

# Metadata og nye roller

Forskerforbundet

Rica Nidelven, 8/10 2014

// VI GJØR  
KUNNSKAP  
TILGJENGELIG //

# Agenda

## // Hvilke data finnes?

- Hva finnes av tilgjengelige metadata og på hvilke format?
- Lisens

## // Hva er bibliotekarenes rolle?

## // Hvordan skaffe seg nødvendig kompetanse?

- MOOC
- NMC Horizon

# Hvem er jeg?



Asbjørn Risan  
Produkteier, Oria  
[Produkteiergruppen](#)

- Masterutdannet bibliotekar ved Høgskolen i Oslo med spesialisering innen kunnskapsorganisasjon og gjenfinning
- Produkteier i BIBSYS med ansvar for Søk og gjenfinning og søkeverktøyet Oria (oria.no)
- Involvert i prosjektet Nytt biblioteksystem



# Tilgjengelige datasett

// BIBSYS Autoritetsregister

// BIBSYS Bibliotekbase

// BIBSYS Brage

// Kommer

– BIBSYS DLR

– BIBSYS BIRD



# BIBSYS Autoritetsregister

- // Består av ca 1.8 millioner personautoriteter
- // Danner utgangspunkt for det Nasjonale autoritetsregisteret som er en delmengde av autoritetsregisteret (ca 73.000)
- // Nasjonalbiblioteket bruker BIBSYS autoritetsregister i sitt arbeid og vedlikeholder det nasjonale autoritetsregisteret ved at de oppgraderer autoriteter i sitt bibliografiske arbeid



# BIBSYS Autoritetsregister

- // I tillegg til personautoriteter inneholder registeret autoriteter for:
  - Konferanser, korporasjoner og organisasjoner (ca 340.000)
  - Standardtitler (ca 25.500)
  - Verk (under utarbeidelse/planlegging)

// Nasjonalt autoritetsregister blir eksportert til VIAF: The Virtual International Authority File

// Inneholder koblinger til dokumenter (bøker) hvor autoriteten har blitt brukt

// Tilgjengelig for søk maskin til maskin (SRU, OAI, webservice) (spørsmål inn, XML tilbake)

**BIBSYS** – Format: Dublin Core eller MARC



# Autoritetsregisteret som Linked data

- // Prosjekt (Rådata nå!) mellom NTNU Universitetsbiblioteket og BIBSYS med økonomisk støtte fra ABM-utvikling (Arkiv, Bibliotek og Museum)
- // Mål: Lage linked data av BIBSYS Autoritetsregister
- // 1,5 millioner personautoriteter representert i RDF med ca. 9 millioner tripler (utsagn)
- // Tilgjengelig som nedlastbar fil eller via et SPARQL-søk-grensesnitt ([http://www.bibsys.no/files/out/linked\\_data/autreg/](http://www.bibsys.no/files/out/linked_data/autreg/))
- // Lisensiert under en åpen lisens (PDDL)  
“Public Domain for Data/Databases Licence”



# Autoritetsregisteret som Linked data

// Rådata nå! ble kåret til årets bibliotekprosjekt innen fagbibliotek 2011 ([http://www.bokogbibliotek.no/images/stories/pdf\\_2012/bob-nr1-2012-web-s20-26.pdf](http://www.bokogbibliotek.no/images/stories/pdf_2012/bob-nr1-2012-web-s20-26.pdf))

// Oppdatert med autoriteter pr. 2011

// Inneholder koblinger til andre datasett:

- Dbpedia (dataversjonen av Wikipedia)
- Deutsche National Bibliothek
- VIAF: The Virtual International Authority File





# Eksempel – Helge Ingstad

```
<rdf:RDF>
  <rdf:Description rdf:about="http://data.bibsys.no/data/notrbib/authorityentry/x90106700">
    <dc:identifier>x90106700</dc:identifier>
    <foaf:name>Helge Ingstad</foaf:name>
    <whois:since rdf:datatype="http://www.w3.org/2001/XMLSchema#gYear">1899</whois:since>
    <whois:until rdf:datatype="http://www.w3.org/2001/XMLSchema#gYear">2001</whois:until>
    <radatana:catalogueName>Ingstad, Helge</radatana:catalogueName>
    <rdfs:label>Helge Ingstad</rdfs:label>
    <owl:sameAs rdf:resource="http://d-nb.info/gnd/124939805"/>
    <owl:sameAs rdf:resource="http://viaf.org/viaf/34580367"/>
    <owl:sameAs rdf:resource="http://dbpedia.org/resource/Helge_Ingstad"/>
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
    <dc:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2010-10-18</dc:modified>
  </rdf:Description>
</rdf:RDF>
```

