

Visjoner om datadeling og framtidens biblioteksystemer

Biblioteklaboratoriet, Bergen

november 2007

ole.husby@bibsys.no

"Store deler av forskningen er
finansiert via offentlige midler, og
forskerne gir fra seg resultatene sine
uten vederlag til forlagene for å få
dem publisert. Hvordan kan det da ha
seg at ett enkelt internasjonalt
vitenskaplig tidsskrift kan koste så
mye som 100000 kroner i året å
abonnere på?"

Jens Vigen, CERN

Dagbladet, 28. september 2007

Vi trenger åpenhet

- Åpen publisering
- Åpne arkiver
- Åpne arkitektur
- Åpne standarder
- Åpen kildekode
- Åpne biblioteker

fordi vi ønsker (trenger) å dele

- Dokumenter
- Tradisjonelle metadata
 - Bibliografiske data
 - Emnebeskrivelser
- Supplerende metadata fra utgivere
 - Omslagsbilder
 - Sammendrag, omtaler, innholdsfortegnelse
- Supplerende metadata fra brukerne
 - Tagging
 - Omtaler, anbefalinger

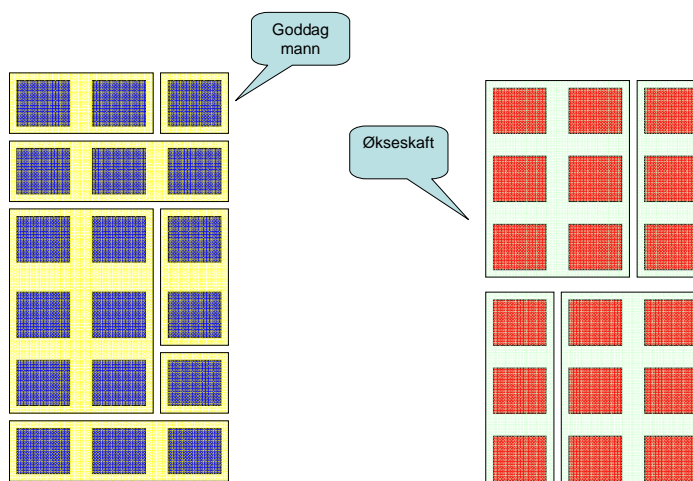
Åpen arkitektur er viktig

fordi det tillater oss å dele opp i biter som ikke er større enn nødvendig!

Arkitektur angår hvordan vi deler opp applikasjonene i enkelte tjenester

- hvordan vi deler opp innhold i enkelte dokumenter
- og hvordan vi deler opp metadata i enkelte poster

Arkitektur



Arkitektur

- Hvorfor kopiere hele databasen når vi trenger en post?
- Hvorfor abonnere på hele tidsskriftet når vi trenger en artikkel?
- Hvorfor katalogisere hele verket når vi får et nytt eksemplar?
- (FRBR)

Hva er SOA?

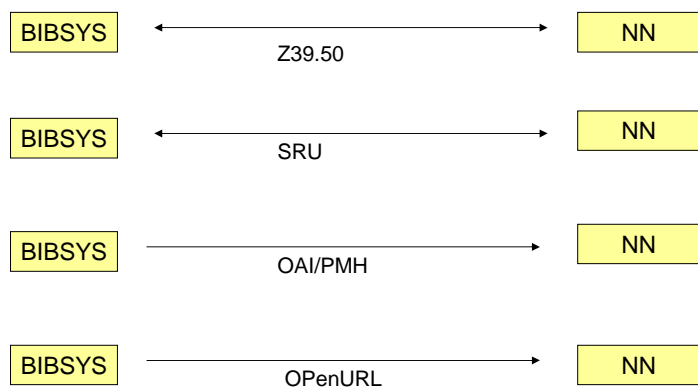
SOA = Service Oriented Architecture =
Tjenesteorientert Arkitektur

- Fokus på **tjenester**, ikke **data** eller **innhold**
- Viktig med **stabile** og **godt dokumenterte** grensesnitt (protokoller, API, formater etc)
- Viktig å bruke **åpne grensesnitt** (standarder)

