

Internett

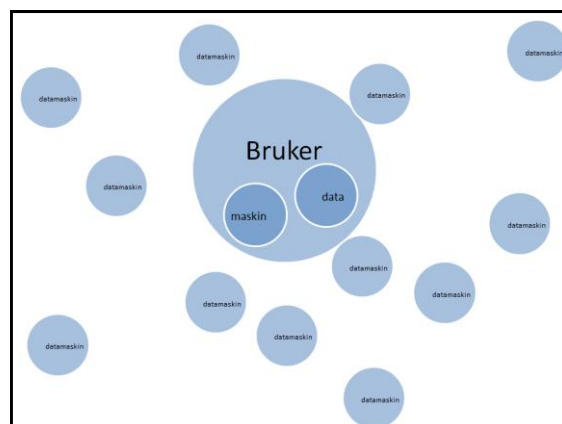
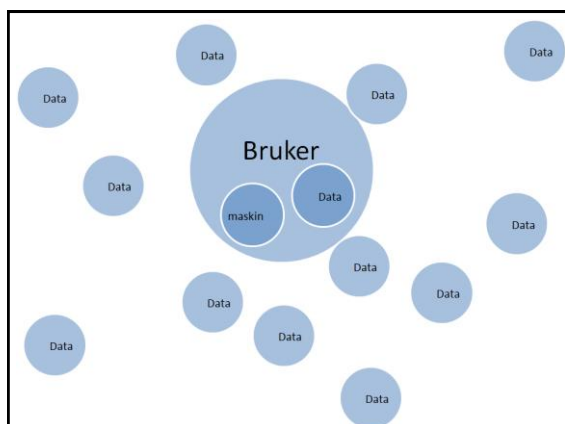
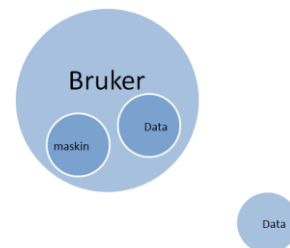
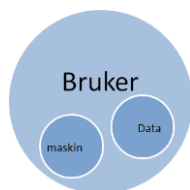
Fra tekniske standarder til samhandlingsarena

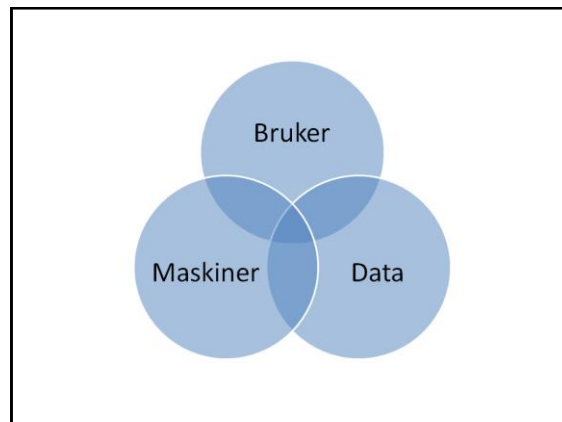
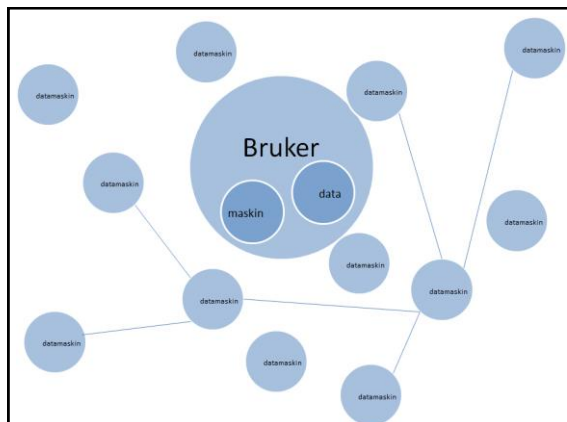
DRI 3010 - 23. september 2015

Steinar Skagemo

Pensum

1. Seres - <http://www.brreg.no/samordning/semantikk/>
→ 404 Not found
Isteden: <https://altinnett.brreg.no/no/SERES/>
2. Om Internett på Wikipedia:
<http://no.wikipedia.org/wiki/Internett>
3. Om XML: <http://www.w3.org/TR/REC-xml/>
4. Semantisk web
5. Om emnekart: <http://no.wikipedia.org/wiki/Emnekart>





Hva vet dere om internett?

