



CATCHPlus

CATCH meeting AGORA, June 11th 2010

Julia Vytopil

Project office CATCHPlus

CATCH meeting Agora, June 11 2010, Julia Vytopil, Hennie Brugman



CATCHPlus

- CATCHPlus: building on successful CATCH projects
- Valorisation of research results
- From knowledge and prototype to reliable applications



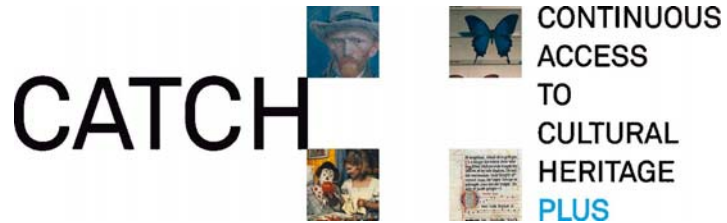
facts & figures

- 3 year project (2009-2011)
- partners: 9 heritage institutions, 3 research institutions and ICT industry
- total costs ± 3,2 M€, financed by:
 - project partners
 - Ministry of Education, Culture and Science
 - Ministry of Economic affairs (through PRIMA measure)
 - Netherlands Organisation for Scientific Research



Expected results

- tools for cultural heritage
- interoperability between institutions and collections
- business models for sustainable collaboration and management of products



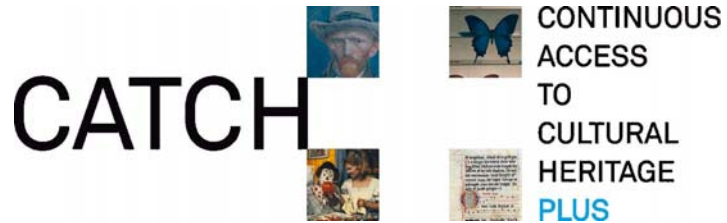
Crucial issues

- generic and open
- tools developed by/for heritage institutions
- tested in at least 1 other institution
- independent commission judges usability and genericity



Business models

- Deliverable in project description
- During and after project
- Collaboration between industry and institutions
- Maintenance and management of software
- Services based on software



Lessons learned so far

- Building on business model ideas from CATCH
- Bridging results from CATCH to plans in CATCHPlus
- ICT industry wants to be engaged



Overview of CATCHPlus technical results

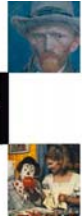
CATCH meeting Agora, June 11 2010, Julia Vytopil, Hennie Brugman



Technical coordination connects sub projects by stimulating:

- Tools that interoperate via web services
- Exchange and accessibility of data
 - Formats: SKOS, annotation format, DublinCore
 - Access: OAI-PMH, (REST) web services, Linked Open Data → Vocabulary Repository
- Identify and develop widely usable components
 - Checkers, User Profile Repository
- Persistence of products, including references/links
 - Persistent identifiers
- Efficiency of development: bundle efforts
 - Example: Workspaces

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Technical coordination connects **sub** **projects** by stimulating:

- Products that interoperate via web services
- Exchange and accessibility of data
 - Formats: SKOS, annotation format, DublinCore
 - Access: **OAI-PMH**, (REST) web services, Linked Open Data → **Vocabulary Repository**
- Identify and develop widely usable components
 - **Checkers, User Profile Repository**
- Persistence of products, including references/links
 - **Persistent identifiers**
- Efficiency of development: bundle efforts
 - Example: **Workspaces**

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS



- Advanced **MuSEUM**
- Entity detection **Checkers**
- Automatic **Multiply**
- Toolset for **WITCHCF**
- Searching **SCRATCH**
- Alignment **STITCHPI**
- Art Recon **CHIP API**

- Speech Recognition tools
Spraakherkenning

Checkers (MITCH and RICH spin offs – Naturalis, RCE)

- Detect named entities in uploaded PDF and HTML texts
- Detect potential errors in (database) table data

User Profile Repository (CHIP spin off - RMA)

- Store and share (semantic) user preferences and ratings in a central repository
- Show and provide ratings for objects on any web page
- Status: technical design is made

/ Gemeentearchief Rotterdam

<http://vocrep.q42.net/index.html>

CATCH ■ ■ ■ **Voculairebank Beta**

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

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.

Top concepten

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

woordenboeken

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

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

Niet-voorkeurstermen

Nederlands

Gevonden in
[Onderwerpen](#)

SKOS bredere

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

SKOS nauwere

Geen bekend.

