

A nation's presence on the Semantic Web



Semantic Days Tutorial 31. may 2010

Per Myrseth (Semicolon)

Robert H.P. Engels (Sesam4)





Why are we here?



Promote common information design principles on the internet!

Represent everything as objects

and provide (defined!) relations between objects



Open enterprise/ gov data

National enterprise master data

Statistical data

Open public/social data

LOD central in global/European Cultural Heritage

LOD bringing together public information and information by the public

Enabling you to join LOD!

WP leader Semicolon - semantics

Project leader SESAM4

Chairman DND FG semantic web

The stairway to heaven (not Led Zeppelin this time)

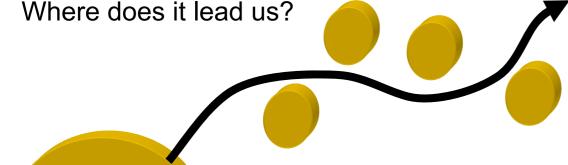


Where does the stairway start?

What are the stairs made of?

What is the sequence of the stairs?

A nation's presence on the Semantic Web



Target situation

Public sectors big hairy goals and the effect of using LOD



Public sector goal is to be **open**, **transparent**, **accountable**, **accessible**, **user-friendly**, **efficient**, **effective** and **service-oriented**.

The anticipated effect of Linked open data:

- Will open up PSI for innovation and makes public sector more transparent and accountable
- Utilize the internet as an infrastructure for access and scalability
- Enables more and cheaper user-friendly eGov services
 - Give the information back to the people
- Enables public sector to remove silos, lower integration cost and become more efficient and effective.
- Is a prerequisite in a service oriented public sector

Status today



Current Best Practice – examples from CH

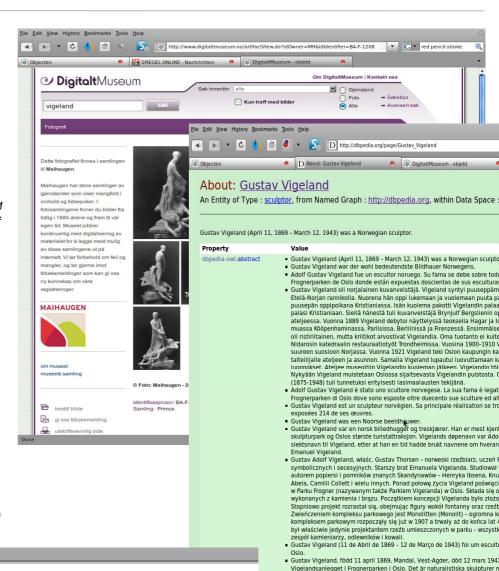


giort den medali som utdelas i samband med Nobels fredspris

Густав Вигеланд — норвежский скульптор, создатель Парка скульпт



- Collection of Norwegian Musea
- Digital representations of Cultural Heritage
 - Grieg, Vigeland, Munch, etc
- No globally identifiable objects
 - No explicit URI, but search on URN (internal foreign key) possible
 - http://www.digitaltmuseum.no/artifactView.do?image=&imageIndex=0&idOwner=M H&idIdentifier=PMF.7.0.01452&pageNo=1&noOnPage=12&noInResult=1&owner= &criteria=PMF.7.0.01452&searchObjectType=Unknown&onlyWithPictures=&lastP ageNo=&filterCriterias=
- Centered around works of art
- Positive:
 - Merging of collections from different sources
 - Artworks as objects with characteristics
- Negative:
 - No global identifiers
 - No relationships → navigation through unfocussed search in text strings
 - No automated access possible (no merging of sources)
 → why would you like to do that?



http

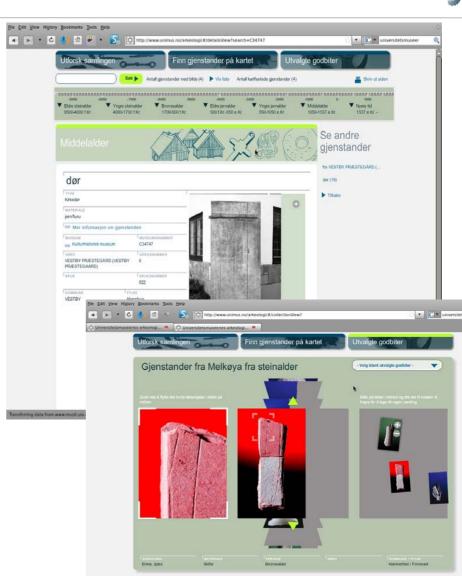
http://dbpedia.org/resource/Gustav_Vigeland