# Eksempel – Helge Ingstad

```
<rdf:RDF>
  <rdf:Description rdf:about="http://data.bibsys.no/data/notrbib/authorityentry/x90106700">
    <dc:identifier>x90106700</dc:identifier>
    <foaf:name>Helge Ingstad</foaf:name>
    <whois:since rdf:datatype="http://www.w3.org/2001/XMLSchema#gYear">1899</whois:since>
    <whois:until rdf:datatype="http://www.w3.org/2001/XMLSchema#gYear">2001</whois:until>
    <radatana:catalogueName>Ingstad, Helge</radatana:catalogueName>
    <rdfs:label>Helge Ingstad</rdfs:label>
    <owl:sameAs rdf:resource="http://d-nb.info/gnd/124939805"/>
    <owl:sameAs rdf:resource="http://viaf.org/viaf/34580367"/>
    <owl:sameAs rdf:resource="http://dbpedia.org/resource/Helge_Ingstad"/>
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
    <dc:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2010-10-18</dc:modified>
  </rdf:Description>
</rdf:RDF>
```

Navn

Fødsels-/dødsår

Koblinger

# Eksempel – Henrik Ibsen

```
<rdf:RDF>
  <rdf:Description rdf:about="http://data.bibsys.no/data/notrbib/authorityentry/x90061718">
    <dc:identifier>x90061718</dc:identifier>
    <foaf:name>Henrik Ibsen</foaf:name>
    <whois:since rdf:datatype="http://www.w3.org/2001/XMLSchema#gYear">1828</whois:since>
    <whois:until rdf:datatype="http://www.w3.org/2001/XMLSchema#gYear">1906</whois:until>
    <radatana:catalogueName>Ibsen, Henrik</radatana:catalogueName>
    <rdfs:label>Henrik Ibsen</rdfs:label>
    <owl:sameAs rdf:resource="http://d-nb.info/gnd/118555286"/>
    <owl:sameAs rdf:resource="http://viaf.org/viaf/71378383"/>
    <owl:sameAs rdf:resource="http://dbpedia.org/resource/Henrik_Ibsen"/>
    <skos:altLabel>Henrikku Ipusen</skos:altLabel>
    <skos:altLabel>Henrik Johan Ibsen</skos:altLabel>
    <skos:altLabel>Yibusheng</skos:altLabel>
    [...]
    <skos:altLabel>I-pu-sheng</skos:altLabel>
    <skos:altLabel>Brynjolf Bjarme</skos:altLabel>
    <skos:altLabel>Henryk Ibsen</skos:altLabel>
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
    <dc:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2009-11-27</dc:modified>
  </rdf:Description>
</rdf:RDF>
```