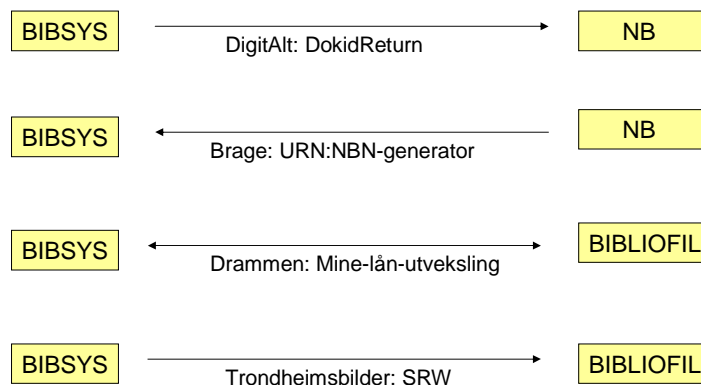
Mer om SOA

- Den enkelte tjenesten defineres gjennom tjenestegrensesnittet
 - det er altså ikke nødvendig å vite hvordan tjenesten ser ut innvendig, hvilke interne formater og databaseløsninger som brukes
- SOA kan realiseres vha Web Services

BIBSYS åpne tjenester



BIBSYS "lukkede" webtjenester



"Flere av systemene bærer preg av å være i "siste fase av en livssyklus". De bygger på en arkitektur og teknologi som innehar mange av de svakhetene som vi kjenner fra våre egne produkter, og er kanskje enda mer modne for modernisering. Det skulle i teorien være duket for et slipp av nye moderne "nestegenerasjons" systemer. Disse er det imidlertid ikke så lett å få øye på.

...

Vi vil hevde at det vi opplever nå er at de store monolittiske systemenes epoke kanskje er over, og at "neste generasjon" vil være bygd på mindre komponenter, åpne systemer og plattformer som kan settes sammen med langt større fleksibilitet enn det vi er vant til.

BIBSYS

Dokument til styret, mai 2007

Et biblioteksystem består av

Søk (BIBSØK, BIBSYS Ask)	Pliktavlevering
Utlån	Emne
Innlån	Katalog
Brukerbestilling	Eksport og import
Mitt BIBSYS	Servere (Z39.50, OAI-PMH, automatservere)
Tilvekst	Meldingssystem
Periodika	Administrativ modul
Lisensadministrasjon (Lisa)	

Men dette er ikke inkludert

Institusjonsarkiv (BIBSYS Brage)	Søkeportal (BIBSYS Mime)
Emneportal	Tidsskriftportal (BIBSYS Tyr)
Forskningsdatabase (BIBSYS ForskDok)	OpenURL tjeneste (BIBSYS X)
Billedatabase (BIBSYS Galleri)	

Markedstendenser

- Markedet for tradisjonelle LMS blir mindre
- Leverandørene blir færre
- Produktene blir færre
- Få nye (moderne) produkter
- Kunder med behov for å skifte ut gamle systemer kommer i beit
- Kundene må også bruke penger på andre produkter: ERM, LMS, portaler etc.

Hva trenger vi

- Åpne modulariserte SOA-baserte systemer som består av byggeklosser
- Horisontal og vertikal integrering
- Effektive felles-systemer for store konsortia, hvor datadeling er lett og hvor totalsystemet framstår som enhetlig
- Større gjenbruk av bibliografiske poster
- Samarbeid mellom systemleverandørene

Andre ting å tenke på

- Kost/nytte ved høykvalitets katalogisering
- Kost/nytte ved bruk av kontrollerte emneord og klassifikasjon
- På hvilke nettarenaer finner vi brukerne?
- Det er stor forskjell på brukere

"SRU is simple!"