[fyll inn]

Universalverktøy uten grenser

Internett er et verdensomspennende [datanettverk](#) basert på en standard kommunikasjonsprotokoll (TCP/IP).

Denne gjør det mulig for datamaskiner å kommunisere på tvers av en rekke forskjellige fysiske nett som f. eks. kabel, fiber, radiolink, satellitt og trådløst nett, og er grunnlaget for at datamaskiner kan kommunisere over hele verden. [Kilde: <http://no.wikipedia.org/wiki/Internett>]

stateless / tilstandsløs

Protokoller

[fyll inn]

Protokoll kan referere til

Et sett med regler som beskriver hvordan to eller flere parter skal oppføre seg i forhold til hverandre, som [diplomatiske protokoll](#) og [protokoller brukt i forbindelse med datamaskiner](#).

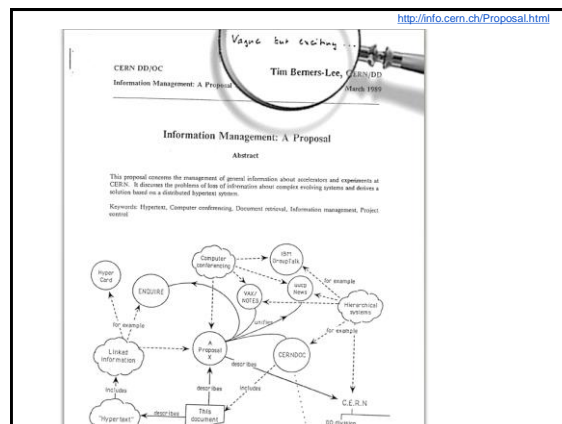
En gjerne noe høytidelig [nedtegnelse](#), som et [referat](#) (f.eks. møteprotokoll) eller annet (for eksempel kondolanseprotokoll)

[<http://no.wikipedia.org/wiki/Protokoll>]

World Wide Web

Hva er byggeklossene i World Wide Web?

1. HyperText Markup Language (HTML)
2. HyperText Transfer Protocol (http)
3. Uniform Resource Locator (URL)



<http://www.idg.no/computerworld/article245409.ece>

HTML

Universelt utformet siden 1992

Da ideen om World Wide Web ble lansert på starten av 90-tallet, var det mange som leiet med tanken på et system for informasjonsutveksling. Et slikt system ville vært alle tiders, det var alle enige om.

I motsetning til sine kolleger, som ville lage spesialbygde datamaskiner for å ta seg av denne informasjonsutvekslingen, ville [Sir Tim Berners-Lee](#) det annerledes. Han mente at en suksessfaktor vil være at hvem som helst, fra hvor som helst i verden, med hvilket som helst utstyr, burde kunne besøke en webside. Så han laget et universelt utformet språk ([HTML](#)) for å gjøre dette.

[Den første nettsiden Sir Tim Berners-Lee laget](#) – for over 20 år siden – er både responsiv og universelt utformet. Det er mer enn man kan si om mye annet som er laget de siste årene.

[<http://jallenkelt.net/2014/06/10/universell-utforming-starter-her/>]

<HEADER>	http://info.cern.ch/hypertext/WWW/TheProject.html
<TITLE>	The World Wide Web project</TITLE>
<NEXTID N="55">	
<HEADER>	World Wide Web
<BODY>	The WorldWideWeb (W3) is a wide-area hypertext information retrieval initiative universal access to a large universe of documents.
<H1>	World Wide Web</H1>The WorldWideWeb (W3) is a
NAME=0 HREF="whats.html">	Everything there is online about W3 is linked directly or indirectly to this document executive summary of the project. Mailing lists . Policy . November's W3 news . FAQ Questions
hypertext	information retrieval
initiative aiming to give universal	What's out there?
access to a large universe of documents.<P>	Pointers to the world's online information , subjects , WWW servers , etc.
Everything there is online about	Help on the browser you are using
W3 is linked directly or indirectly	Software Products
to this document, including an <A	A list of W3 project components and their current state. (e.g. Line Mode , X11 , NeXTStep , Servers , Tools , Mail robot , Library .)
NAME=24 HREF="Summary.html">executive	Technical
summary of the project. <A	Details of protocols, formats, program internals etc.
NAME=29 HREF="Administration/Mailing/Overview.html">M	Bibliography
ailing	Paper documentation on W3 and references.
Overview.html">M	People
ailing	A list of some people involved in the project.
Overview.html">M	History
Overview.html">M	A summary of the history of the project.
Overview.html">M	How can I help?
Overview.html">M	If you would like to support the web.
Overview.html">M	Getting code
Overview.html">M	Getting the code by anonymous FTP , etc.
Overview.html">M	
</A	

