

# CATCH



CONTINUOUS  
ACCESS  
TO  
CULTURAL  
HERITAGE  
PLUS

## Services for Digital Cultural Heritage

Hennie Brugman

Technical coordinator CATCHPlus

Max-Planck-Institute for Psycholinguistics  
Netherlands Institute for Sound and Vision

BigGrid/CLARIN Infrastructure Landscape Workshop - March 8, 2010



## Overview

- CATCH and CATCHPlus
- CATCHPlus and infrastructure for Digital Cultural Heritage
- Case: Vocabulary and Alignment Service
- Concluding remarks



# CATCH & CATCHPlus

- CATCH research program by NWO (14 projects)
- CATCHPlus valorisation project
  - 8 subprojects at large CH institutions
    - Deliver (re)usable tools and services
  - Connected by common services concerning
    - terminology
    - annotations
    - metadata (collection catalogs)
    - Content
- CATCHPlus project bureau hosted by Netherlands Institute for Sound and Vision
- [www.catchplus.nl](http://www.catchplus.nl)

BigGrid/CLARIN Infrastructure Landscape Workshop - March 8, 2010

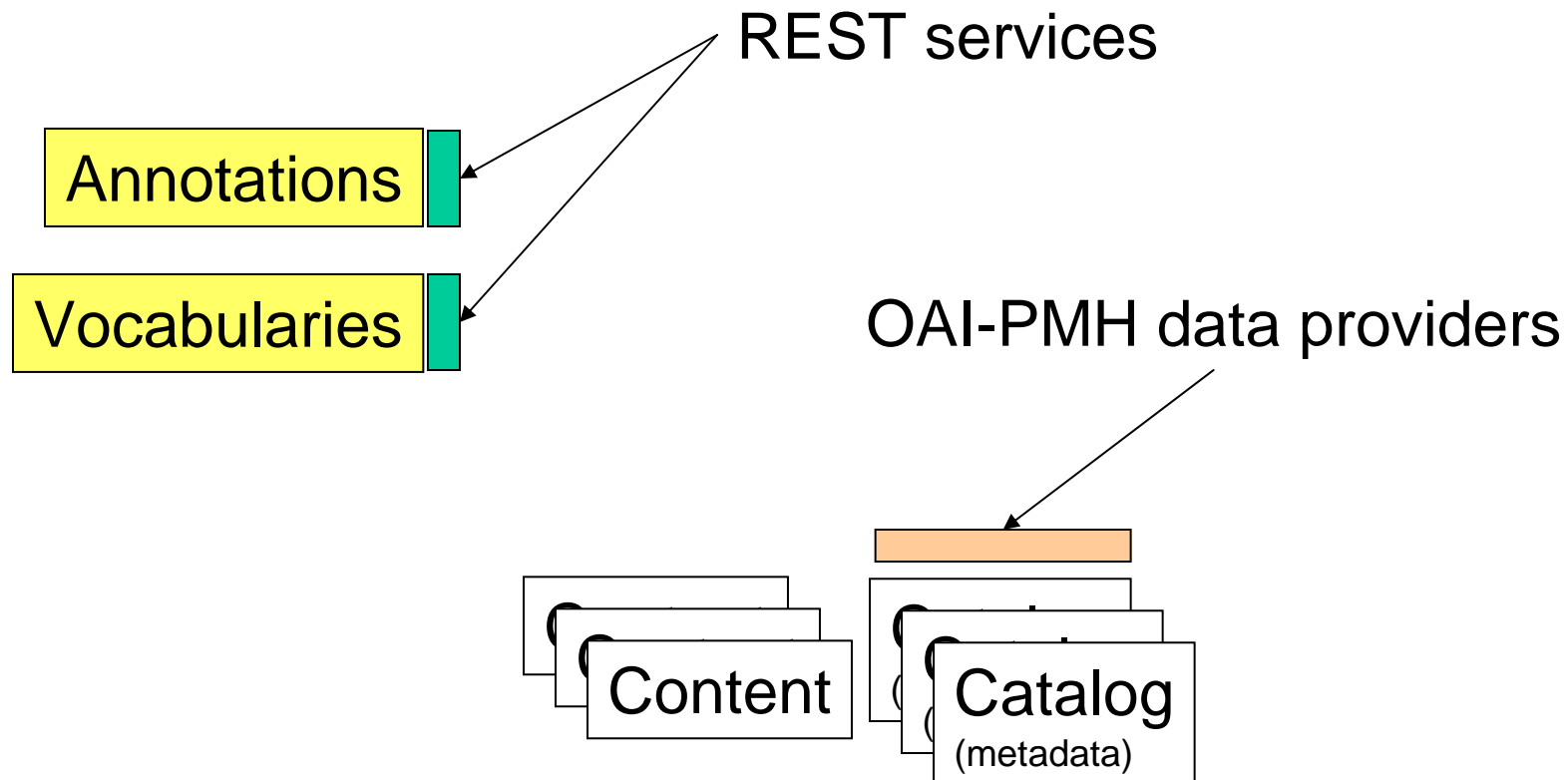


## CATCHPlus and infrastructure for digital cultural heritage

BigGrid/CLARIN Infrastructure Landscape Workshop - March 8, 2010

