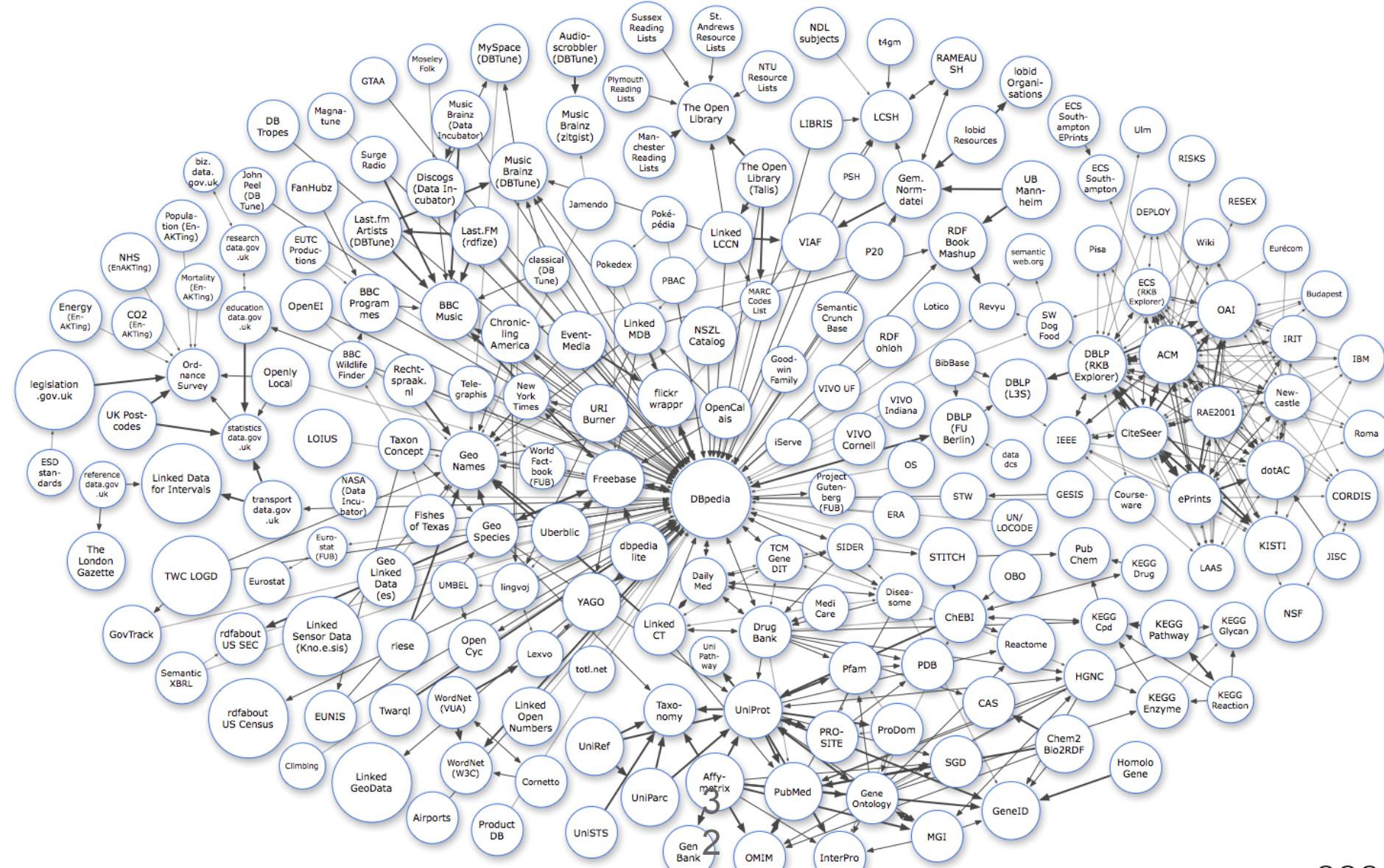
<http://www.gapminder.org/videos/200-years-that-changed-the-world-bbc/>

## Beskriv og drøft:

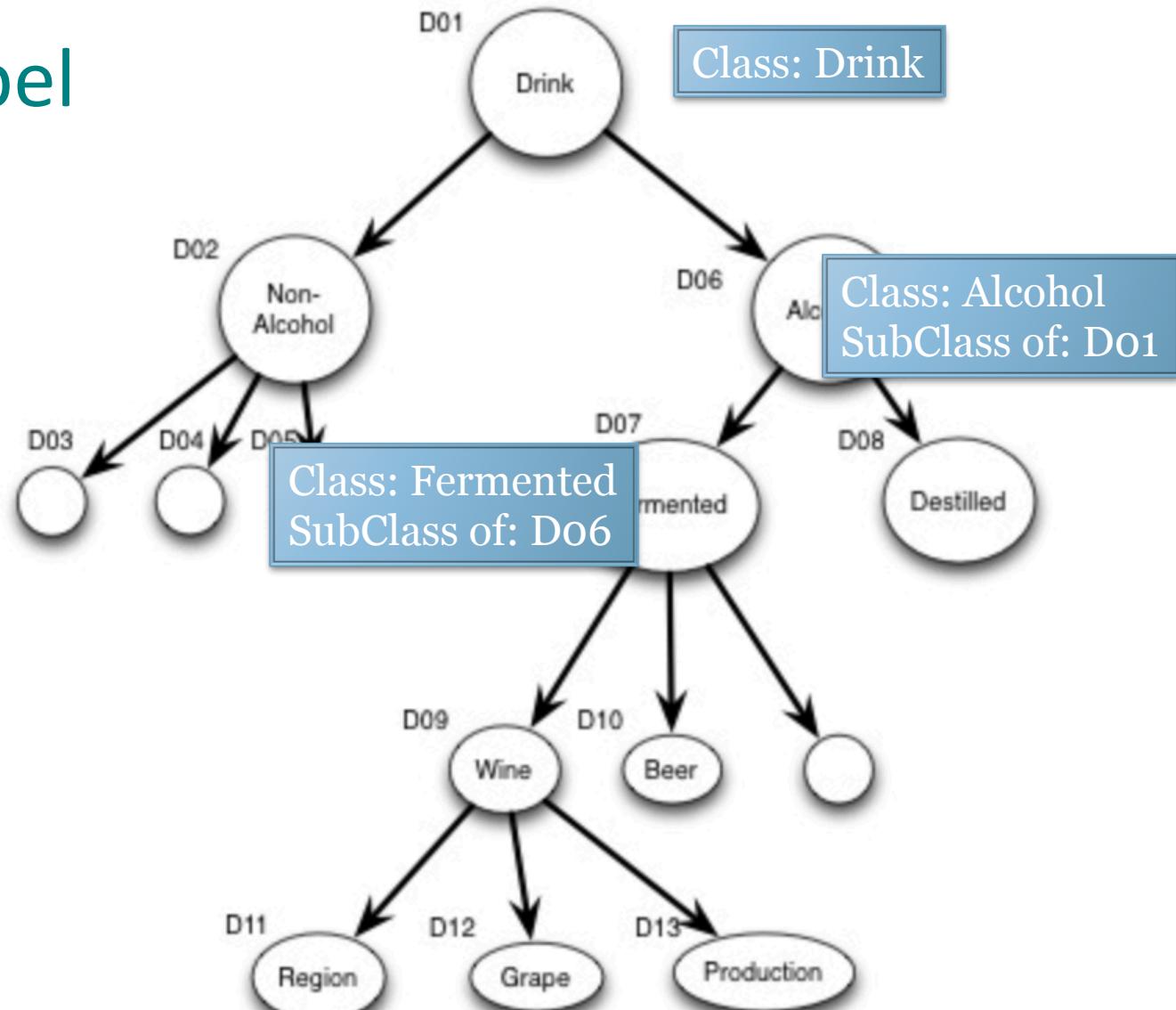
- Hvilke begreper brukes
- Behov for harmonisering av begreper og tilhørende data
- Behov for forvaltning av begreper og tilhørende data