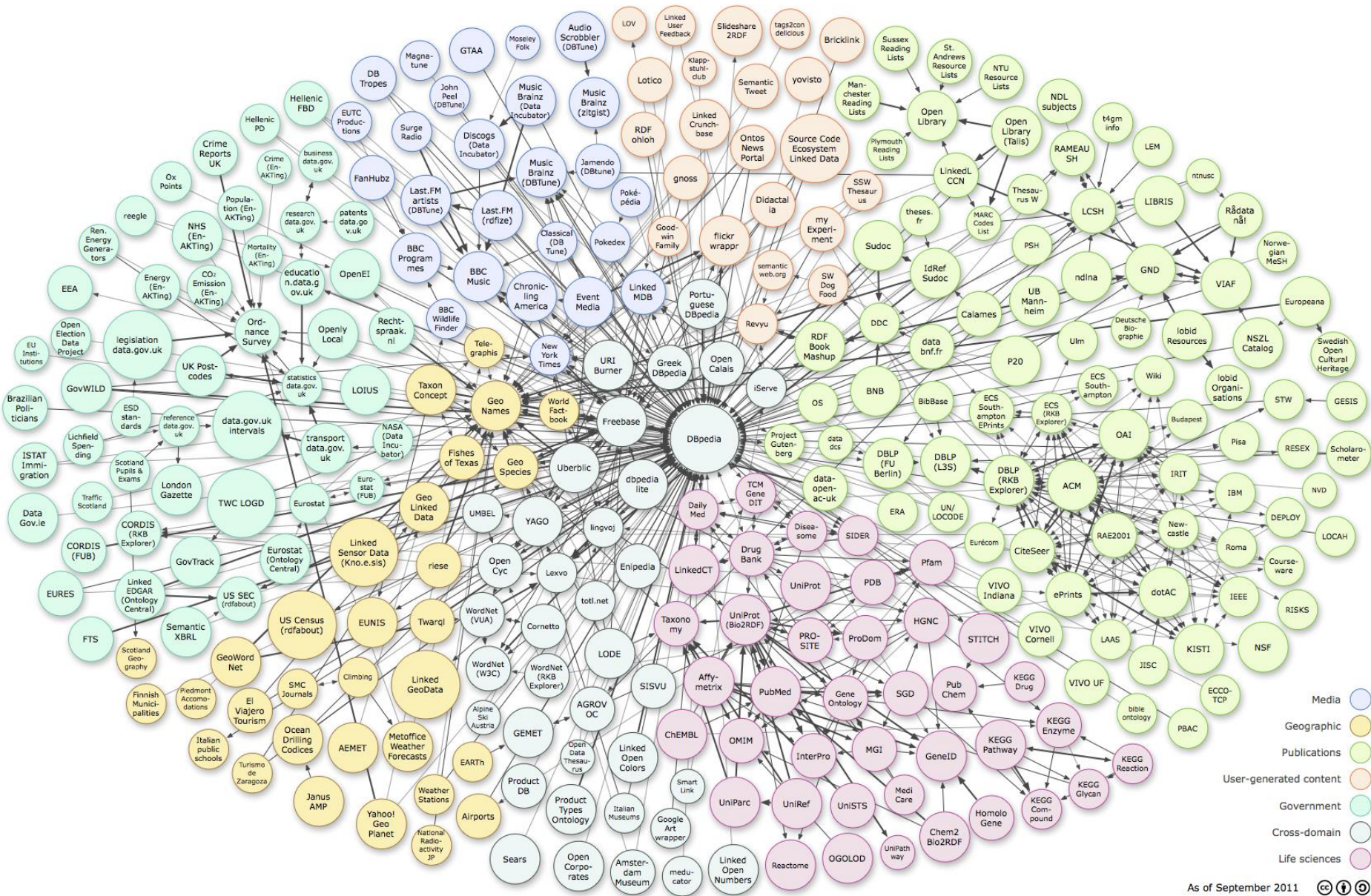
**Navn**

**Fødsels-/ dødsår**

**Koblinger**

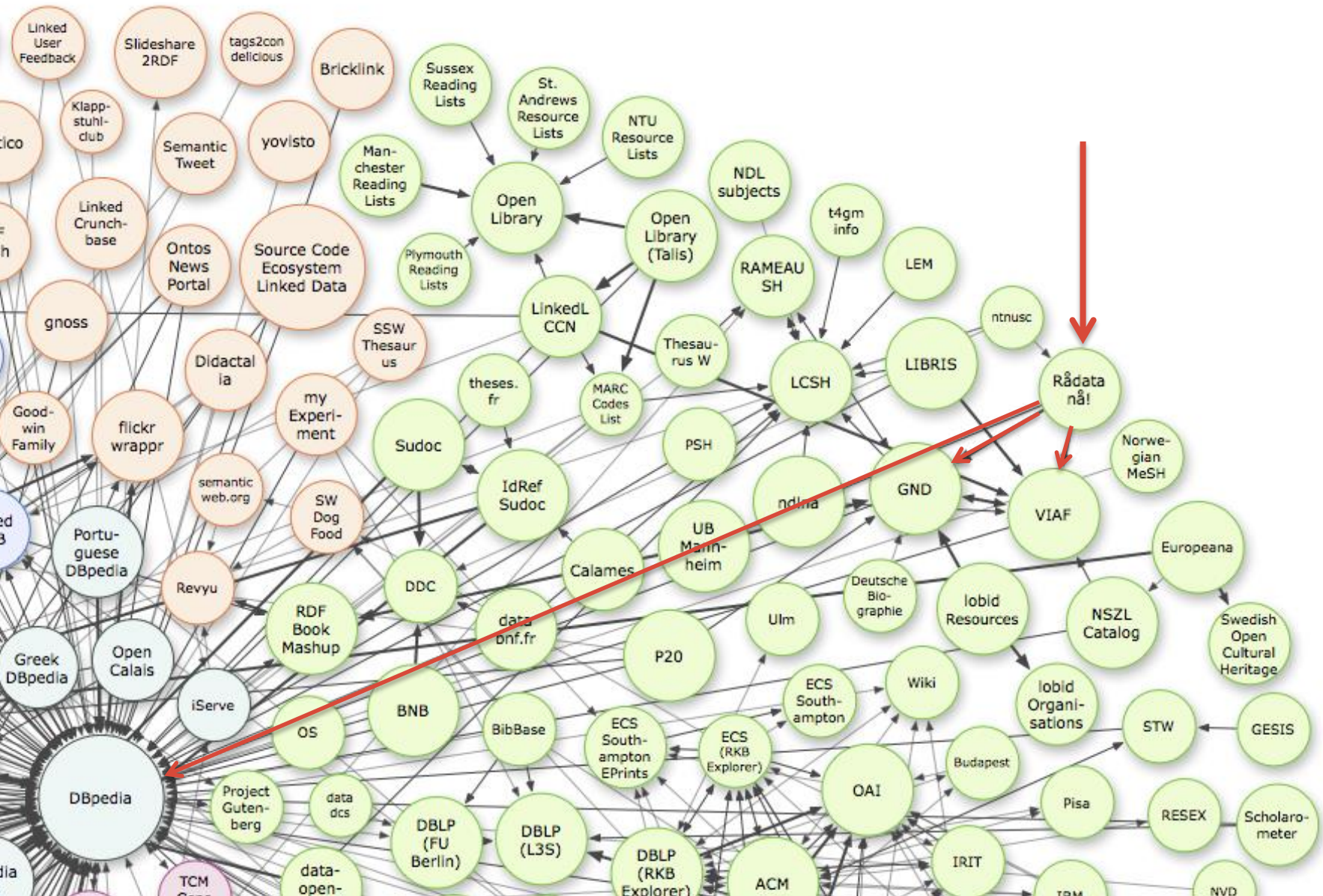
**Alternative navneformer**





As of September 2011   





# BIBSYS Bibliotekbase

- // Basert på den samlede beholdningen til norske fag- og forskningsbibliotek
  - Inneholder mer enn 6,9 millioner bibliografiske beskrivelser av dokumenter
  - Inneholder informasjon om mer enn 23 millioner eksemplarer med plassering og tilgjengelighetsinformasjon
- // Fritt tilgjengelig via søk maskin til maskin (SRU, OAI)
  - Format: Dublin Core eller MARC
- // Gjør det mulig å hente ut lister o.l. på egenhånd
- // Inneholder koblinger til autoritetsregisteret