Live demo: <http://scratchpad.io/mundane-class-6934>

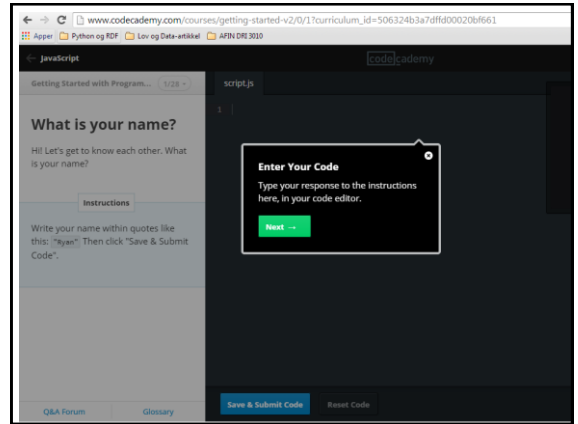
http

OPTIONS / HTTP/1.1
Accept: */*
Accept-Encoding: gzip, deflate
Connection: keep-alive
Content-Length: 0
Host: www.uio.no
User-Agent: HTTPie/0.8.0

HTTP/1.1 200 OK
Accept-Ranges: bytes
Age: 0
Allow: GET, HEAD, POST, OPTIONS, TRACE
Cache-Control: max-age=300
Connection: keep-alive
Content-Language: no
Content-Length: 0
Content-Type: httpd/unix-directory
Date: Wed, 24 Sep 2014 05:24:24 GMT
Server: Apache/2.2.25 (Unix)
Via: 1.1 varnish
X-Cache: MISS
X-Cacheable: NO,Not Cacheable
X-Varnish: 525831185

URL

<https://url.spec.whatwg.org/>



Mennesker

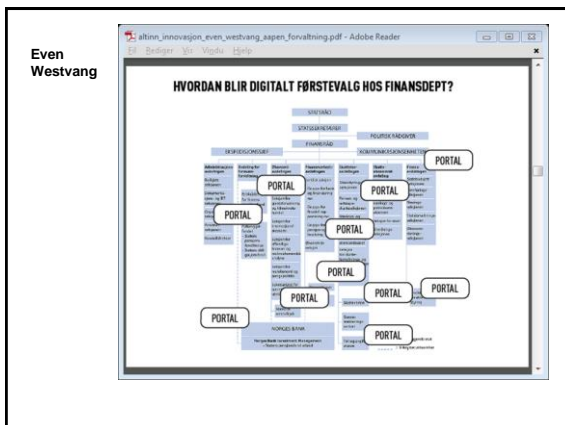
Maskiner

<http://bcove.me/f3564qn7>

Slik gjør vi digitalt til brukernes førstevalg
 Gründer og faglig leder Jostein Magnussen, Netlife Research

Har Tim Berners-Lee lyktes for godt?
 Er det for enkelt å lage nettsteder?
 Er vi kuer?
 Er Internett fullt?

Video fra Digitaliseringskonferansen i 2014
<http://frontend.smartcomtv.no/difi/2014/digitaliseringskonferansen/#>



Even Westvang

DET NOE VAGE KRAVET OM INTEROPERABILITET VI ALLEREDE HAR:

difi

MÅ BLI TIL EN PÅKREVDT PRAKSIS SOM INNEBÆRER F. EKS:

[Systems] architected for openness and expose high-value data and content as web APIs

Hvordan skal maskiner forstå oss?

Lære maskiner å forstå "naturlig språk" - og ta hensyn til oss
Lære oss å ta hensyn til maskinene - slik at vi blir forstått

"Jordan" og "Jordan"

<http://www.apollon.uio.no/artikler/2008/jordan.html>

English

FORSKNINGSMAGASINET

APOLLON

Forsiden Artikler Portretter Bolanmeldelser Ledere Video Om Apollon Tema Fagområder

2014
2013
2012
2011
2010
2009
2008
2007
2006

Ser forskjell på Jordan og Jordan

Åsne Haaland har utviklet et dataprogram som ser forskjell på egennavn i tekster. Så nå er det mulig å skjelve elven Jordan fra tannbørstemerket Jordan.