# Åpne data

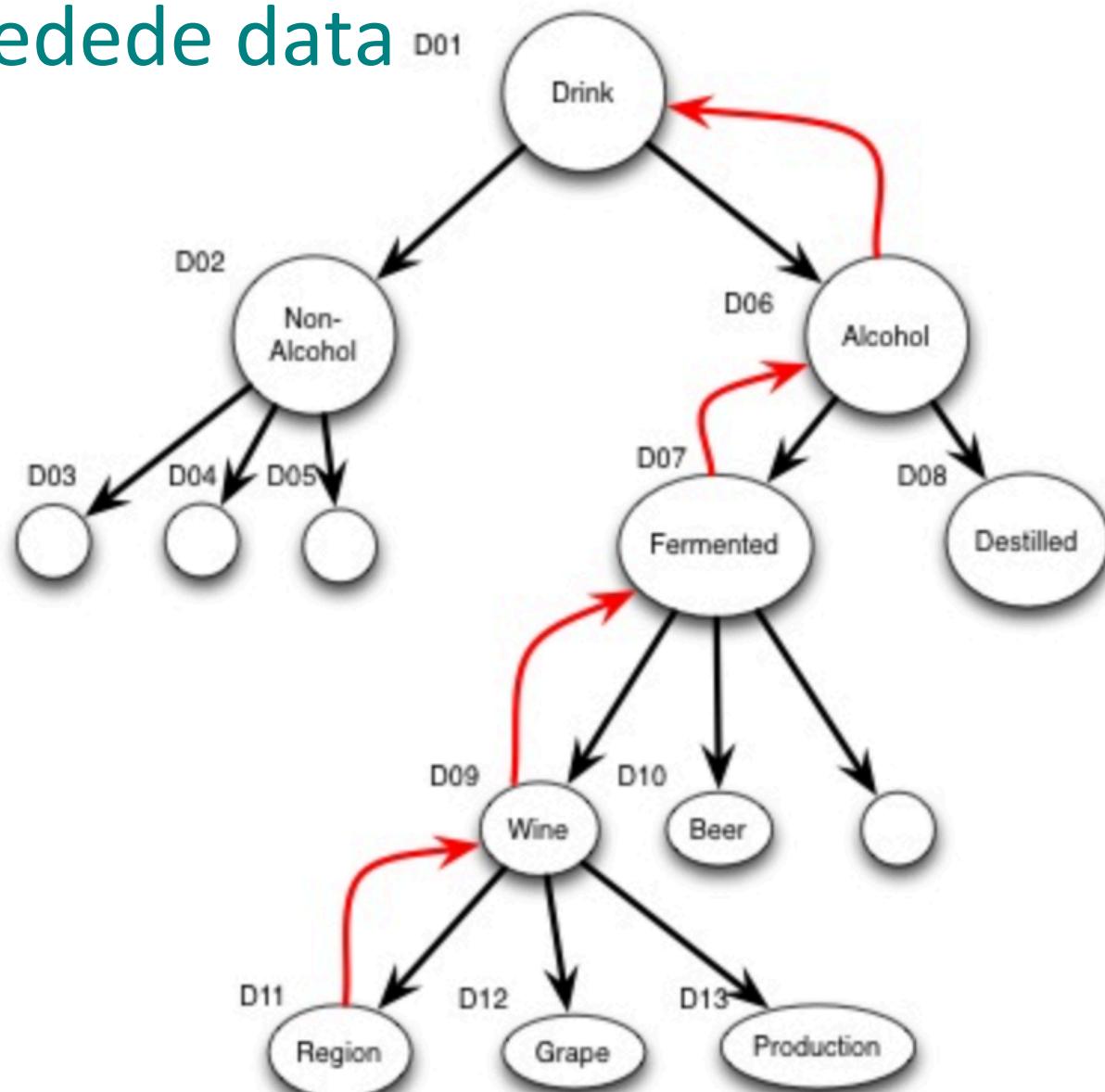
# Linked open data / del av web 3.0



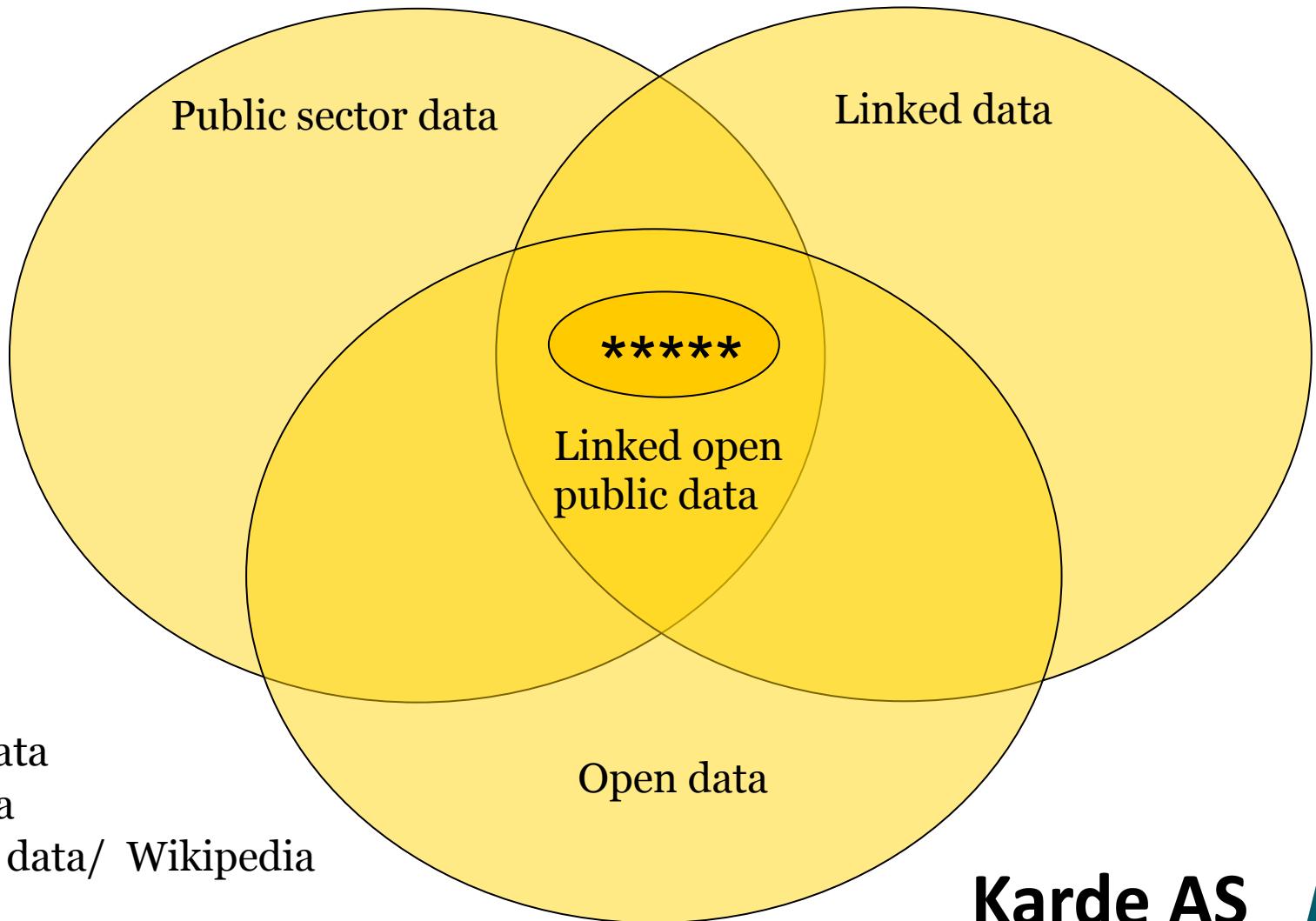
# Enkelt eksempel



# Eksempel på utledede data



# Intersection of some terms



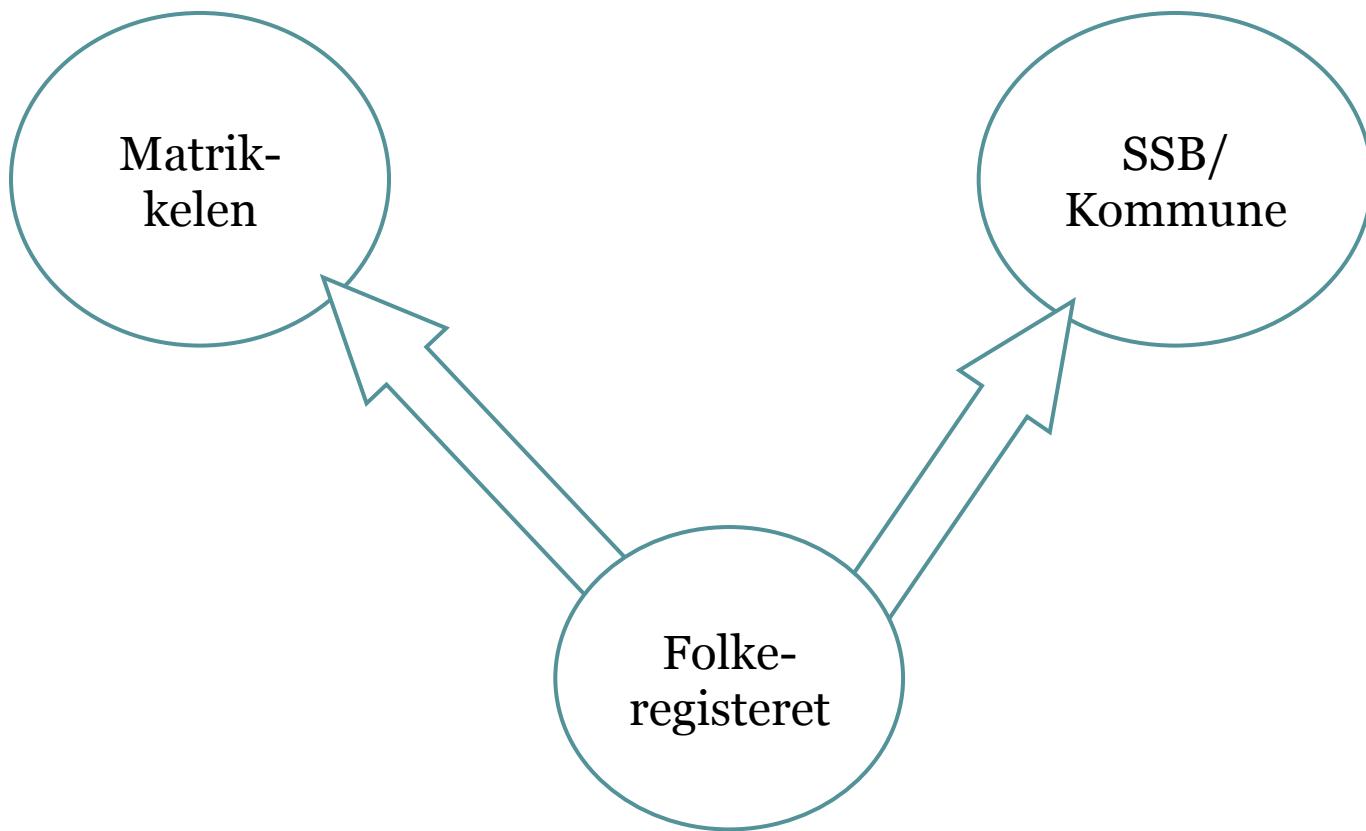
# 5-star deployment scheme for Linked Open Data

★	Available on the web (whatever format) but with an open licence, to be Open Data
★★	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
★★★	As (2) plus non-proprietary format (e.g. CSV instead of excel)
★★★★	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
★★★★★	All the above, plus: Link your data to other people's data to provide context

# Hva betyr dette for offentlig sektor?

- Folkeregisteret og SSB
- Folkeregisteret har ansvaret for opplysninger om personer og familie
- SSB har ansvaret for kommunenr og navn
- Matrikkelen er oversikten over eiendommer, adresser og bygninger.