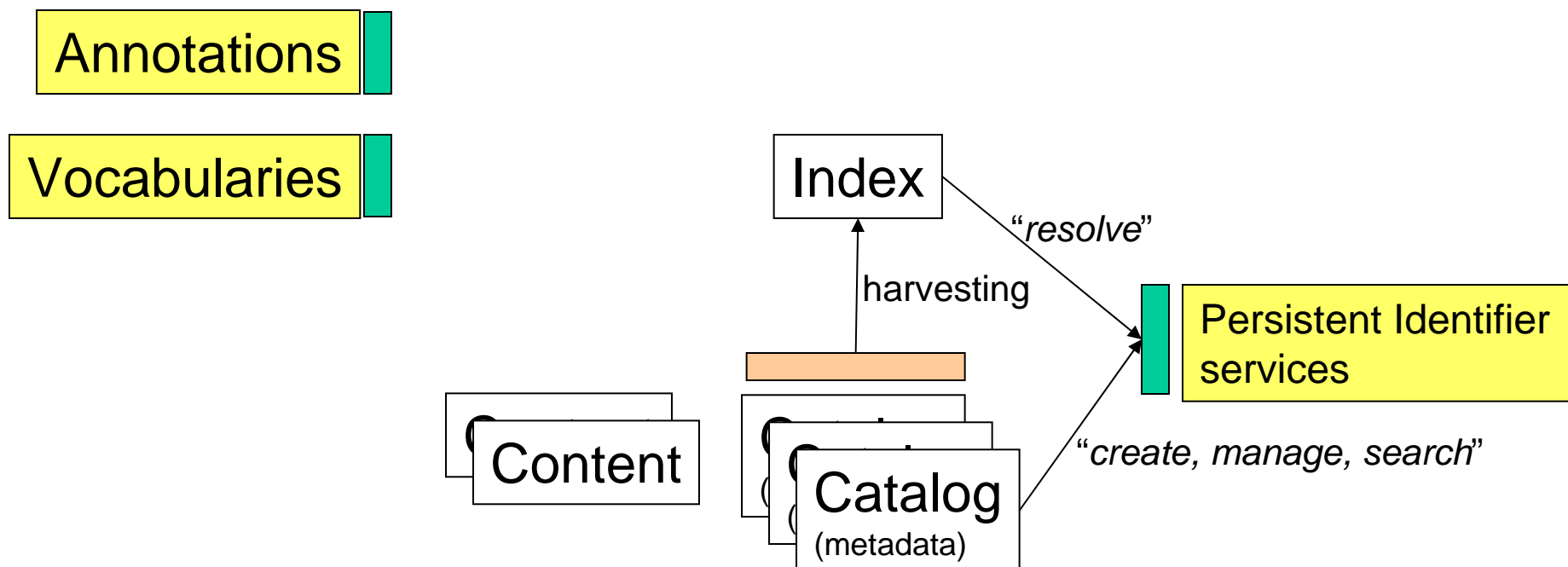


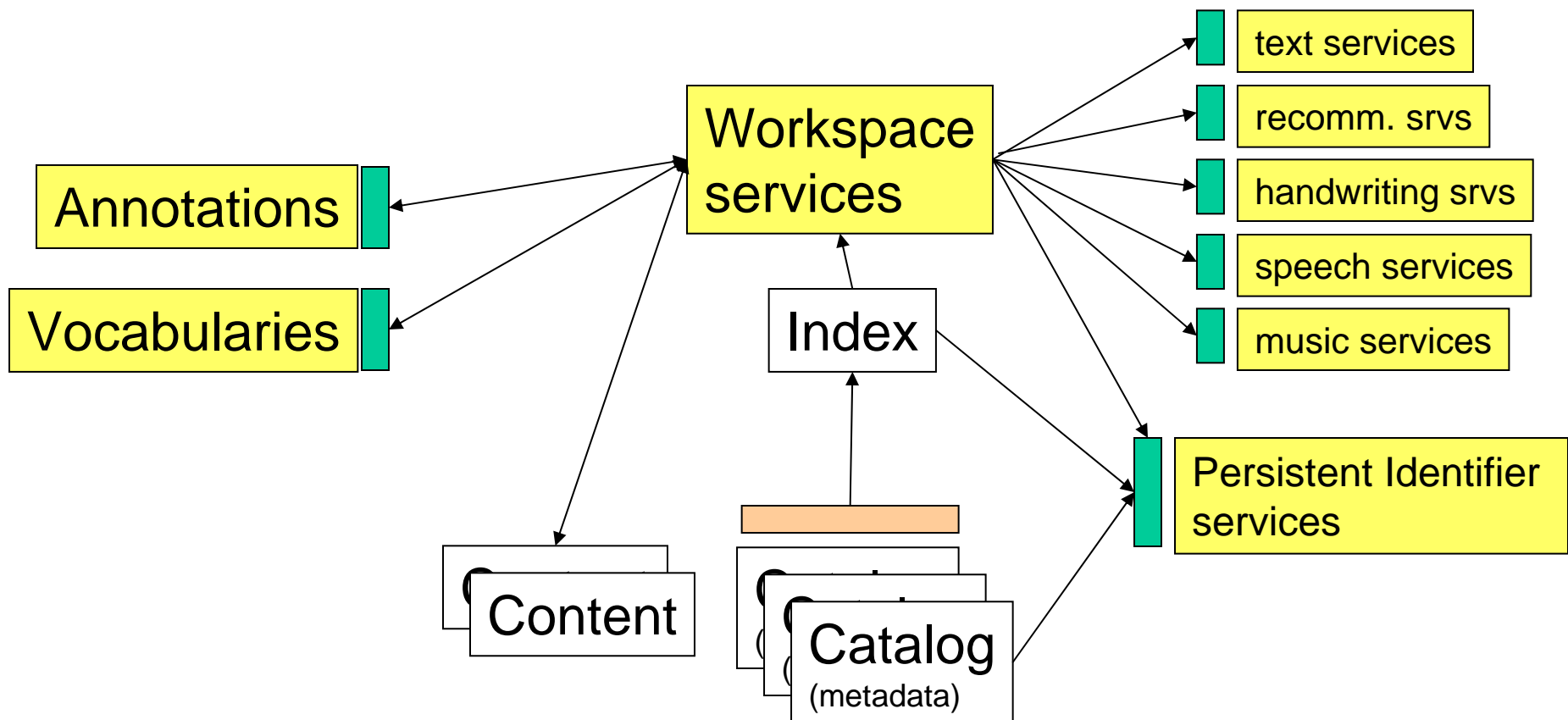
## CATCHPlus service landscape

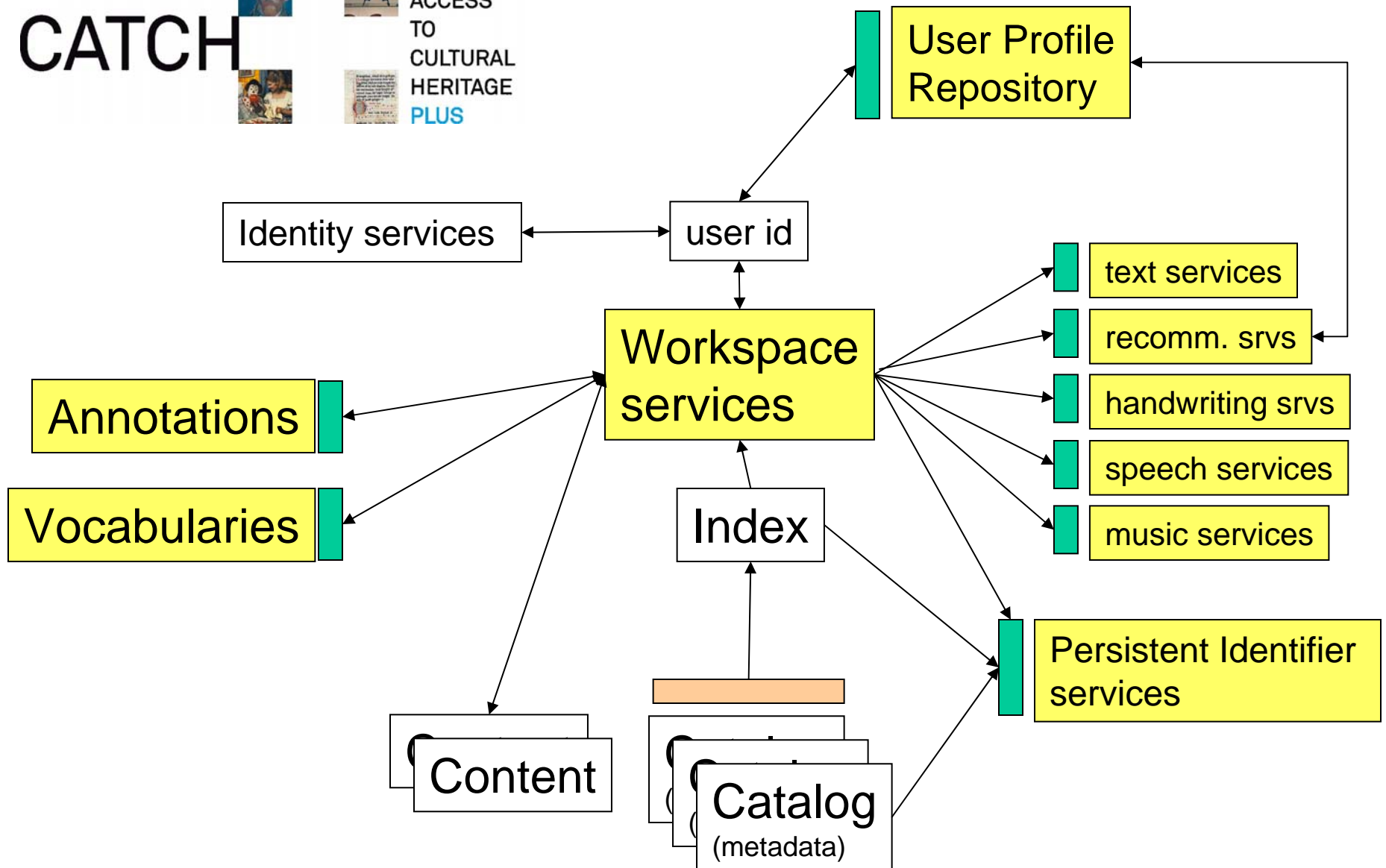




## CATCHPlus service landscape



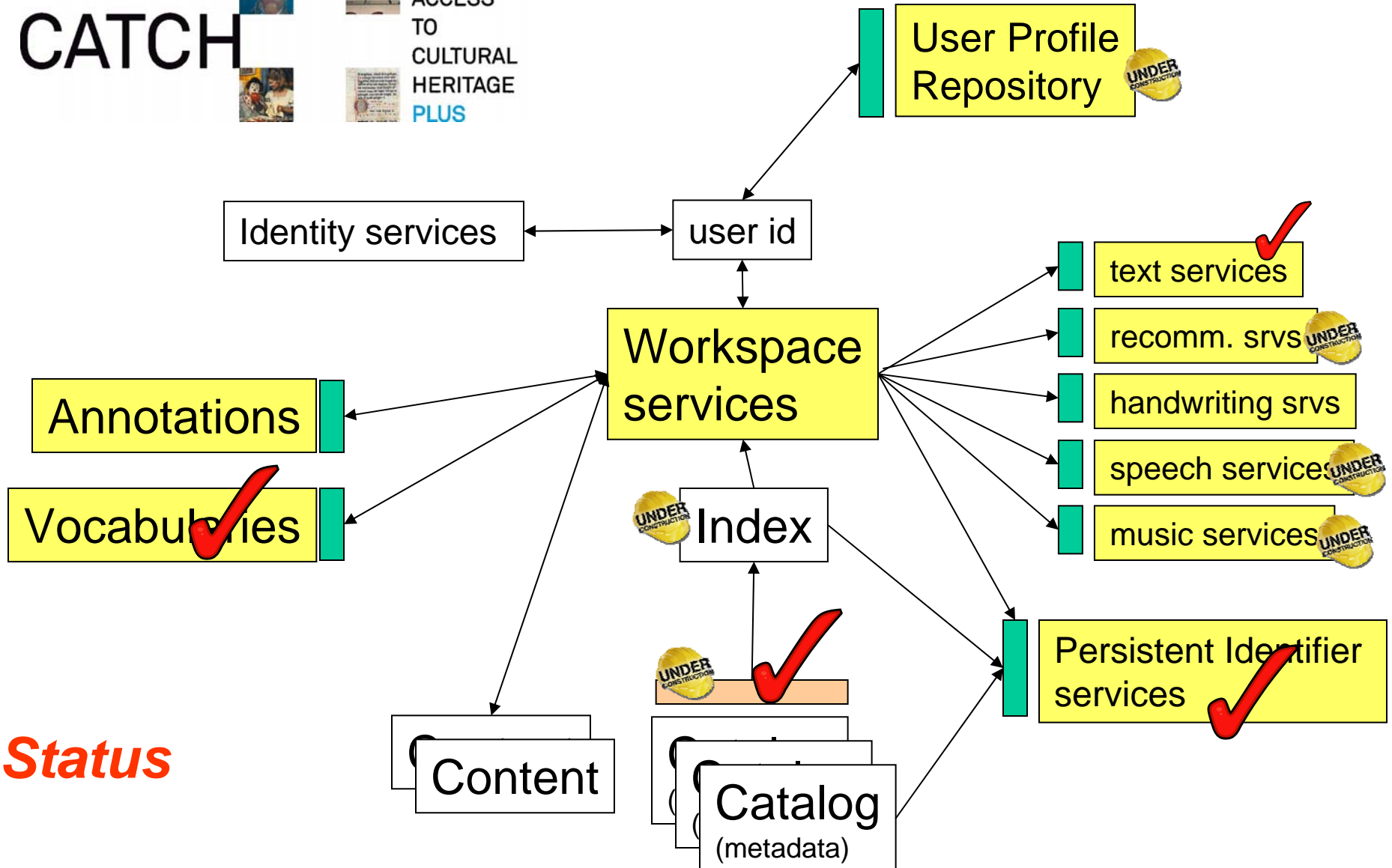






# CATCH

CONTINUOUS ACCESS TO CULTURAL HERITAGE PLUS



**Status**





## Case: Vocabulary and Alignment Service



## VAS aims

- Standard *format* and *access* methods
  - SKOS, SKOS based REST API
- Web *publication* of vocabularies
  - As searchable and browsable dataset → REST API
  - As Linked Data
  - Usable for sustainable references to concepts → PIDs
- Improve *semantic interoperability* by supporting alignments
- Centralised arrangements for *licensing*