Terik om søkemotoren Google kunne skille mellom elven Jordan, nakemodellen Jordan og tannbørstemerket Jordan? En ny doktrahandling ved Universitetet i Oslo viser at dette er mulig. Ved Institutt for lingvistik og nordiske studier har Åsne Haaland utviklet et program som gjør det mulig for en datamaskin å skille forskjellige typer egennavn i en tekst, uavhengig av om egennavnet er en person, et sted, en organisasjon eller noe annet.

– Det er utfordring for maskinlæringsteknikere



"Things, not strings"

<http://googleblog.blogspot.no/2012/05/introducing-knowledge-graph-things-not.html>



Introducing the Knowledge Graph: things, not strings

Posted: Wednesday, May 16, 2012

841 8417 1 Tweet 1.157 1 Like 1.5

Cross-posted on the Inside Search Blog

Search is a lot about discovery—the basic human need to learn and broaden your horizons. But searching still requires a lot of hard work by you, the user. So today I'm really excited to launch the Knowledge Graph, which will help you discover new information quickly and easily.

Take a query like [aj mahal]. For more than four decades, search has essentially been about matching keywords to queries. To a search engine the words [aj mahal] have been just that—two words.

But we all know that [aj mahal] has a much richer meaning. You might think of one of the world's most beautiful monuments, or a Grammy Award-winning musician, or possibly even a casino in Atlantic City, NJ. Or, depending on when you last ate, the nearest Indian restaurant. It's why we've been working on an intelligent model—in geek-speak, a "graph"—that understands real-world entities and their relationships to one another. Things, not strings.

Schema.org

http://semanticweb.com/schema-org-fires-lit_b44380

Schema.Org: The Fire's Been Lit

By Jennifer Davis on September 15, 2014 3:00 PM



Why has schema.org made the following strides since its debut in 2011?

- In a sample of over 12 billion web pages, 21 percent, or 2.5 billion pages, use it to mark up HTML pages, to the tune of more than 15 billion entities and more than 60 billion triples; in that same sample, this works out to six entities and 26 facts per page with schema.org.
- Just about every major site in every major category, from news to e-commerce (with the exception of Amazon.com), uses it.
- Its ontology counts some 600 properties and 600 classes.

Eksempel: Enhetsregisteret Et format for data: http + JSON

```
C:\Python27>http --verbose GET http://hotel1.difi.no/api/json/brreg/enhetsregistret?query=974778610 HTTP/1.1
GET /api/json/brreg/enhetsregisteret?query=974778610 HTTP/1.1
Accept: */*
Accept-Encoding: gzip, deflate
Connection: keep-alive
Host: hotel1.difi.no
User-Agent: httpie/0.9.0

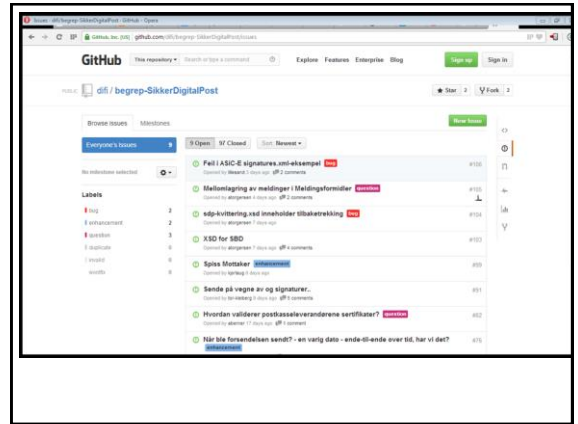
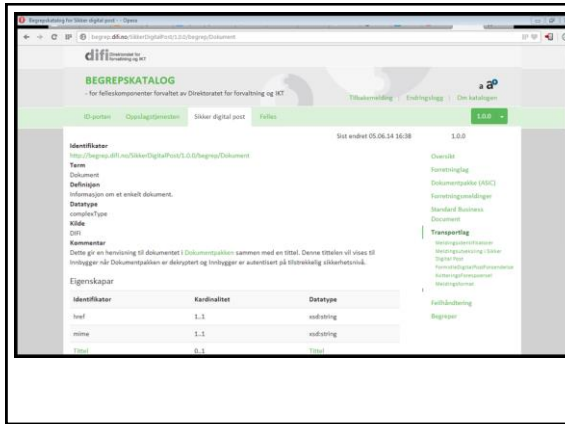
HTTP/1.1 200 OK
Access-Control-Allow-Origin: *
Connection: keep-alive
Content-Type: application/json; charset=utf-8
Date: Wed, 11 Sep 2014 09:36:48 GMT
ETag: 11313057
Server: Apache/2.4.18 (Ubuntu)
Transfer-Encoding: chunked
X-Datahotel-Page: 1
X-Datahotel-Total-Pages: 1
X-Datahotel-Total-Posts: 1

{
  "entries": [
    {
      "ansatte_antal": 148,
      "ansatte_dist": 10,
      "avvikling": "1000000000",
      "forradskommavn": "1234567890",
      "forradskommnr": "123456",
      "forradslaud": "1000000000",
      "forradspostnr": "123456",
      "forradspoststed": "1234567890",
      "forretningsadr": "1234567890",
      "foretaksnr": "1234567890",
      "konkur": "1234567890",
      "navn": "1234567890"
    }
  ]
}
```