# Folkeregisteret



# Eksempler på Linked Open Data

- Enhetsregisteret Brønnøysund
- Wikipedia
- Utdanningsdirektoratet og læreplanmål
- Kjennetegnes ved at de har et SPARQL grensesnitt som det kan gjøres spørninger mot

# Data.computas.com - Enhetsregisteret

- Tilgjengelig som åpen data
- Brønnøysundregistrene legger nå ut sine data som åpne lenkede data for bruk av andre etater
  - I dag stor grad av kopiering inn i systemer (daglig/ukentlig)
- Gir informasjon om en enhet



## Search results

Show search form

## SLAVIANOR AS

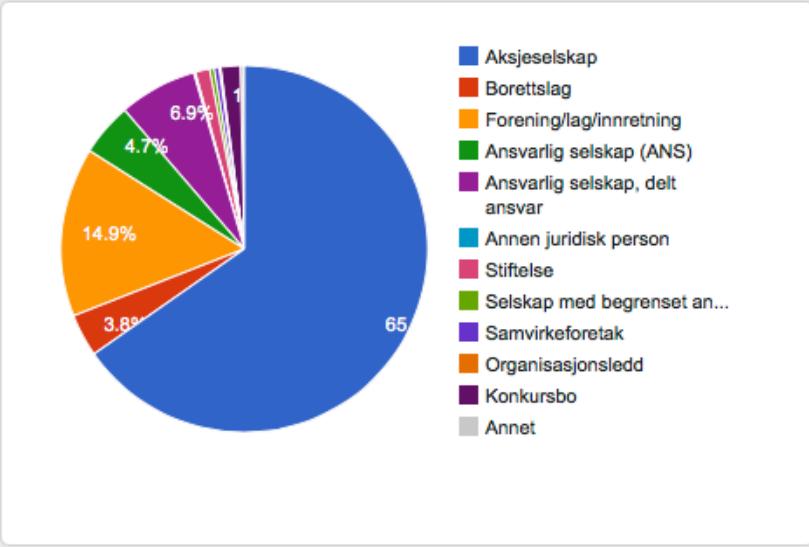
navn SLAVIANOR AS

## VIDEONETT NORGE AS

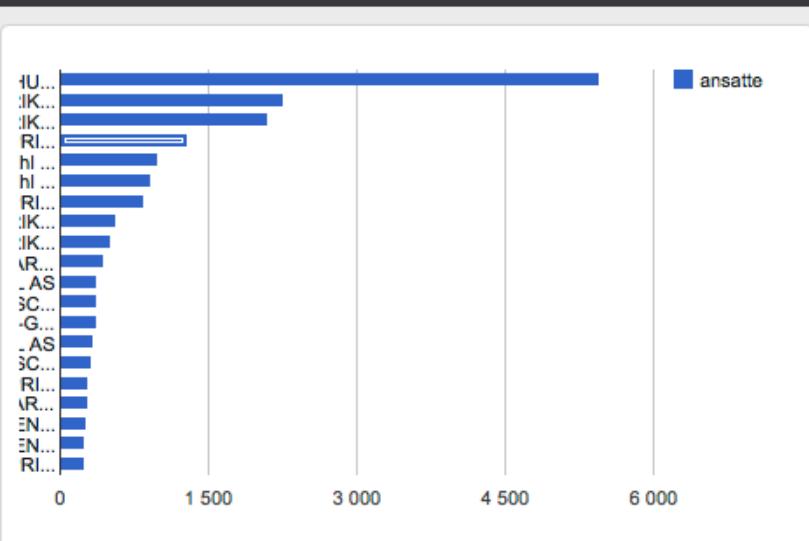
navn

STE  
navn  
BORG  
navn  
HOLMEN BORETTSLAG  
navn  
FREDRIKSTAD BAPTISTMENIGHET  
navn  
STAR EIENDOM ANS  
navn  
PROFFER GLASS ENGROS AS  
navnVirksomheter i  
Fredrikstad