# BIBSYS Bibliotekbase

- // Søknad inne hos Nasjonalbiblioteket om å konvertere og gjøre om Bibliotekbasen til RDF
- // Ønske om pilot i løpet av 2015
- // Inneholder koblinger til autoritetsregisteret
- // Fritt tilgjengelig for alle
- // Vil se på initiativ som BIBFRAME, Deichman, Libris etc og koble mot disse



# Andre registre

- // BIBSYS Brage – institusjonelle arkiv – mulig å høste via OAI-PMH
- // BIBSYS DLR – Læringsobjekter – vil bli mulig å høste
- // BIBSYS BIRD – Forskningsdata (prosjekt med BI) – deler vil bli mulig å høste
- // Nye emneregistre inn i autoritetsregisteret – under utredning





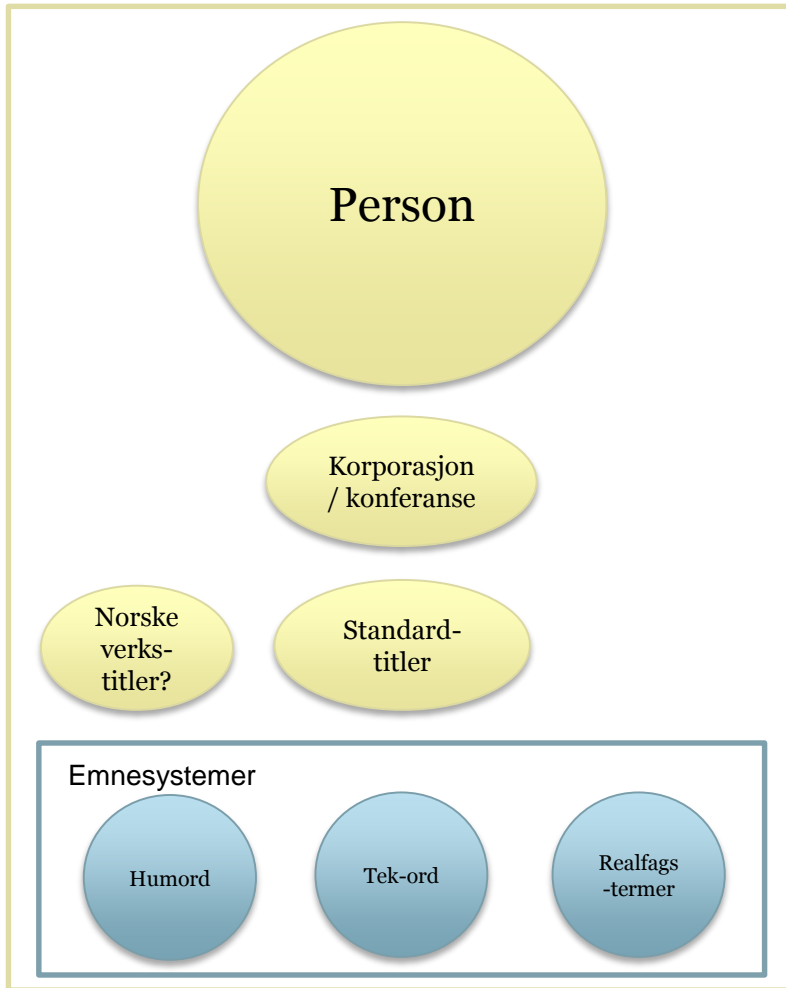
# Nå-situasjonen

**BIBSYS**  
Autoritets-  
register for  
person



# Fremtidig situasjon ?

## BIBSYS Autoritetsregister



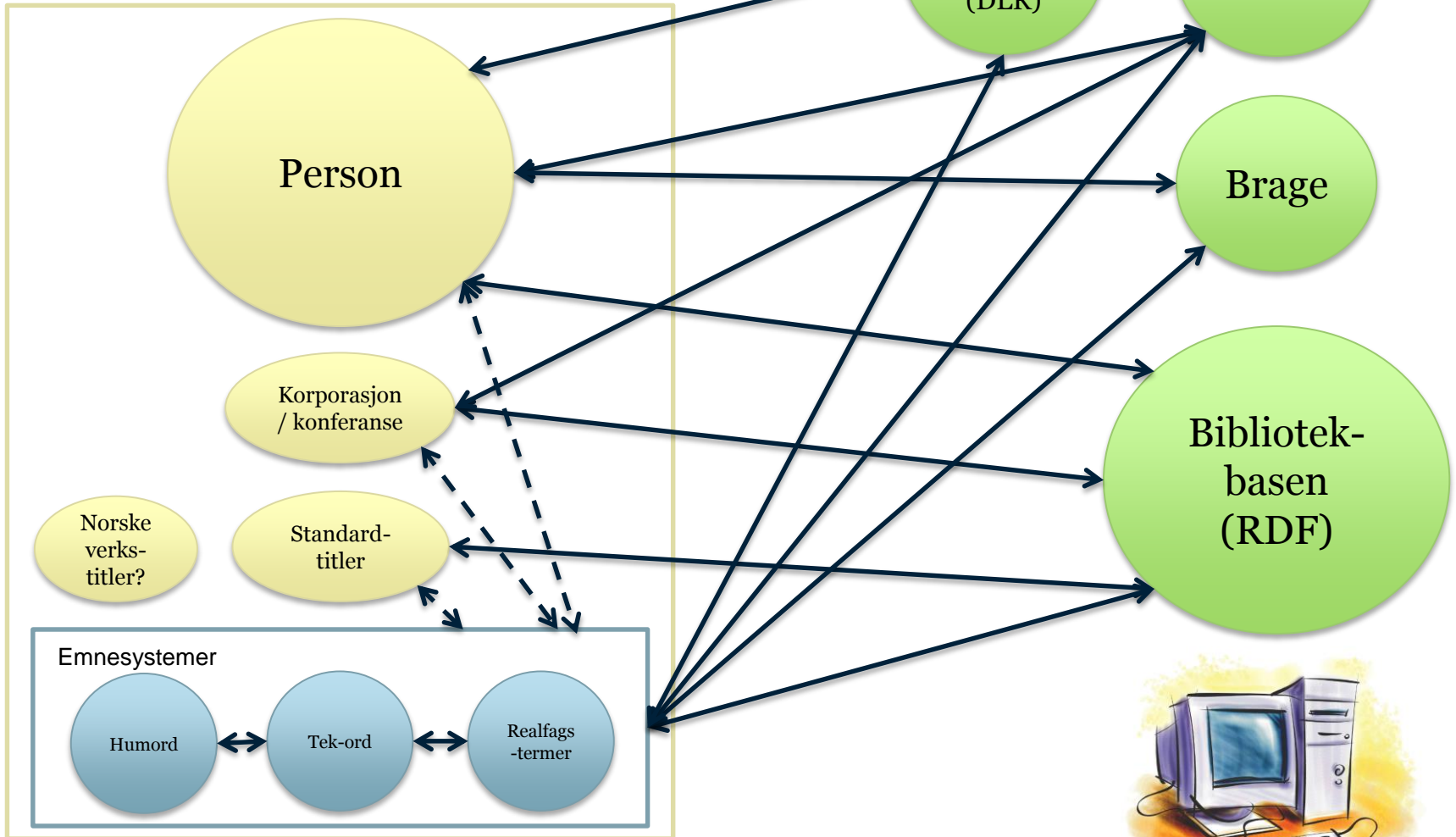
ID'er basert på handle



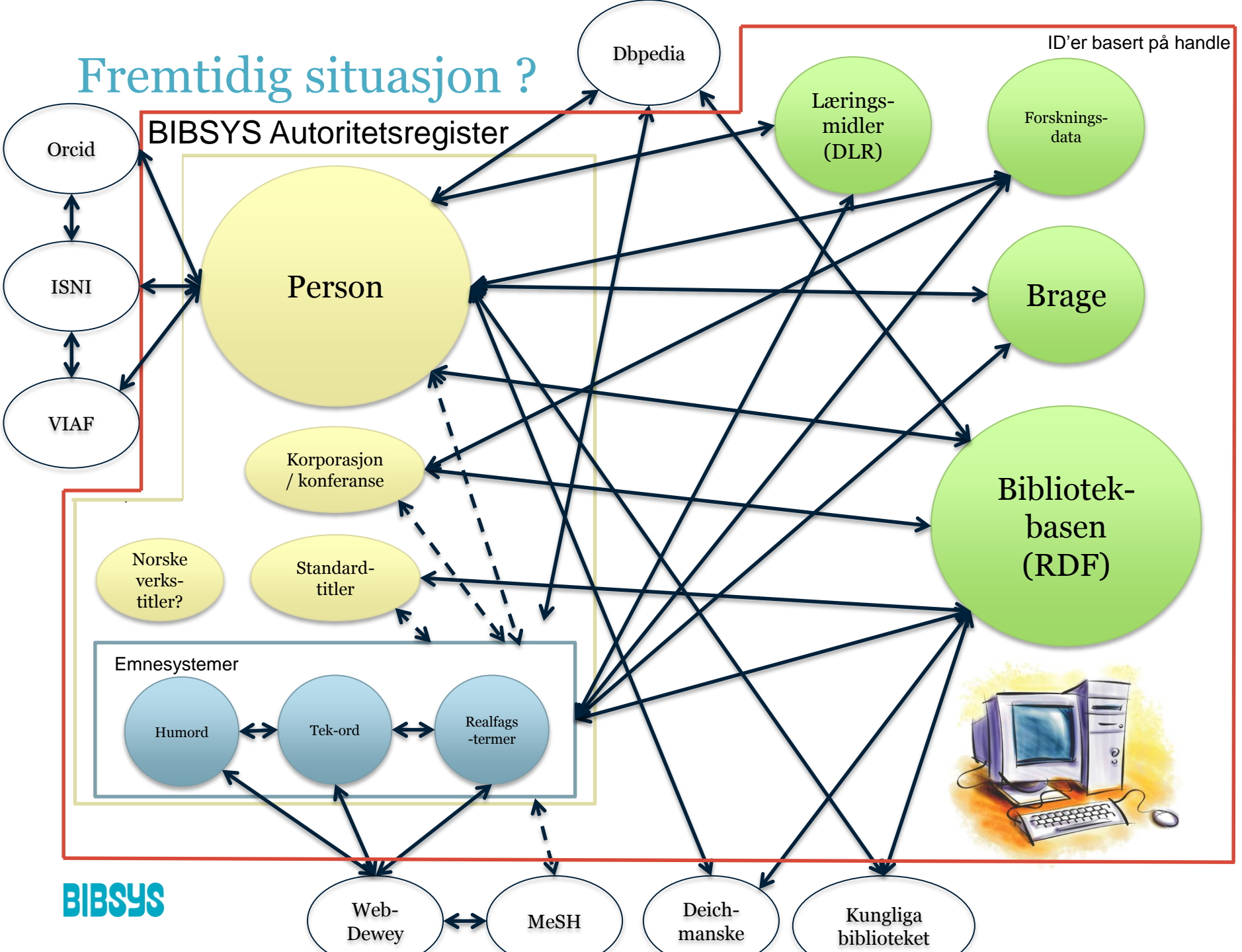
# Fremtidig situasjon ?