Current Best Practice – examples from CH





- Collection of Norwegian Archeological Objects
- Digital representations of Cultural Heritage
 - Doors, tools, objects from diverse epocs
- No identifiable objects
 - No explicit URI, but search on URN (internal foreign key)
- Centered around objects
- Positive:
 - Merging of collections from different sources
 - Nice flash-based interactive user interface
- Negative:
 - No globally identifiable objects (URI)
 - No publishing of metadata according to standards (f.ex DC)
 - Metadata not object-based → navigation through unfocussed search in text strings
 - No automated access possible (no merging of sources)



BRs Register of Business Enterprises

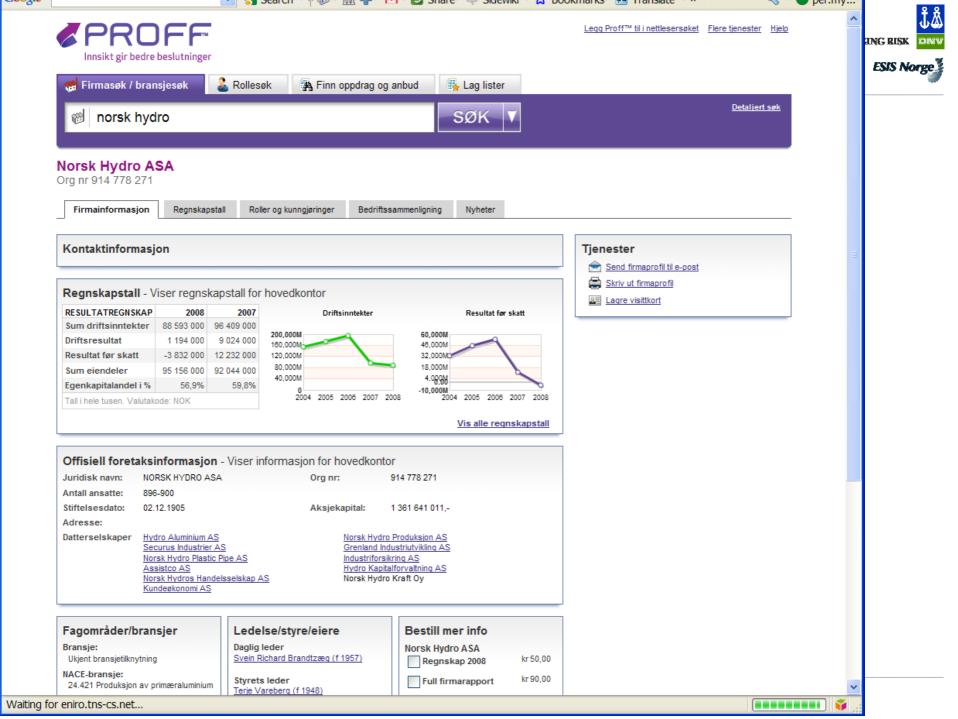
Done

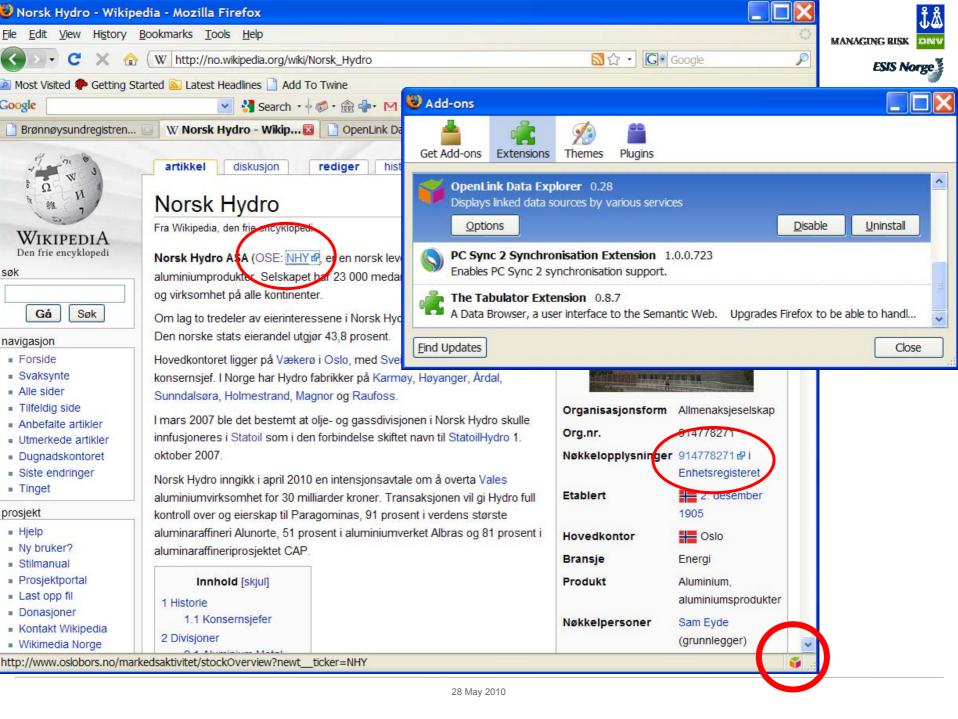


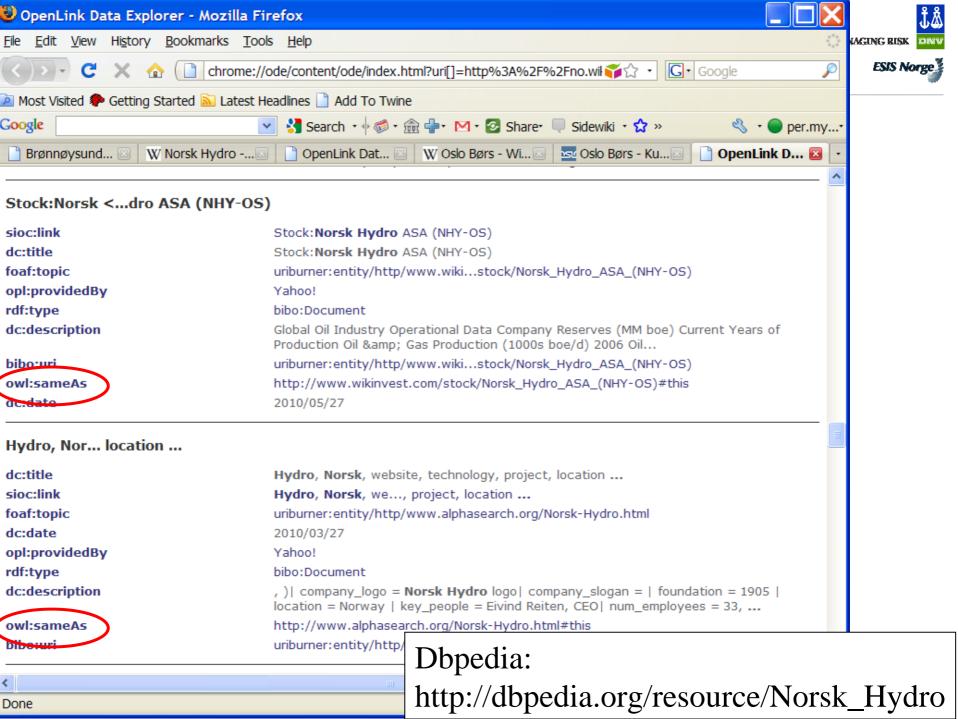
Nytt søk

Slide 8









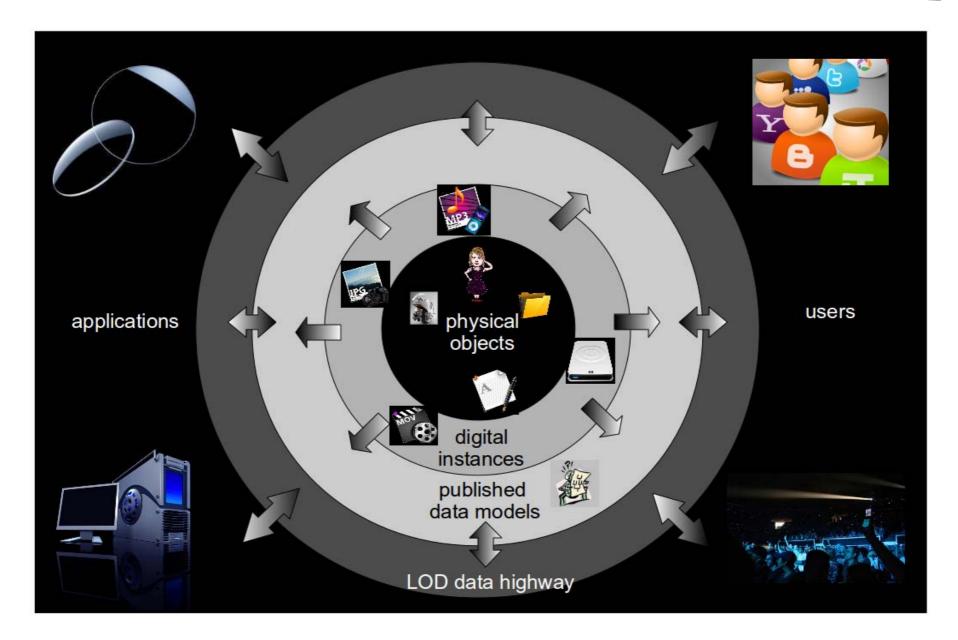
Target situation

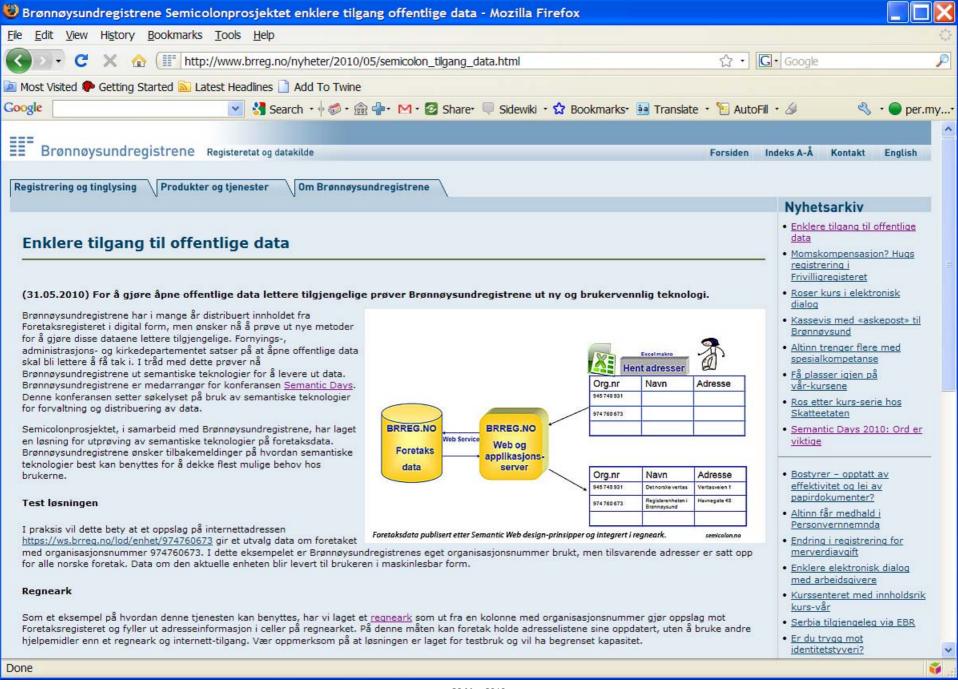




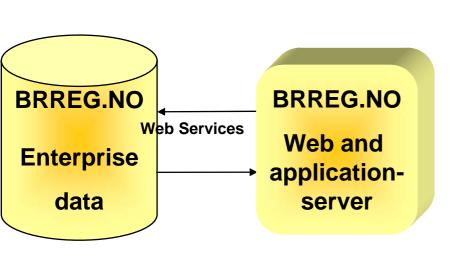
Open Information Architecture with UPI/URI