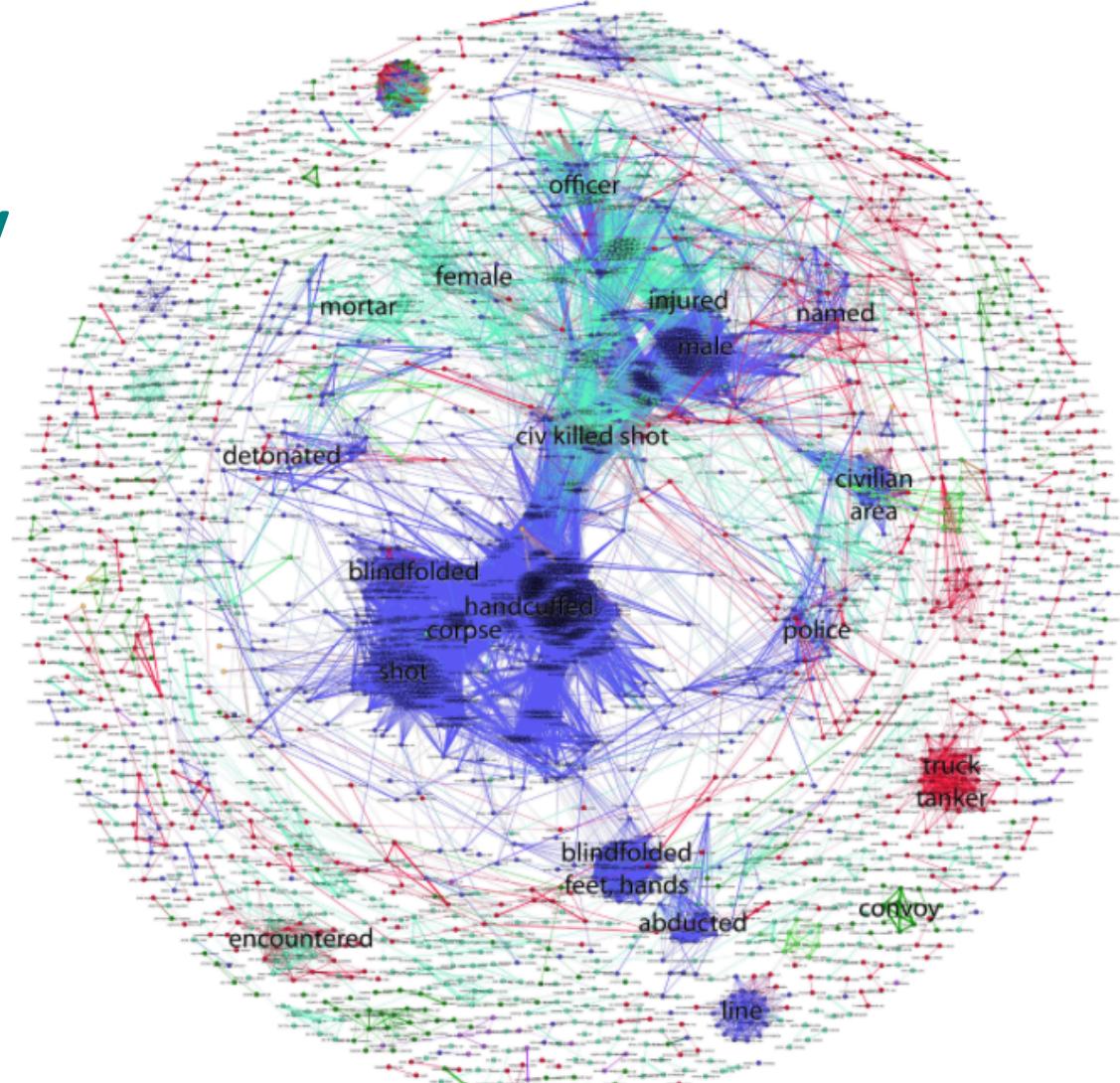
## Organization form distribution for municipality 0106



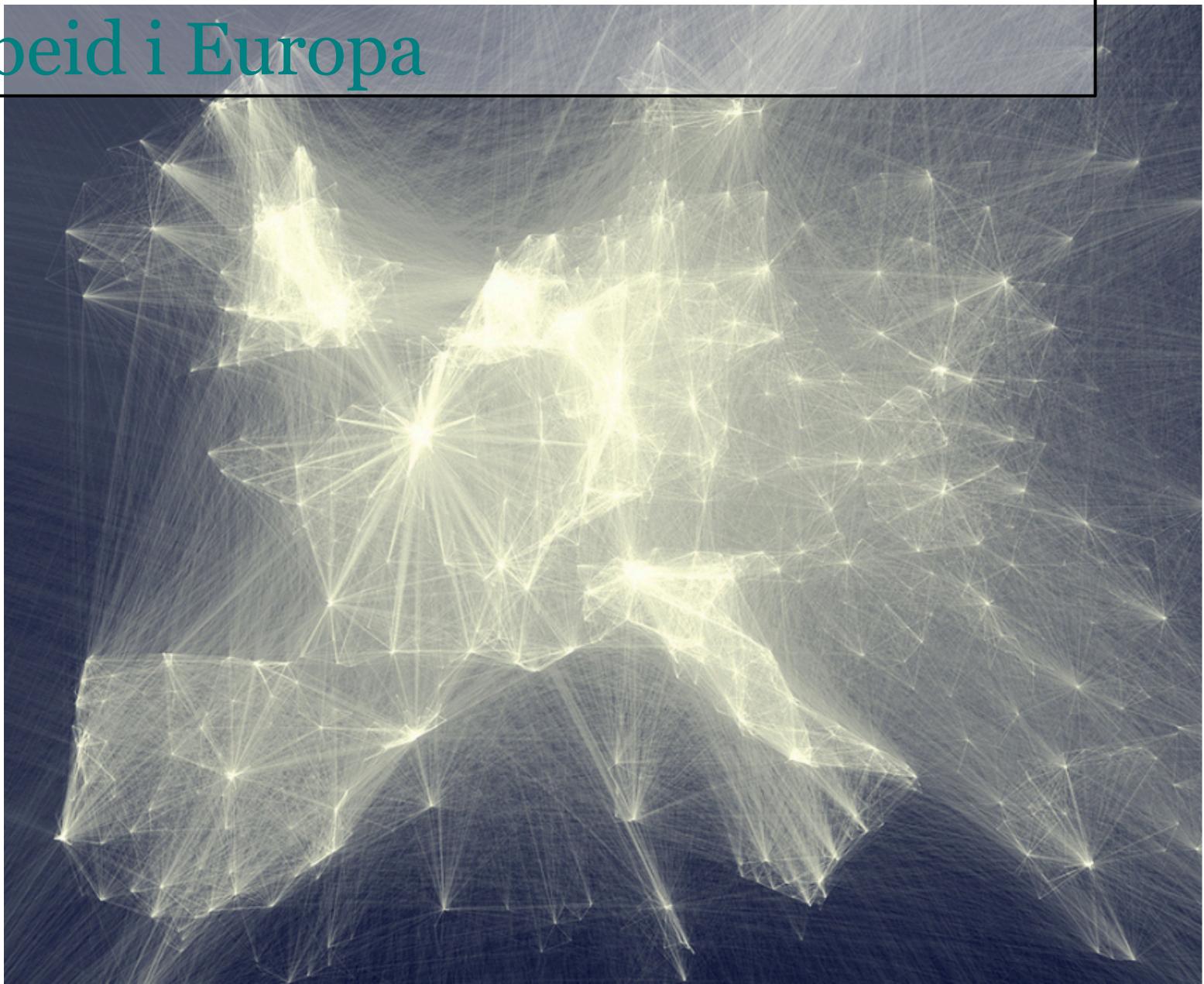
## Organizations with most employees in the municipality 0106



