



Det Humanistiske Fakultet



DK-CLARIN WP5 Workshop Metadatasøgning

29. juni 2010

DK-CLARIN WP5:
Christian Wolfgang, CST
Lene Offersgaard, CST
Tomas G. Christensten, DSL



Dagsorden

Intro og formålet med workshoppen

Typer af søgning:

- Ren metadatasøgning
- Fuldtekstsøgning
- Kombineret fuldtekstsøgning og metadatasøgning

Søgesyntaks

Prioritering af søgemuligheder

Evt



Intro

Dialog mellem brugere og udviklere:

- Hvordan vil I gerne kunne søge på metadata?
- Diskutere typer af søgning og syntaks

Metadata har flere funktioner

- beskriver ressourcen (ved import/export)
- beskriver relationer til andre ressourcer
- indeholder infrastrukturens administrative oplysninger
- bruges i søgning

Når en ressource er importeret i repositoret så har den metadata tilknyttet

**Et antal metadata-elementer er fælles for alle ressourcer.
Hver ressourcestype har også sine 'egne' metadata-elementer.**

Metadata-formater i forskellige standarder:

bl.a. DublinCore, OLAC, TEIP5, IMDI, CMDI



Intro -2

Generelle metadataelementer, bl.a.

- Type
- Title
- Language
- Format
- Publisher
- PublicationDate
- Rights(license)
- ContentProvider
- Creator(forfatter/værktøj)
- Description
- ConformsTo(fx link til xml-schema)
- InfoAbout (link til dokumentation)

Repositoriet tilføjer desuden nogle metadata, bl.a.

PID, ModificationDate



Intro -3

Typer af søgning

Ren metadatasøgning

- Hvad skal man kunne søge på?
- syntaks

Fuldtekstsøgning

- Hvad skal man kunne søge på?
- Syntaks

Kombineret fuldtekstsøgning og metadatasøgning

- Hvad skal man kunne søge på?
- syntaks



Ren metadatasøgning - Simpel

Simple søgning 1: Fast formular

- Faste metadataelementer givet ressourcetypen
- Alle elementer

Simple søgning 2: Dynamisk ekspanderende formular

Metadata Relation Værdi		
Type	= ▼	Type
Title	= ▼	Title
Date	> ▼	Date
<input type="button" value="Søg"/>		

Metadata Relation Værdi		
Type ▼	= ▼	Type
Title ▼	= ▼	Title
Date ▼	> ▼	Date
<input type="button" value="Tilføj nyt element"/>		
<input type="button" value="Søg"/>		



Ren metadatasøgning

Avanceret søgning

- Fri brug af metadataelementer
- Sammensæt elementer via boolske udtryk
- Grupperinger med parenteser

Title = title OR Creator = author

Title = title OR (Creator = author AND Date > 2000)

CQL-søgning

- Proximity, ord der er i nærheden
- Fuzzy, ord der minder om



Fuldttekstsøgning

Rettigheder og søgning

Søgning i 'pub' og 'pub+aca'

Søgesyntaks

Brugeren skriver selv sin søgestreng

Googlelike-syntaks

CQL-syntaks



Kombineret fuldtekstsøgning og metadatasøgning

Formbaseret-søgning

Googlelike syntaks

CQL



Prioritering af søgemuligheder

- **Ren metadatasøgning**
 - Simpel søgning
 - Avanceret søgning
 - CQL-søgning
- **Fuldtekstsøgning**
 - Søgning i 'pub' og 'pub+aca'
 - Googlelike
 - CQL-syntaks
- **Kombineret fuldtekstsøgning og metadatasøgning**
 - Formbaseret-søgning
 - Googlelike syntaks
 - CQL
- **Andre ønsker**



Næste skridt

Udarbejde opsummering

- Prioriterede søgemuligheder og syntaks
- Andre ønsker/tanker frem
- Tilbage melding først omkring 9. juli

Implementeringsplanlægning

Tak fordi I kom!