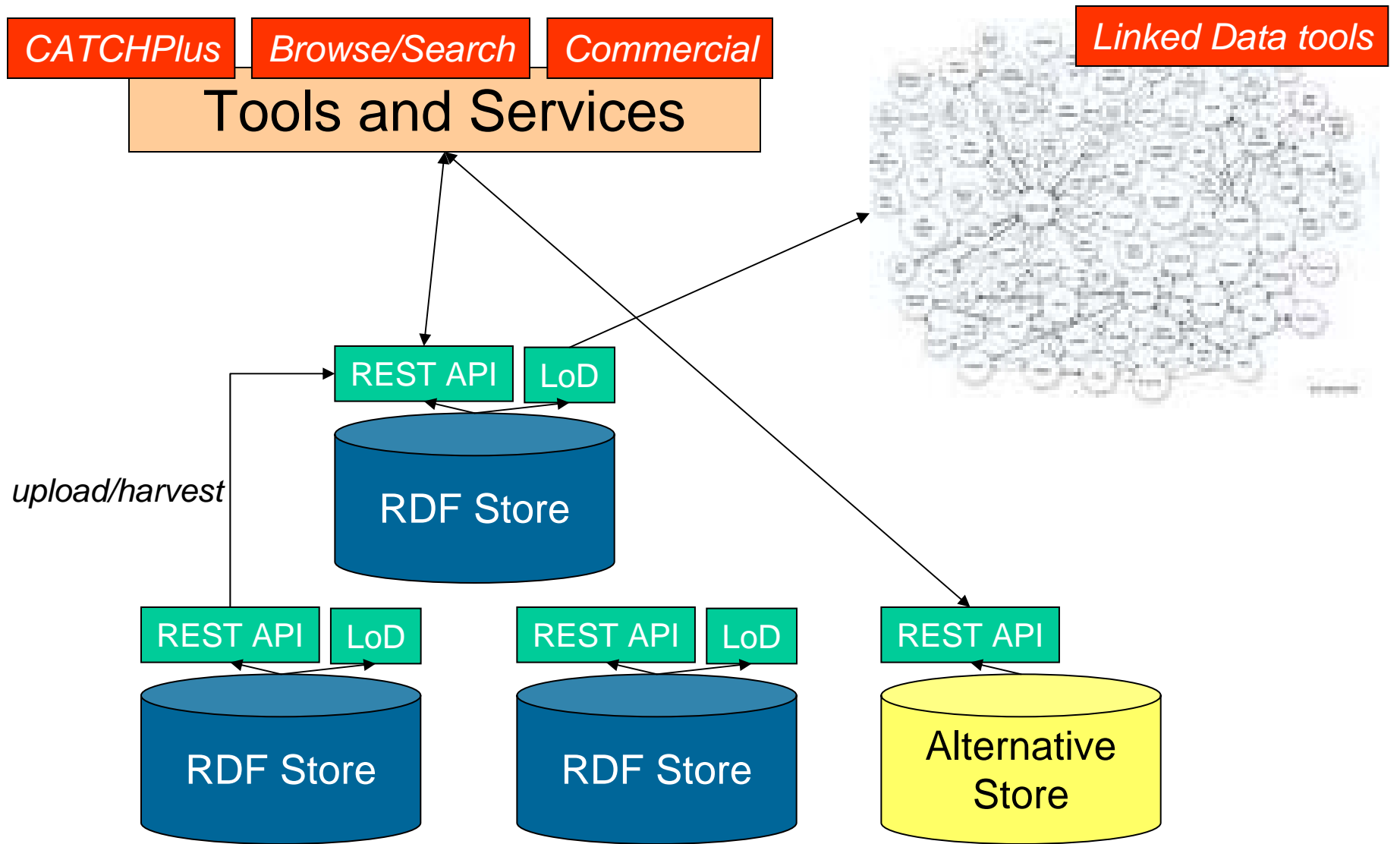
## Use cases

- Use cases from CATCHPlus and Cultural Heritage
  - Publish your thesaurus: import SKOS vocabulary, then get REST access, tool support and Linked Data for free.
  - Use for resource description: concept selection
  - Use for browse and search (both terminology and collections)
    - VAS Repository as topic map for CH collections
  - Use for thesaurus maintenance by online communities
  - Query translation, expansion, refinement
  - Etc.