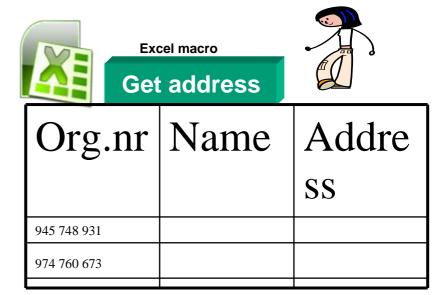






BRs Register of Business Enterprises as LOD and integrated into excel





Name	Addre
	SS
Det norske veritas	Veritasveien 1
Registerenheten i Brønnøysund	Havnegate 48
	Det norske veritas Registerenheten i

deliverable from www.semicolon.no, design and implementation by Per Myrseth, Martin Giese and Jørgen Stanger.myrseth@dnv.com, martingi@ifi.uio.no, jorgen.stang@dnv.com.

How the server responds



Identifier for an Norwegian enterprise:

https://ws.brreg.no/lod/enhet/914778214

No data behind this URI, based on how you ask, it redirects to either:

https://ws.brreg.no/lod/page/914778214

a html response

Or

https://ws.brreg.no/lod/data/914778214

a RDF response

What is requested ...



- Open up the national master data
- Greater efforts underway to get arid of information silos
 - Triggered by newly gained ease of information sharing
 - Triggered by increasing availability of valuable information on the net
- Getting aligned (inter-)nationally....
 - Cultural heritage (collections and information)
 - Private domains (health, finance, publishing)
 - Governmental data (statistics, information, interaction)
- through the use of open standards!
 - World Wide Web consortium: HTTP, XML, RDF, RDFS OWL
- Governments take up the challenge
 - Library of Congress → use of semantic standards to open up information
 - data.gov.uk → portal + entry to national data
 - Digitaal Erfgoed Nederland → "de Basis"



... and how to deliver that



- Help needed!
 - Competence
 - Insight
- Initiatives to ensure interoperability and avoid high "trail-and-error" costs
 - Builds upon experiences with many different sources
 - Minimal requirements in order to align future initiatives
 - source/organisation internal:
 - Objects must have URI
 - Relations between objects are required
 - Web service using information access standards like SPARQL for automated access
 - Between sources/organisations
 - Relations within a graph/model are common, but linking to other graphs is required!
 - Means → reuse schemes/ontologies and instance data with all your might!



Open Data, how to find the sources



- Several national initiatives
- Several domains have thousands of data sets
- Who should maintain the trustworthy list of sources

Examples of sources:

- Data.norge.no a list of open sources
- ORIS, a Semicolon/ IFI initiative, semi automatic updates a list of sources in a structured portal
- Semantic web search engines, indexes the rdf sources found.

Relevant legislative topics for mashup providers based on open data



- Protecting own rights
 - Establish clear rights to reuse and publish others open data
 - Is there a need for limiting others ability to reuse/ miss-use my data?
- Protecting the rights of 3rd parties
 - Is there a need for limiting others ability to reuse my sub-vendors data and/ or my own proprietary data?
- Limiting own legal responsibility
 - Based on SLA, data quality, IPR violations etc.

MANAGING RISK DNV ESIS Norge

Intellectual Property Rights for open data

Most jurisdictions recognize the following forms of IPR:

- Trade secret
- Trademark
- Patent
- Copyrights

For data openly accessible on the web it is no longer possible to claim trade secrets, trademark or patent protection of their data.

Internationally, quite uniform that one cannot claim copyright protection for individual entries of facts stored in a database.

1996, EU Introduced the database directive, to grant database creators a sui generis right towards unauthorized extraction of all or a substantial part of the database. Relevant if such extraction substantially harms the database creators investments.