<http://www.loc.gov/standards/sru/simple.html>

SRU request URL

```
http://z3950.loc.gov:7090/voyager  
?operation=searchRetrieve  
&version=1.1  
&query=chaos  
&recordSchema=info%3Asrw/schema/1/dc-v1.1  
&maximumRecords=10  
&startRecord=1
```

<http://z3950.loc.gov:7090/voyager?operation=searchRetrieve&version=1.1&query=chaos&recordSchema=info%3Asrw/schema/1/dc-v1.1&maximumRecords=10&startRecord=1>

En enkel SRU-klient i HTML

```
<html>
<head>
<title>SRU</title>
</head>
<body>
<h1>SRU søk i Library of Congress</h1>
<form method="GET" action="http://z3950.loc.gov:7090/voyager">
<input type="hidden" name="operation" value="searchRetrieve">
<input type="hidden" name="version" value="1.1">
<input type="hidden" name="recordSchema" value="dc">
<input type="hidden" name="startRecord" value="1">
<input type="hidden" name="maximumRecords" value="10">
<input type="text" name="query">
<input type="submit" value="Søk">
</form>
</body>
</html>
```

```
<?xml version="1.0"?>
<zs:searchRetrieveResponse xmlns:zs="http://www.loc.gov/zing/srw/"><zs:version>1.1</zs:version><zs:
<title>28 days later [sound recording] : the soundtrack album.</title>
<type>sound recording</type>
<publisher>United States : XL Recordings Ltd.</publisher>
<date>p2003.</date>
<language>eng</language>
<description>Strings, Andrew Price, Filmworks ; drums and percussion, Ged Lynch ; with various p
<description>Original score composed by John Murphy ; score mixed and produced by Daniel L. Grif
<description>Score recorded at Wonderland studios, Los Angeles; Chaos studios, Los Angeles; Parr
<description>The beginning -- Rage -- The church -- Jim's parents (Abide with me) -- Then there
<subject>Motion picture music.</subject>
<subject>Popular music--2001-2010.</subject>
</srw_dc:dc></zs:recordData><zs:recordPosition>1</zs:recordPosition></zs:record><zs:record><zs:rec
<title>The $5.00 EP [sound recording] </title>
<type>sound recording</type>
<publisher/>
<date>2000.</date>
<language>eng</language>
<description>Reaction time -- Better days -- Nothing's wrong -- Million miles away -- Nothing to
</srw_dc:dc></zs:recordData><zs:recordPosition>2</zs:recordPosition></zs:record><zs:record><zs:rec
<title>The Abbott and Costello show.</title>
<creator>Yarborough, Jean, direction, production.</creator>
<creator>Abbott, Bud, 1895-1974, cast.</creator>
<creator>Costello, Lou, cast.</creator>
<creator>Copyright Collection (Library of Congress) DLC</creator>
<type>moving image</type>
<type>Comedy Television series. migfg</type>
<publisher>United States : [s.n.], United States : Shanachie,</publisher>
<date>1952 ;</date>
<date>1992.</date>
<language>eng</language>
<description>Lou's selling pots and pans and cooks a dinner for the whole gang at the roominghou
```

SRU Response - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

SRU search client

CQL query: [search again](#)

Showing 1 - 10 of 3188 hits [get more](#)

- 1  **28 days later [sound recording] : the soundtrack album.**
United States : XL Recordings Ltd. : p2003.
(Strings, Andrew Price, Filmworks ; drums and percussion, Ged Lynch ; with various performers.)
[\(fetch\)](#)
- 2  **The \$5.00 EP [sound recording]**
2000.
(Reaction time -- Better days -- Nothing's wrong -- Million miles away -- Nothing to lose -- One.)
[\(fetch\)](#)
- 3 **The Abbott and Costello show.** / Yarbrough, Jean, direction, production.
United States : [s.n.], United States : Shanachie, ; 1958.
[\(fetch\)](#)
- 4 **Acting on impulse** / Irvin, Sam, direction.
United States : Academy Entertainment, ; c1993.
[\(fetch\)](#)
- 5 **Advanced dungeons & dragons : official game accessory.**
Lake Geneva, WI, U.S.A. : TSR, ; c1996-c1998.. ISBN: 0786903686
(Some vols. include bibliographical references or indexes.)
[\(fetch\)](#)
- 6  **[Advanced dungeons & dragons publications]**
Renton, WA : Wizards of the Coast, ; c1999-cc2001 >. ISBN: 0786913835
[\(fetch\)](#)
- 7  **Advanced topics in nonlinear control systems** / Leung, T. P.
Singapore ; River Edge, NJ : World Scientific, ; c2001.. ISBN: 9810244428
(Includes bibliographical references.)
[\(more\)](#) [\(fetch\)](#)