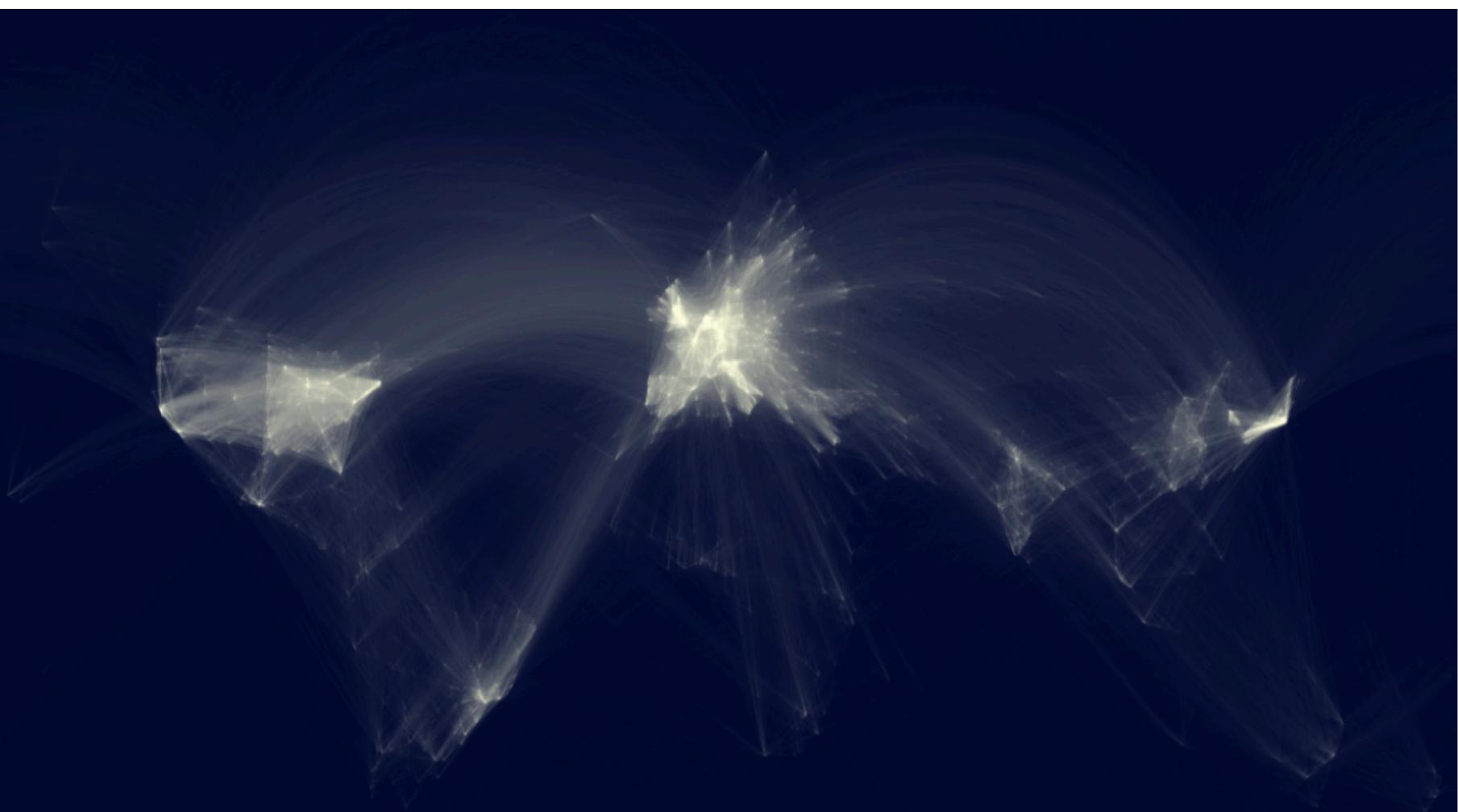
# Visualisering av åpne data



# Samarbeid i Europa



# Map of Science Collaboration 2005-2009



Map of scientific collaborations from 2005 to 2009

Computed by Olivier H. Beauchesne @ Science-Metrix, Inc.