Fra dokumentformat til dataformat

HTMLs suksess inspirerte til XML
XML er blitt brukt til å utveksle strukturerte data
Men er det så vellykket?
Idag: De som utvikler foretrekker ofte JSON

Vi har protokollen for å overføre, URL-er for å referere til/identifisere, og stadig bedre format for å "pakke inn" dataene som skal utveksles.
Grunnlag for samhandling



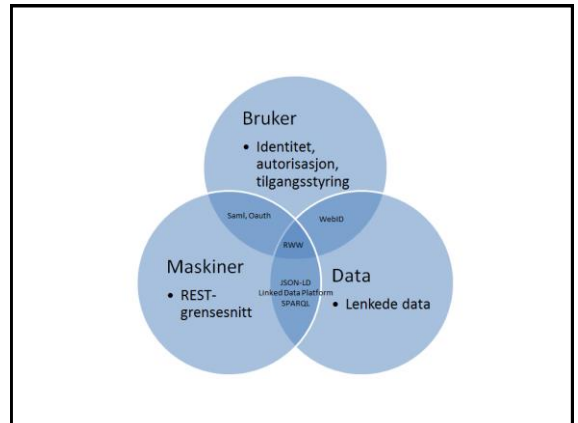
<http://schema.org/Person>

Seres - hva bør det være?

Thing > Person
A person (alive, dead, undead, or fictional).

Property	Expected Type	Description
Properties from Person		
additionalName	Text	An additional name for a Person, can be used for a middle name.
address	PostalAddress	Physical address of the item.
affiliation	Organization	An organization that this person is affiliated with. For example, a school/university, a club, or a company.
alumniOf	EducationalOrganization	An educational organization that the person is an alumni of.
award	Text	An award won by this person or for this creative work. Supercedes awards .
birthdate	Date	Date of birth.
brand	Brand or Organization	The brand(s) associated with a product or service, or the brand(s) maintained by an organization or business person.
children	Person	A child of the person.
colleague	Person	A colleague of the person. Supercedes colleagues .
contactPoint	ContactPoint	A contact point for a person or organization. Supercedes contactpoints .
deathDate	Date	Date of death.
done	Text	The Dun & Bradstreet DUNS number for identifying an organization or business person.

Se notatens fra forelesningen i 2012 for mer informasjon om Seres og formålet.



http://semanticweb.com/semantic-webs-rocking-aint-stopping-now_b44478

2014 SEMTECH@ WEST

The Semantic Web's Rocking, And There Ain't No Stopping It Now

By Jennifer Zizzo on September 23, 2014 @ 04:48

Make no mistake about it. The semantic web has been a success and that's not about to stop now. That was essentially the message delivered by W3C Data Activity Lead Phil Archer, during his keynote address celebrating the semantic web's ten years of achievement at last month's Semantic Technology & Business Conference in San Jose.

After acknowledging that he's heard it all about the semantic web being a failure, about it being rebranded as Linked Data and that being a failure too, he summed up those impressions in one distinctly British word: "Bollocks." The list of successes ranged across the spectrum, from the use in federated data portals of the Data Catalogue Vocabulary from the W3C's Government Linked Data Working Group, to the 47-million triples strong *Open Phacts* pharmacology discovery.

https://www.youtube.com/watch?v=WFR3Om_xhE

IBM's Watson Supercomputer Destroys Humans in Jeopardy! Engadget

Engadget 1:35:10