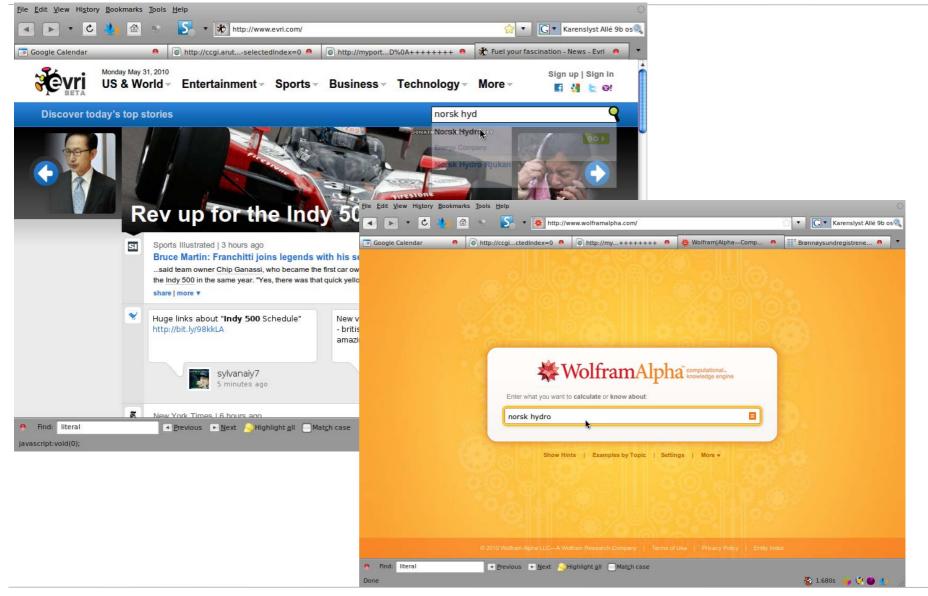
1996-2004 US: six bills suggested to implement a US-database directive. Non of them has pas into laws.