Data from Scopus, using books, trade journals and peer-reviewed journals

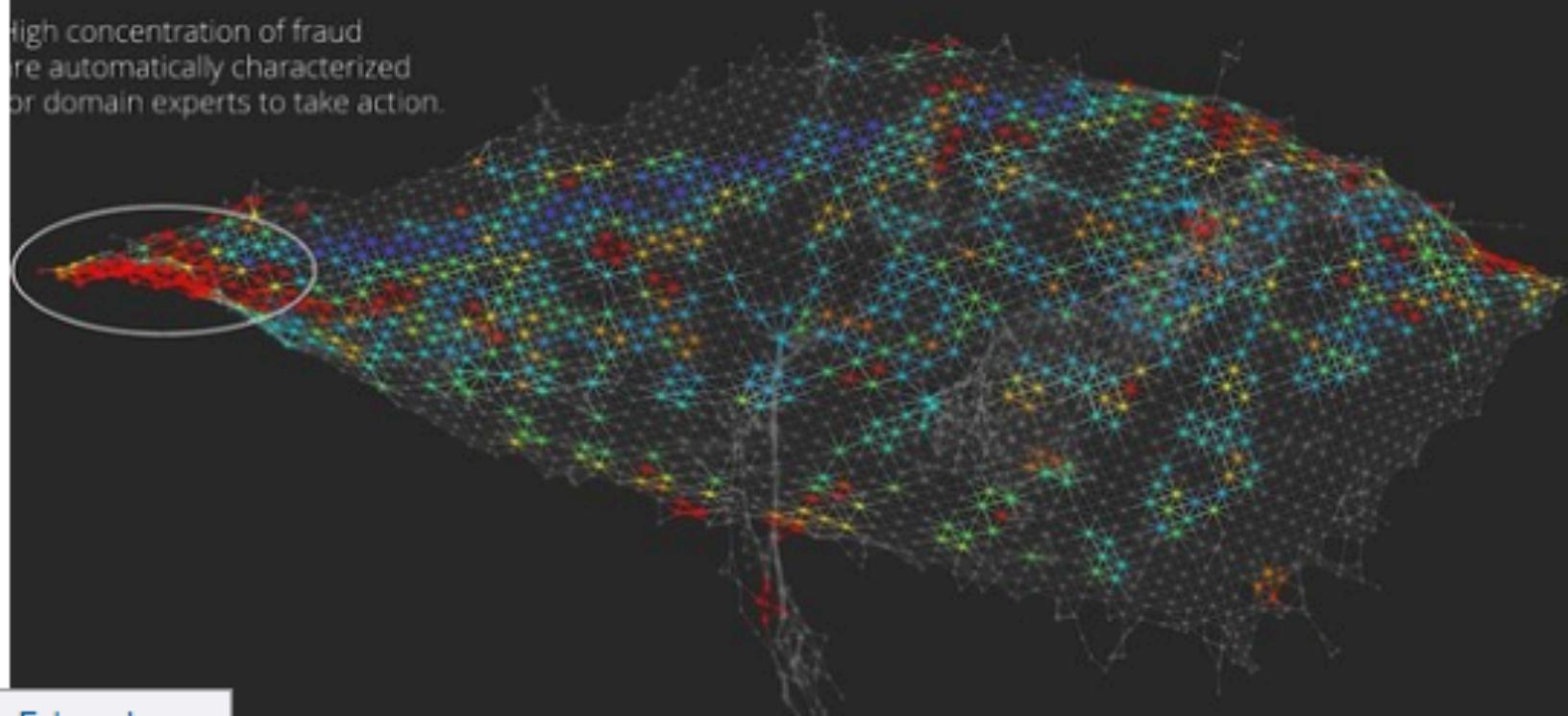
# System brukt av banker

## Automatic Fraud Discovery

AYASD

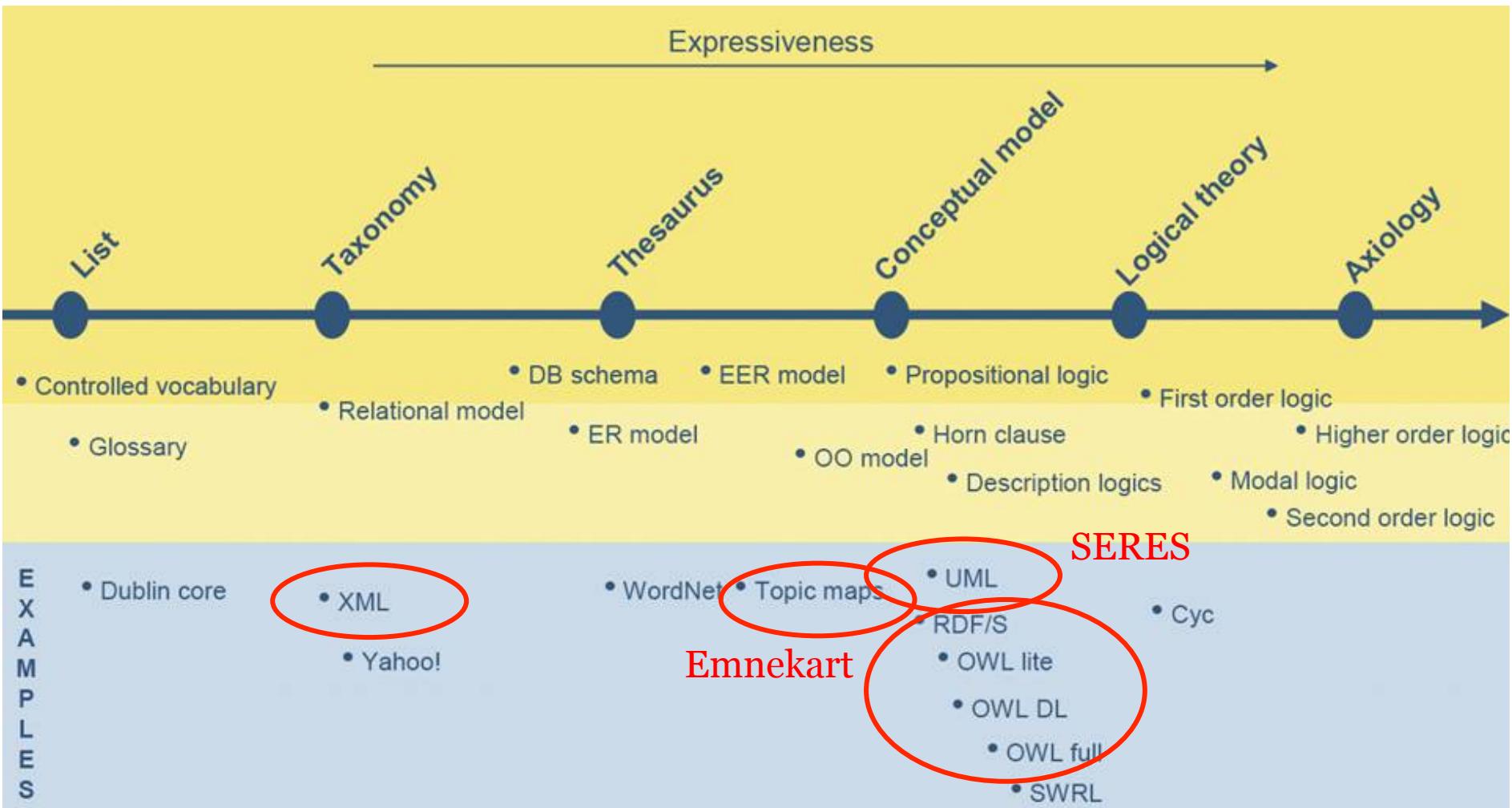
Fraudulent Transactions

high concentration of fraud  
are automatically characterized  
for domain experts to take action.



[Enlarge Image](#)

# Uttrykkskraft og formalisme i begrepsmodeller II

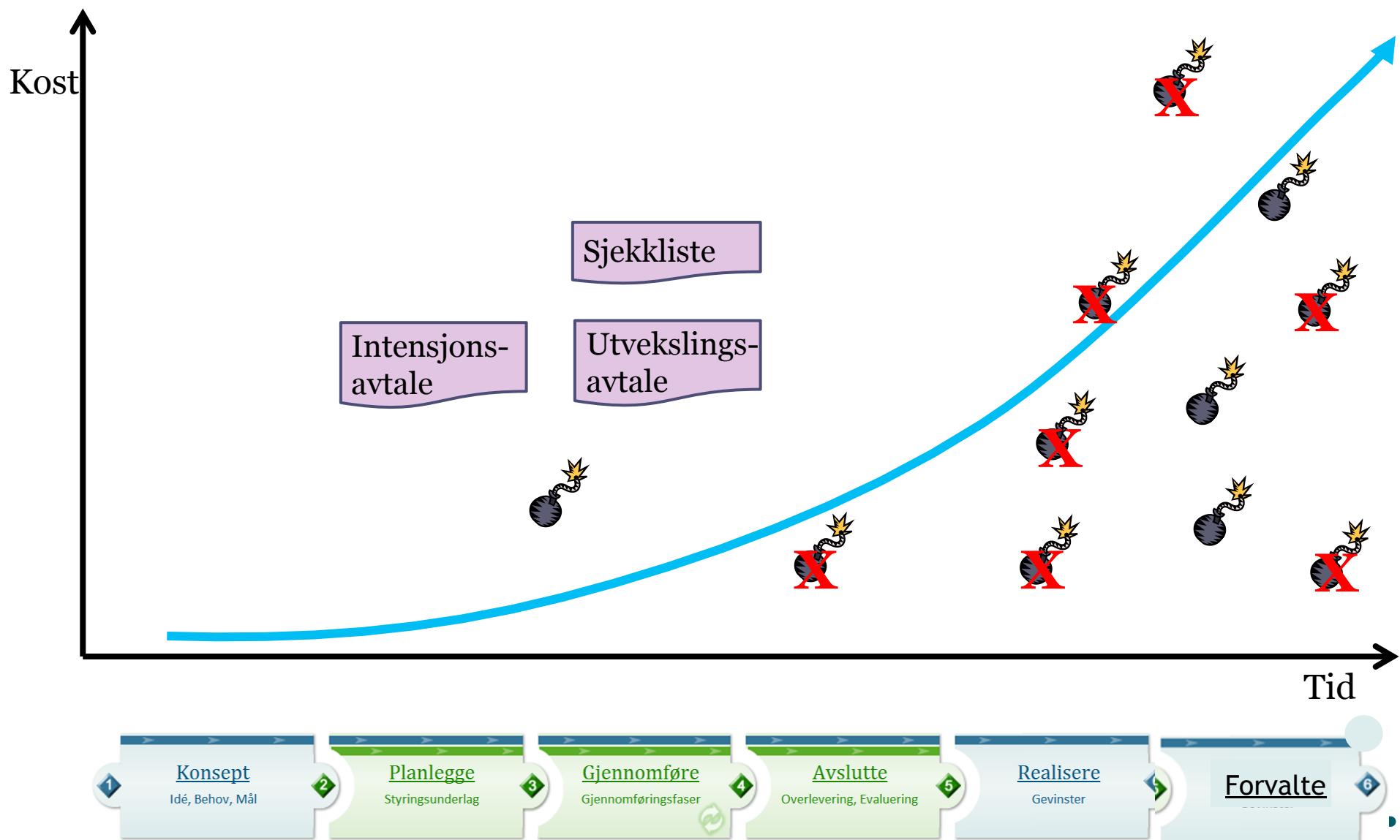


Source: Davis, M. 2006, Semantic Wave 2006 [Davis M]