ID'er basert på handle

## BIBSYS Autoritetsregister



# Fremtidig situasjon ?



# Lisens



Cielo

Cielo

Cielo



Værsågod,

- bare forsyn deg!



Værsågod,

- bare forsyn deg,

men kun en om gangen!





# Lisens

- // En lisens er en eksplisitt forhåndsdefinert tillatelse som definerer rettigheter og forutsetninger for bruk av datasett/åndsverk
- // Klargjør hvorvidt et datasett kan brukes, i hvilke sammenhenger og under hvilke forutsetninger
- // Besparende for bedriften som har en definert politikk
- // Besparende for brukerne som vet hva de må forholde seg til



# Lisenser i BIBSYS

// Autoritetsregisteret som Linked Data er lisensiert som PDDL (Public Domain Database Licence)

- Helt åpen for bruk uten krav eller begrensninger

// BIBSYS Autoritetsregister og BIBSYS Bibliotekbase er lisensiert med NLOD (Norsk lisens for offentlige data)

## Du har lov til:

- å kopiere og tilgjengeliggjøre
- å endre og/eller sette sammen med andre datasett
- å kopiere og tilgjengeliggjøre en endret eller sammensatt versjon
- å benytte datasettet kommersiell

## På følgende vilkår:

- at du navngir lisensgiver slik lisensgiver ber om, men ikke på en måte som indikerer at disse har godkjent eller anbefaler deg eller din bruk av datasettet
- at du ikke bruker dataene på en måte som fremstår som villedende, og heller ikke fordreier eller uriktig fremstiller dataene

# Hva er bibliotekenes rolle?

## // Tradisjonelt sett:

- Sikre gode og korrekte metadata
- Bygge opp autoritetsregistre
- Sikre søkeinnganger (kontrollerte emneord, klassifikasjon, henvisninger)

Bibliotekarer har lagt til rette for den semantiske veven siden starten!



# Hva er bibliotekenes rolle?

## // Fremtidige roller:

- Metadataeksperter
- Bestiller av tjenester
- Aktiv del av læringsprosessen og produksjon og tilgjengeliggjøring av læringsobjekter
- Aktiv del av forskningsprosessen
- Jobbe for åpen publisering og tilgang



# Hva er bibliotekenes rolle?

## // Fremtidig rolle: Metadataeksperter

- Det er fortsatt behov for gode metadata og gode innganger til materialet
- Det vil være nye metadataformat
  - MLR ("Information technology - Learning, education and training - Metadata for learning resources") også kjent som "ISO/IEC 19788"
  - MODS (Metadata Object Description Schema)
- Kobling mellom ulike metadataformat
- Andre typer materiale som skal beskrives
- Andre nivåer/granularitet på det som beskrives
- Økt bruk av systemgenererte data – sikre at nødvendige data blir registrert og brukt
- Økt verdi av kontrollerte data/innganger – semantisk web

# Hva er bibliotekenes rolle?

## // Fremtidig rolle: Bestiller av tjenester

- Kjenne brukerne
- Kjenne til teknologi og utviklingstrender
- Definere behov og ønsket funksjonalitet
- Stille krav til funksjoner og tjenester/kravspesifisere
  - Modellere virkeligheten: Hvem gjør hva, i hvilken rekkefølge, i hvilken hensikt
  - En før og etter situasjon, er ofte tilstrekkelig kravspesifisering. Fokus på hvem som handler og hvilken informasjon som blir utvekslet.
- Kanskje den viktigste rollen
- Har dere kompetanse på programmering og teknologi er det fint, men ikke nødvendig...