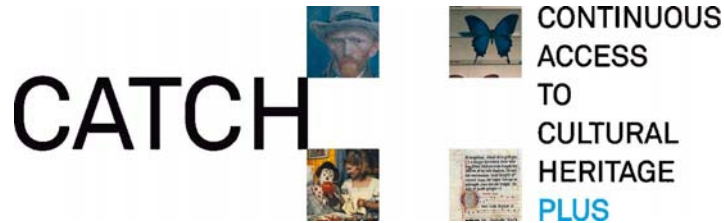
SKOS gerelateerde

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

Done

start | Inbox - Microsoft Out... | D:\CHOICE\CATCH-d... | REST sample calls - Mi... | Microsoft PowerPoint ... | Catch Plus Vocubular... | 6:27 PM

CATCH meeting Agora, June 11 2010, Julia Vytopil, Hennie Brugman



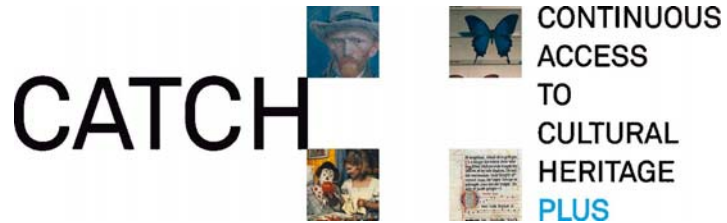
Vocabulary and Alignment Repository (1)

– Products

All GTAA concepts starting with “cul”

http://catchplus.tuxic.nl:18080/vas/api/find/concept?format=xml&query=^cul&match_kind=regex&limit=100&format=xml&cs=http%3A%2F%2Fwww.beeldengeluid.nl%2FThesaurus%2FOnderwerpen&sort=match_label&match_language=nl

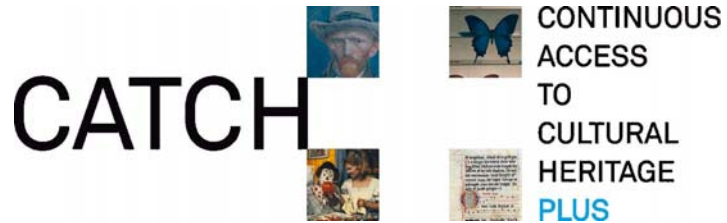
- GUI – generic thesaurus browsing and searching web application



Vocabulary and Alignment Repository (2)

– Applications

- Primarily for CATCHPlus sub projects
- Beeld en Geluid and National Archive collaboration
 - Pictura uses the API from their Memorix software
- Interest expressed by CLARIN-NL, Werkgroep Erfgoedthesauri, RKD, Adlib etc.



Vocabulary and Alignment Repository (3)

– Future

- Distributed operation
- “live” connection to thesaurus source databases
- Link to DEN thesaurus overview, both ways



Persistent identifiers (1)

– Products

- Requirements analysis and technology overview (see www.catchplus.nl)
- Basic choice: Handle technology
- Implementation:
 - One “Local Handle System” per participating organisation → you own your own ‘domain’
 - REST web service for identifier management
 - » From context of collection management system
 - » From context of metadata publishing



Persistent identifiers (2)

- Pilot projects/applications
 - Beeld en Geluid
 - Interested: Naturalis, RCE
 - National Archive uses Handles for NA4All
- Organisational
 - EPIC consortium (SARA, MPG/GWDG, CSC) provides sustainable hosting and services



Metadata standards and harvesting

- Support introduction of OAI-PMH “data providers” for standardised publication of metadata for harvesting
 - DublinCore
 - Qualified DublinCore
 - Proprietary formats
- Beeld en Geluid, KB, RMA, Naturalis
- First application: Nederlands Erfgoed: Digitaal!
 - Harvests to central index



Workspaces and annotations

- Workspaces
 - Shared solution for login, uploading resources, add metadata, bundle resources
 - This functionality is part of most CATCHPlus sub projects
 - Search, browse, annotate
 - Launch web services on workspace content
 - Possibly collaboration with BigGrid, CLARIN-NL and LifeWatch
- Annotations:
 - Exchange format
 - Repository for storing, searching and browsing annotations

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Conclusions

- Vocabulary Repository and Persistent Identifier services are working and attracting substantial interest
- Common services can save sub projects (and others) work
- Common services stimulate standardisation and integration
- Common services contribute to digital cultural heritage infrastructure
- One year on the way: conditions not always easy, but products are showing up

CATCH meeting Agora, June 11 2010, Julia Vytopil, Hennie Brugman



Thank you. Questions?

For more information:

www.catchplus.nl

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

CATCH Symposium, June 14-20, 2010, Uillevoytpil, Hennie Brugman