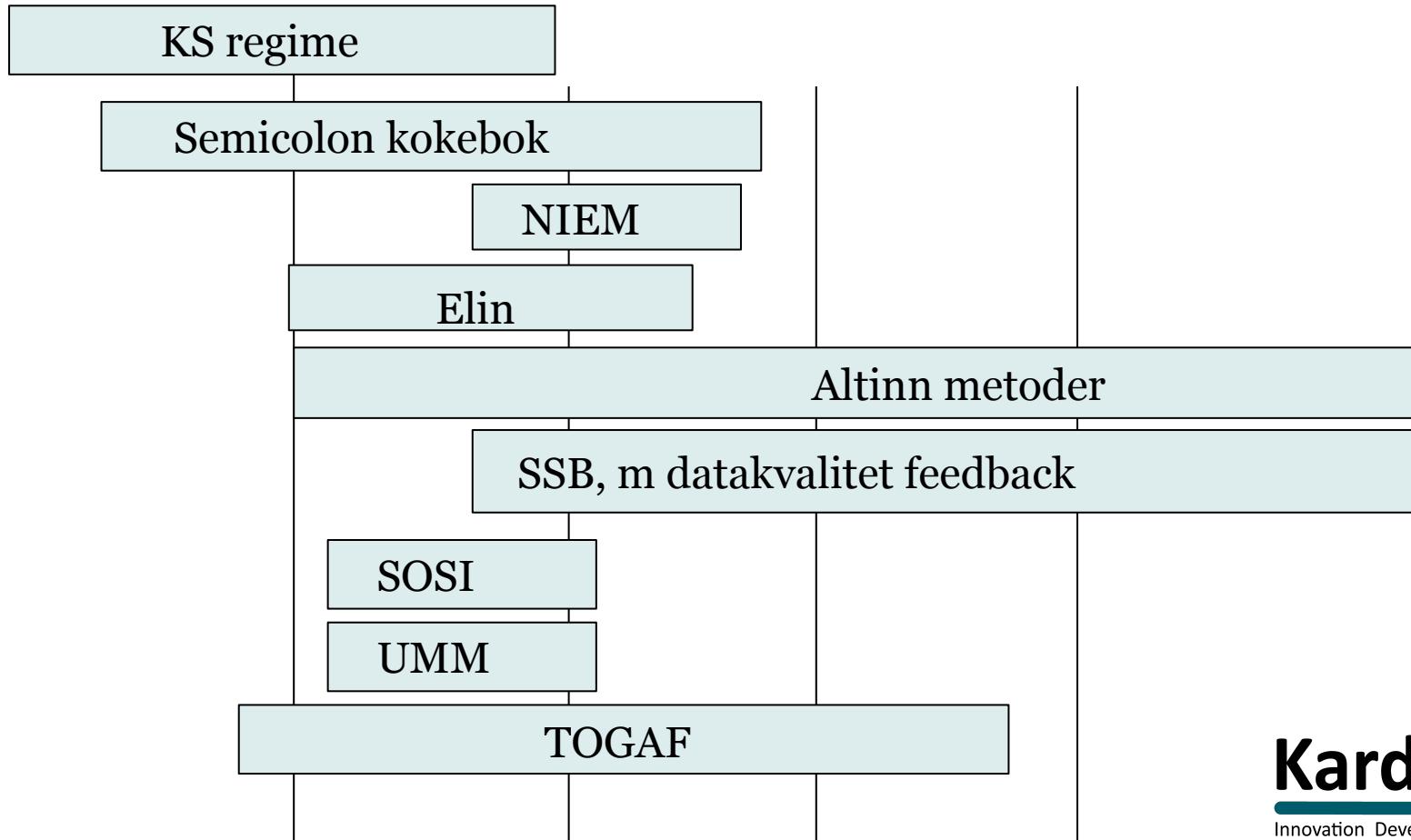
# Metoder for å etablere elektronisk samhandling

# Erfaring:

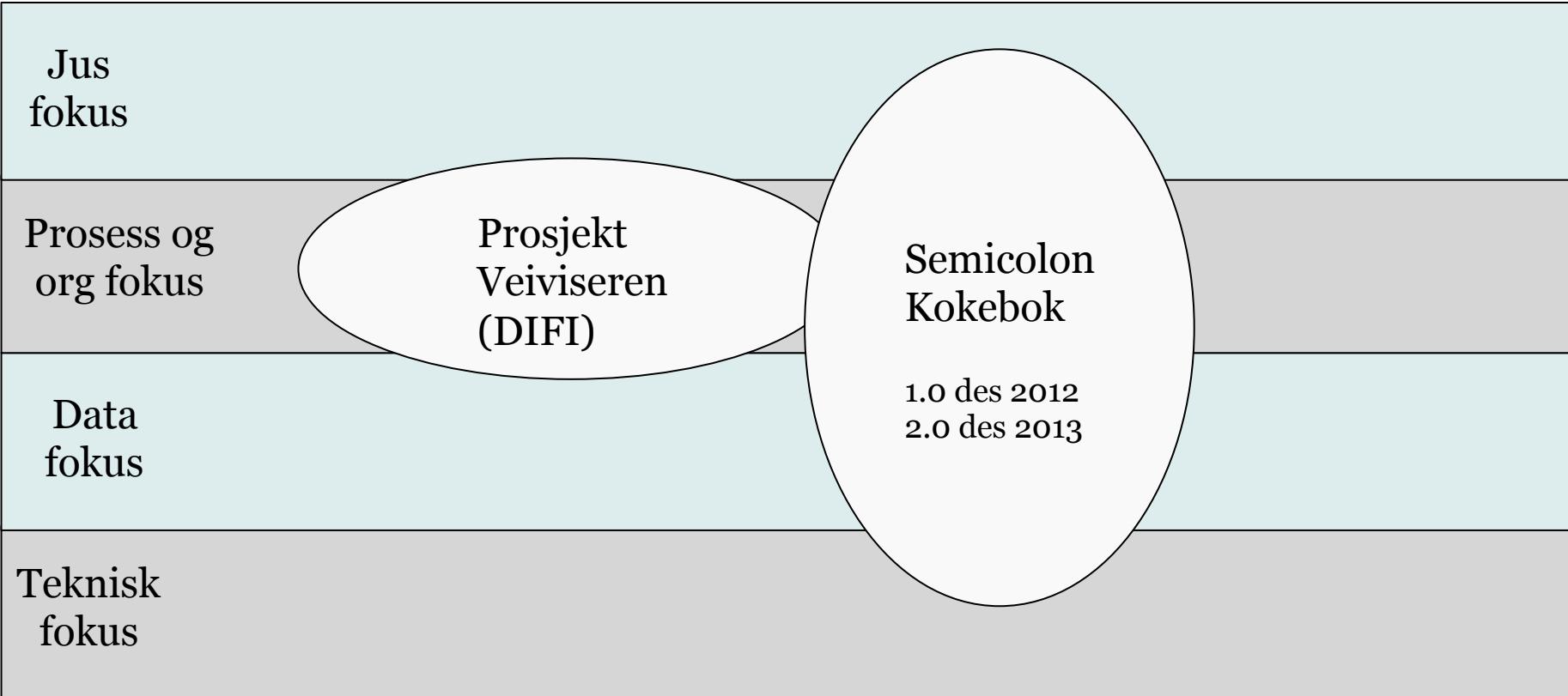
1. Elektronisk samhandling er utfordrende
2. Kostnad ved både å etablere og endre samhandling er høy