Done

Søk i Library of Congress via SRU + XSL

SRU Response - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

SRU search client

CQL query: chaos [search again](#)

Showing 1 - 10 of 1285 hits [get more](#)

- 1 **Evolution of order and chaos in physics, chemistry and biology proceedings of the International Symposium on Synergetics at Schloss Elmau, Bavaria, April 26-May 1, 1982** / International Symposium on Synergetics Schloss Elmau
Berlin Springer : 1982. ISBN: 3540119043
[\(fetch\)](#)
- 2  **Lords of chaos the bloody rise of the satanic metal underground** / Moynihan, Michael
Venice, CA Feral House : 1998. ISBN: 0922915482
(Omhandler satanismen i Norge, og bl.a. Varg Vikernes).
[\(fetch\)](#)
- 3 **Laws of chaos a probabilistic approach to political economy** / Farjoun, Emmanuel
London Verso : 1983. ISBN: 086091058x
[\(fetch\)](#)
- 4  **Beyond heterochrony the evolution of development** / Zelditch, Miriam Leah
New York Wiley-Liss : 2001. ISBN: 0471379735
[\(more\)](#) [\(fetch\)](#)
- 5 **Challenging cancer from chaos to control** / Kfir, Nira
London Tavistock/Routledge : 1991. ISBN: 0415063434
[\(fetch\)](#)
- 6 **Chaos en beheersing documenten uit aeneolitic Egypte** / Asselberghs, Henri
Leiden : 1961
[\(fetch\)](#)
- 7  **Exchange rate chaos twenty-five years of finance and consumer democracy** / Geisst, Charles R.
London Routledge : 1995. ISBN: 0415109817
[\(fetch\)](#)

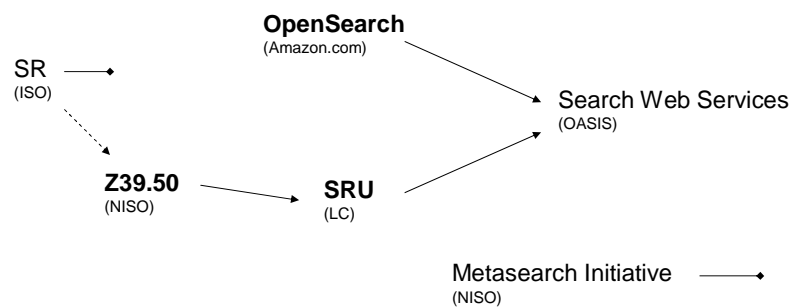
Done

Samme søk i BIBSYS, samme XSL

BIBSYS SRU beta

- BaseURL:
 - <http://sru.bibsys.no/services/sru>
- Contextsets:
 - dc
 - bibsys
- Recordschemas:
 - dc
 - marcxml

Alternativer til SRU



"Eiendomsrett" til poster i BIBSYS-sammenheng betyr ikke at postene er tilgjengelige på børs. Tvert imot, alle kan fritt søke fram og kopiere poster fra Bibliotekbasen. Reglene for kopikatalogisering fra BIBSYS er: "Any user has free access to search and to download single records in any of the record syntaxes (formats) that are supported. It is prohibited to do bulk download of records."

Roy Gundersen, BIBSYS

BIBLIOTEKNORGE@nb.no, 22. august 2007

"Everyone a librarian

The response has been startling to say the least: In three days, users have combined 17,000 times, mashing together 42,000 works!"

Tim Spalding

The LibraryThing Blog, 23. februar 2006