## What is it?

- Repository for SKOS data (including alignment data)
  - RDF store (Virtuoso)
- REST API on top (search, autocomplete, upload, download), based on SKOS data model
- Linked Data interface
- Both persistent identifiers and stable URIs
- Future functionality:
  - Distributed operation
  - “live connections” with thesaurus databases → automatic updates



BigGrid/CLARIN Infrastructure Landscape Workshop - March 8, 2010



## Client tools and services

- CATCHPlus cases (semantic annotation, ranking, art recommender, ...)
- Commercial collection management software builder uses API to include thesaurus information
- Generic browse and search web application (using the REST API)



**Browse**

- Brinkman Trefwoorden
- UNESCO KB
- NUR:Nederlandstalige Uniforme Rubrieksindeling
- Gemeenschappelijke Thesaurus Audiovisuele Archieven**
- NBD/Biblion
- Regio Thesaurus KB
- OCLC NL:Karakteristiek in gecodeerde vorm
- Referentie Collectie Glas

Geografische namen

- Namen
- Classificatie**
- Genre
- Persoonsnamen
- Maker
- Onderwerpen

Zoeken op

Top concepten

- aardrijkskunde 15A
- algemeen 00X
- bestuur en geschiedenis 03B
- communicatie en media 11C**
- economie 05E
- gezondheid 06G
- kunst en cultuur 12K

Geen bredere bekend.

**SKOS nauwere**

- informatievoorziening overig 11C3**
- communicatie en media algemeen 11C0
- massacommunicatie en -media 11C2
- audiovisuele en elektronische media 11C4
- communicatietransport 11C1
- Geen gerelateerde bekend.

woordenboeken

- public relations
- audiovisuele archieven
- conferenties
- websites
- jaarverslagen
- naslagwerken
- bibliotheken
- archieven**
- encyclopedieën
- hoorzittingen
- bibliotheken

**SKOS bredere**

- informatievoorziening overig 11C3
- non-profitorganisaties
- SKOS nauwere**
- audiovisuele archieven**
- SKOS gerelateerde**
- collecties
- documenten
- opslagplaatsen
- musea
- bibliotheken

**SKOS bredere**

- informatievoorziening overig 11C3
- audiovisuele en elektronische media 11C4
- archieven
- Geen nauwere bekend.
- SKOS gerelateerde**
- filmconservering
- beeld- en geluidsdragers
- audiovisuele media

**Details**

**Concept: audiovisuele archieven**

**Identifier**  
<http://www.beeldengeluid.nl/Thesaurus/27697>

**Voorkeurstermen**

Nederlands  audiovisuele archieven

**Niet-voorkeurstermen**

Nederlands  filmarchieven  geluidsarchieven

**Gevonden in**  
[Onderwerpen](#)

**SKOS bredere**

- [informatievoorziening overig 11C3](#)
- [audiovisuele en elektronische media 11C4](#)
- [archieven](#)

**SKOS nauwere**

Geen bekend.

**SKOS gerelateerde**

- [filmconservering](#)
- [beeld- en geluidsdragers](#)
- [audiovisuele media](#)



## Status

- Currently contains 12 thesauri (most are not yet licensed)
- Browse/search tool (version 1) is ready
- Attracting interest from
  - Thesaurus providers
    - VU, Wageningen SemWeb group, RKD, CLARIN-NL
  - Tool builders
    - collection management software builders
  - Opportunity for API and/or technology harmonisation
- Used for collaboration of Beeld en Geluid and National Archive on their GTAA thesaurus
- Candidate for Open Source development?



## Concluding remarks

- Many services that CATCHPlus builds or needs are quite generic
  - We have services to offer and services to ask
- Cultural Heritage ICT departments are interested in infrastructural services
- Harmonisation of APIs
- We started with REST (+mashups). Additional need for SOAP (+service bus)?
  - Current CATCHPlus answer: no.
- Most CATCHPlus services need to be reliable and performant. Storage capacity is less of an issue.



Thank you. Questions?