# Hva er bibliotekenes rolle?

## // Fremtidig rolle: Aktiv del av forskningsprosessen

- Registrering, publisering, tilgjengeliggjøring og formidling av forskning
- Dataforvaltning og kurator/fasilitator
- Bistå med dataforvaltningsplaner



# Hva er bibliotekenes rolle?

## // Fremtidig rolle: Aktiv del av forskningsprosessen

- Biblioteket må bidra med å strukturere og langtidspreservere digitale data, samt kunne bistå forskeren med innsamling og analyse av dataene
- Alle bibliotek MÅ ha minst en bibliotekar som kan bistå forskere med å utarbeide "Dataforvaltningsplan" (Data Management Plan)
  - uten å kunne bistå i starten av forskningsprosessen, der forskeren prøver å finne budsjett, så blir det vanskelig å få bli med på resten av forskningsprosessen.
- Alle bibliotek MÅ forstå hvordan forskning blir utført på egen institusjon - uten denne kunnskapen er biblioteket koblet av resten av både forsknings- og utdanningsprosesser.





# Hva er bibliotekenes rolle?

- // Fremtidig rolle: Aktiv del av produksjon og tilgjengeliggjøring av læringsobjekter (quiz, video, presentasjon, tekst, lyd m.m.)
  - Kjennskap til læringsprosessene, pedagogikk og kommunikasjon
    - Hvordan skal læringsobjektet brukes, i hvilken kontekst av hvilken målgruppe
  - Kjennskap til redigering av digitale objekter, for eksempel video/lyd, er ofte nødvendig
  - Copyright-avklaringer
  - Metadata
  - Når noen andre har levert innholdet, så må dette struktureres, dokumenteres, pakkes og presenteres - ofte i flere versjoner - for ulike målgrupper

# Hva er bibliotekenes rolle?

- // Fremtidig rolle: Jobbe for åpen publisering og tilgang
  - Lisenser – hvem kan bruke under hvilke forutsetninger
  - Åpne format – ikke-proprietære format (EPUB fremfor pdf, odf fremfor doc, html5)
  - Tilgjengelig på alle enheter – unngå låsing til en bestemt type enhet, plattform (iOS, Android, Kindle)
  - Copyright – avklare rettigheter og muligheter for bruk. F.eks. i utvikling av MOOCs



# Hva er bibliotekenes rolle?

- // Fremtidig rolle: Jobbe for åpen publisering og tilgang
  - Biblioteket må kjenne til hva som gjør en EPUB fil bedre enn en annen EPUB fil - med tanke på bruk av åpne formater som kan brukes på alle enheter - uten denne kunnskapen vil biblioteket vanskelig kunne kravstille innkjøp og forvalte budsjettet på en god måte.
  - Biblioteket må ha kompetanse til å identifisere når det dukker opp et nytt og bedre alternativ til EPUB.



# Kilder til kunnskap

## // MOOC om semantisk teknologi

- Online fritt tilgjengelig kurs som dekker web- og semantiske teknologier
- Arrangert juli-august-september over 8 uker
- Mulig å melde seg på for å bli med neste gang det arrangeres <https://www.coursera.org/course/metadata>
- Fritt hvor mye innsats man ønsker å legge ned i det
- Ressursene er tilgjengelig etter at kurset er avsluttet



 [Quick Questions](#)

## ABOUT THE COURSE

[Syllabus](#)[Grading and Logistics](#)

## COURSE

[Announcements](#)[Video Lectures](#)

## EXERCISES

[Homeworks](#)[Surveys](#)

## Syllabus

# Metadata: Organizing and Discovering Information

Dr. Jeffrey Pomerantz - School of Information and Library Science - University of North Carolina at Chapel Hill

## About the Course

If you use nearly any digital technology, you make use of metadata. Use an ATM today? You interacted with metadata about your account objects, so that those objects may be organized now and found later.

Metadata is a tool that enables the information age functions performed by humans as well as those performed by computers. Metadata is an interdisciplinary field that studies how people create and manage information.

## Course Syllabus

### Unit 1: Organizing Information

- An introduction to the concept of metadata, and what it means to describe something. You'll also be introduced to the phenomenon of

### Unit 2: Dublin Core

- The Dublin Core Metadata Initiative is designed as a lowest-common-denominator metadata schema. It's the logical place to begin.

### Unit 3: How to Build a Metadata Schema

- You'll be introduced to the documents that have to exist for a metadata schema to be computer-processable. It's all about XML.