Some examples



MANAGING RISK DAV

Tools & Infrastructure that benefit from LOD



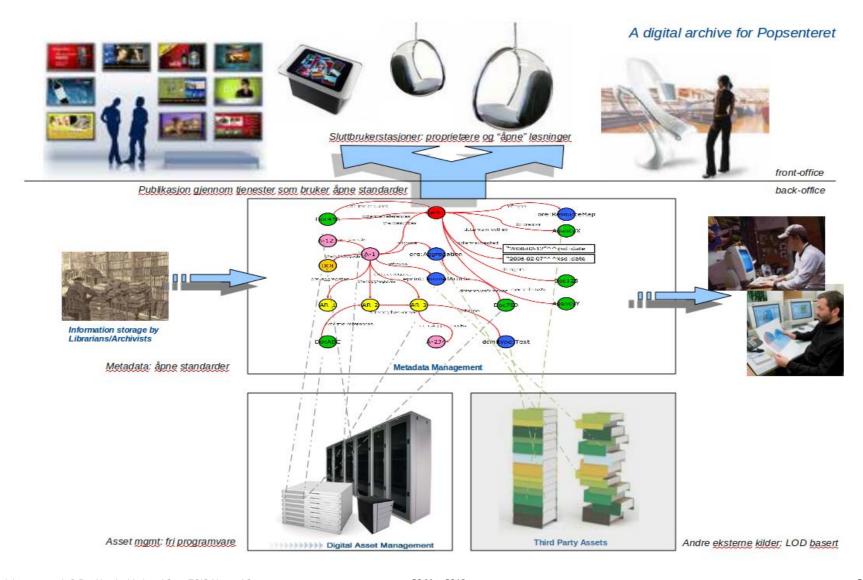
Example from Cultural Heritage sector





Example from Cultural Heritage sector

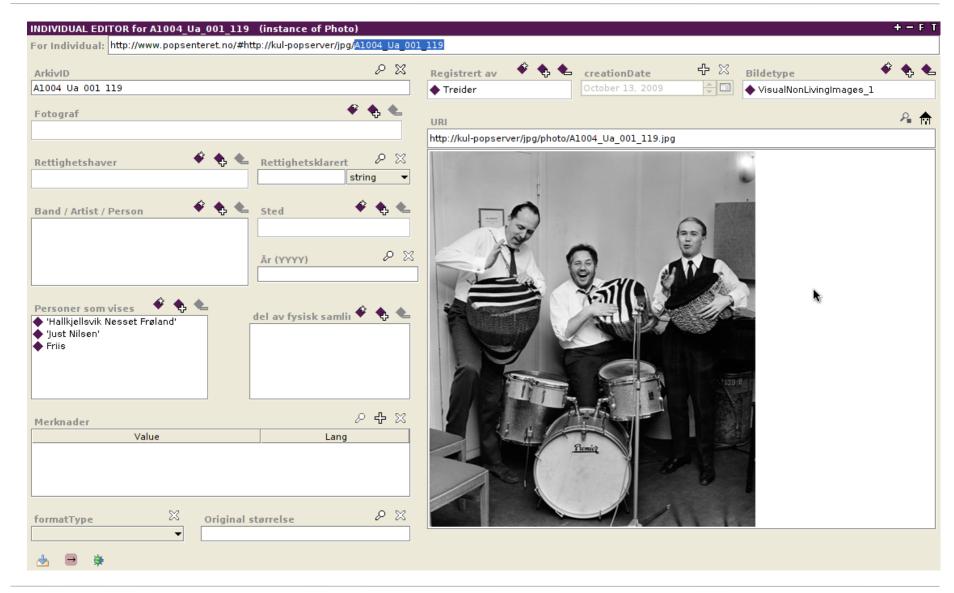




Closer look on a semantic solution



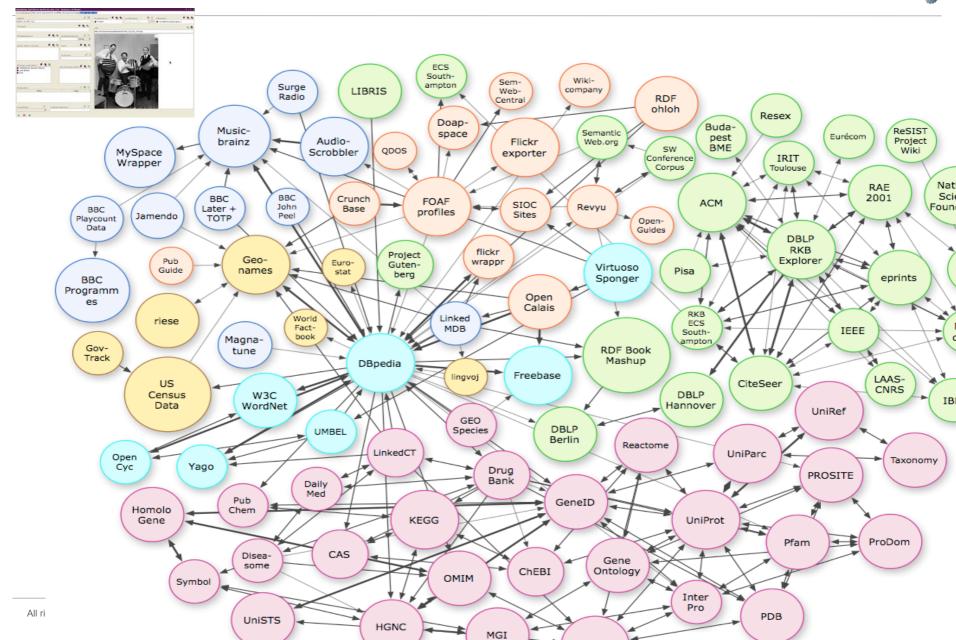




JÅ MANAGING RISK DNV

Why do we want this?





Summing up





Public sectors big hairy goals and the effect of using LOD



Public sector goal is to be **open**, **transparent**, **accountable**, **accessible**, **user-friendly**, **efficient**, **effective** and **service-oriented**.

The anticipated effect of Linked open data:

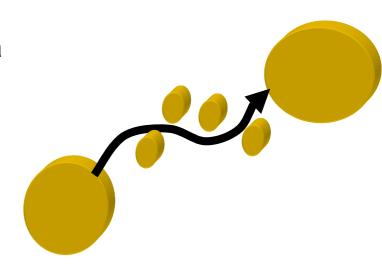
- Will open up PSI for innovation and makes public sector more transparent and accountable
- Utilize the internet as an infrastructure for access and scalability
- Enables more and cheaper user-friendly eGov services
 - Give the information back to the people
- Enables public sector to remove silos, lower integration cost and become more efficient and effective.
- Is a prerequisite in a service oriented public sector



The stairs to a nation's presence on the Semantic Web



- Open access to data
 - Licenses, costs/ fees, IPR issues handled
 - Infrastructure for access and reuse that scales
- Open access to **metadata**/ meaning of data
 - Publish
 - Reuse
- Follow LOD design principles
 - Identifiers, Relations and aut. query facilities
- Open access to provenance metadata
 - Describtion of the life cycle of the data
 - Quality attributes of the data
- A **trusted rating** of who holds what data (registry)





Thanks for your attention

Per Myrseth (Semicolon)
Robert H.P. Engels (Sesam4)