# Metoder for å etablere krav og design til elektronisk samhandling, metodenes



# Metoder for å etablere krav og design for elektronisk samhandling

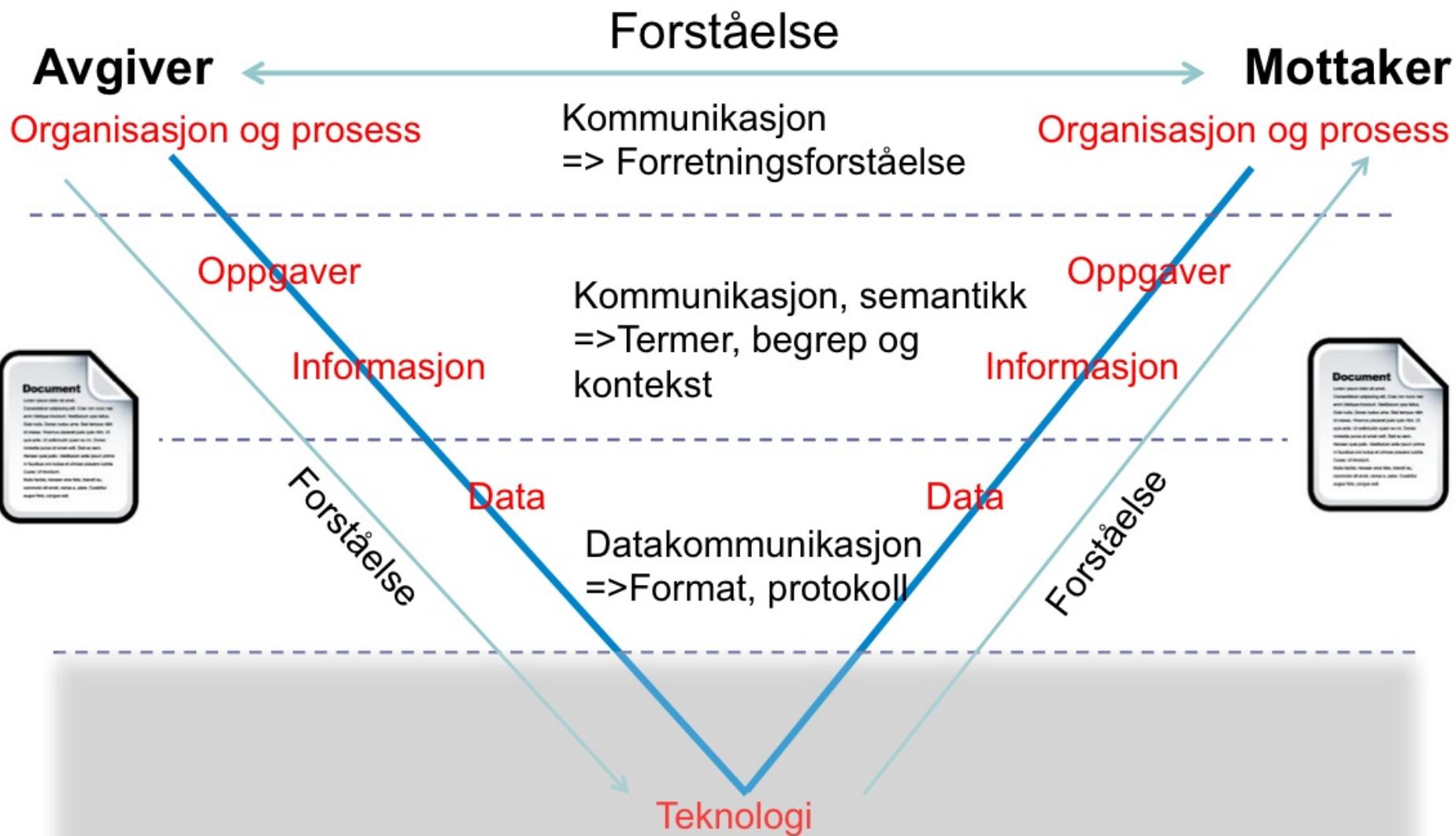


Egnet for en  
virksomhet



Egnet for flere  
virksomheter

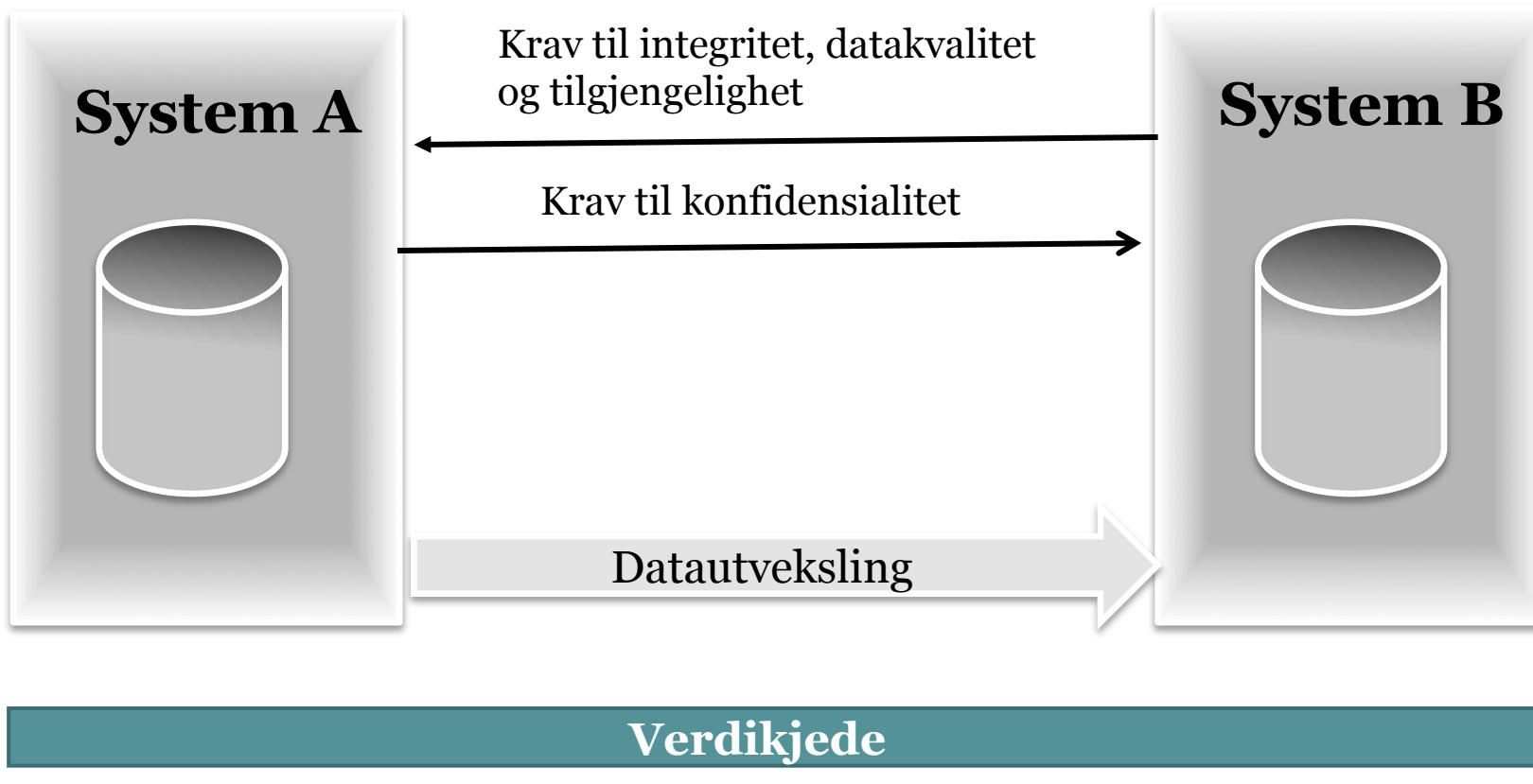
# Kommunikasjon på alle plan



# Egenskaper ved informasjon i verdikjeder

Sjekk hjemmel til å avgjøre aktuelle data?

Sjekk hjemmel til å få eller bruke aktuelle data?



# Andre modeller for samhandling



# CEN-CENELEC-ETSI Smart Grid Coordination Group

## Smart Grid Reference Architecture

System A

Economic / Regulatory Policy

Business Objectives

Business Procedures

Business Context

Semantic Understanding

Syntactic Interoperability

Network Interoperability

Basic Connectivity

Interoperation

System B

Business Layer

Function Layer

Information Layer

Communication Layer

Component Layer

# CEN-CENELEC-ETSI Smart Grid Coordination Group

## Smart Grid Reference Architecture