### Unit 4: Alphabet Soup: Metadata Schemas That You (Will) Know and Love

- Here we start a 3-unit series in which we'll look at metadata schemas for a variety of purposes. In this unit we'll discuss (non-Dublin C

# Video Lectures

[Help](#)

Having trouble viewing lectures? Try changing players. Your current player format is html5. [Change to flash.](#)

## ▼ Unit 1: Organizing Information

✓	1-1: Introduction - 12:26	    
✓	1-2: NSA - 7:28	        
	1-3: OSX Mavericks - 4:13	     
	1-4: Data About Data - 8:13	   
	1-5: Description - 11:55	  
	1-6: Item vs. Collection - 6:29	  
✓	1-7: LCSH - 10:42	  
	1-8: Thesauri - 14:17	   
	1-9: Faceted Classification - 6:40	  

# Kilder til kunnskap

## // NMC Horizon Library edition

- 6 trender
- 6 utfordringer
- 6 teknologier
- NMC Horizon EdTech Weekly App
- [library.wiki.nmc.org](http://library.wiki.nmc.org)
- Kort form: Forskningsdata, nye formater og «Web of Objects»



# NMC: Trender:

## 1) Research Data Management

- // (Forsknings-)data blir generert i stadig større tempo og omfang – av alle, overalt
- // Det er naturlig at bibliotekarer tar ledelsen i forvaltningen (kurator-/fasilitator-rolle) og sikrer gjenfinning
- // Forskere må ha "Data management plans" (DMP) for å få utbetalt forskningsmidler
  - Bibliotekarer må kunne bistå med å skrive DMP'er
  - Hjelpemidler:
    - "DCC Checklist" (<http://www.dcc.ac.uk/resources/data-management-plans/checklist>)
    - re3data.org
    - OpenAIRE.eu





# NMC: Trender:

## 2) Mobilt innhold

- // 50+% av alle voksne (amerikanere) har en enhet/brett **dedikert** til lesing av digitale dokument
- // Bibliotekene må fokusere på varige løsninger som er leverandørnøytrale (EPUB3 og HTML5)
- // Alle søketjenester må optimaliseres for bruk fra mobile enheter
- // Prioriter (også økonomisk) innhold som gir tilgang;
  - "The Right to e-Read"-kampanjen (<http://www.eblida.org/e-read/home-campaign/>)



# NMC: Utfordring:

## 1) Integrere biblioteket i læreplanen

- // Bibliotekarer må utvide sin oppfatning av hva deres rolle skal være i forbindelse med læreplanen og pensum
- // Strekk ut en hjelpende hand til fakultetet og hjelp dem med å ta biblioteket/bibliotekarene bedre i bruk
- // Forskningsbibliotekarer bør forstå og understøtte undervisnings- og forskningsprosessen ved egen institusjon



# NMC: Utfordring:

## 2) Bibliotekarens nye rolle

- // Klar tendens i ansettelse av mer funksjonelle spesialister, med sterk teknologisk bakgrunn
- // Akademiske- og forskningsbibliotekarer har utviklet seg mot emnespesialister som bistår forskning innen enkeltdisipliner
- // Økende kurator fokus; mer dybdekunnskap, preservering av data, vitenskaplig kommunikasjon, publisering (innsikt i copyright utfordringer), peer-review og gjenfinning



# NMC: Teknologi:

## 1) Åpent innhold

- // Open Educational Resources (OER) bruker Creative Commons (CC) og andre former for lisenser for å sikre fri gjenbruk
- // Biblioteket bør lede lokale OER-initiativ
- // Bibliotekarens nøkkelkompetanse innen OER er å finne relevant fagstoff som blir godkjent av fageksperten innen fagområdet – med avklart åpen lisens
  
- // Mer informasjon: "Nordic OER movement", <http://nordicoer.org/>,



# NMC: Teknologi:

## 2) Semantisk web og lenkede data

- // Den semantiske web gjør informasjon (og ikke bare data) tilgjengelig for datamaskiner
- // Lenkede data pakker den semantiske web'en sammen slik at datamaskiner kan "resonere"/finne sammenhenger over ulike samlinger med informasjon
- // Gjør det mulig å automatisk avgrense søkeresultat mellom øya, kaffen og programmeringsspråket "Java" – basert på konteksten søket er utført i
- // BIBFRAME er Library of Congress sitt initiativ for å lage lenkede data av MARC21-formatet



# Oppsummering

- // Mengder av data tilgjengelig som kan brukes til nye tjenesteutvikling
- // Datakildene vil bli koblet samme til en vev
- // Utfordringene er å finne hvilke gode tjenester man skal tilby, og som det bruk for, basert på de tilgjengelige dataene
- // Bibliotekarene må være gode på metadata og på brukernes behov
- // Biblioteket må ha kjennskap til teknologi og trender
- // Åpen og tilgjengelig formidling og gjenfinning

# Oppsummering

- // Biblioteket må være aktiv i lærings- og i forskningsprosessen
- // Innsikt og deltakelse i forskningsprosessen er vesentlig, for at biblioteket skal være en anerkjent og viktig del av egen institusjon.
- // Undervisningen i høyere utdanning skal være forskningsbasert, dermed er også innsikt og forståelse av forskningsprosessen viktig for at Biblioteket skal få bli en del av utdanningsprosessen.
- // Hvordan skal bibliotekaren kunne ha god kunnskap om egne brukeres behov - om de ikke har god innsikt i de to primærprosessene til brukerne?



# Takk for meg!

// E-post: [asbjorn.risan@bibsys.no](mailto:asbjorn.risan@bibsys.no)

// Telefon: 918 97 938

// Nettsted: <http://www.bibsys.no>

// Informasjon om metadata og tjenester:  
<http://www.bibsys.no/betadata>

