



B2Share

Easy data deposit & share

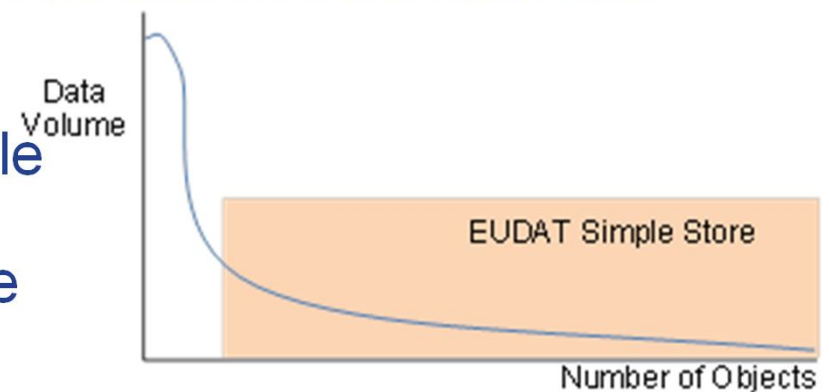
Ilja Livenson
PDC/SNIC
October 28, 2013





Motivation

- Address the “long tail” of “small” data and the researchers/citizen scientists creating/manipulating them
- EUDAT will offer a simple storage service - **B2Share** to
 - Offer a service to *everyone* with a data
 - Get a wide visibility in the research world
 - Address the many researchers and citizen scientists and not just the “few” organized in large communities
 - Make data accessible, shareable and persistent
 - Offer a strong European service





Approach

- Start simple
- A web service with a simple deposit flow
- Integration with EUDAT services behind the curtains
- Open-source and evolving based on user feedback
- Hosted by EUDAT partners



Metadata Catalogue

Aggregated EUDAT metadata domain.
Data inventory



AAI

Network of trust among authentication and authorization actors

PID

Identity
Integrity
Authenticity
Locations

Safe Replication

Data curation and access optimization



Data Staging



Simple Store

Researcher data store (simple upload, share and access)



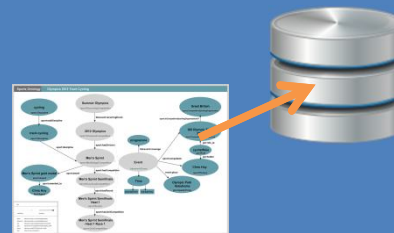
EUDAT Box

dropbox-like service
easy sharing
local syncing



Semantic Anno

checking & referencing



Dynamic Data

immediate handling



services
to come



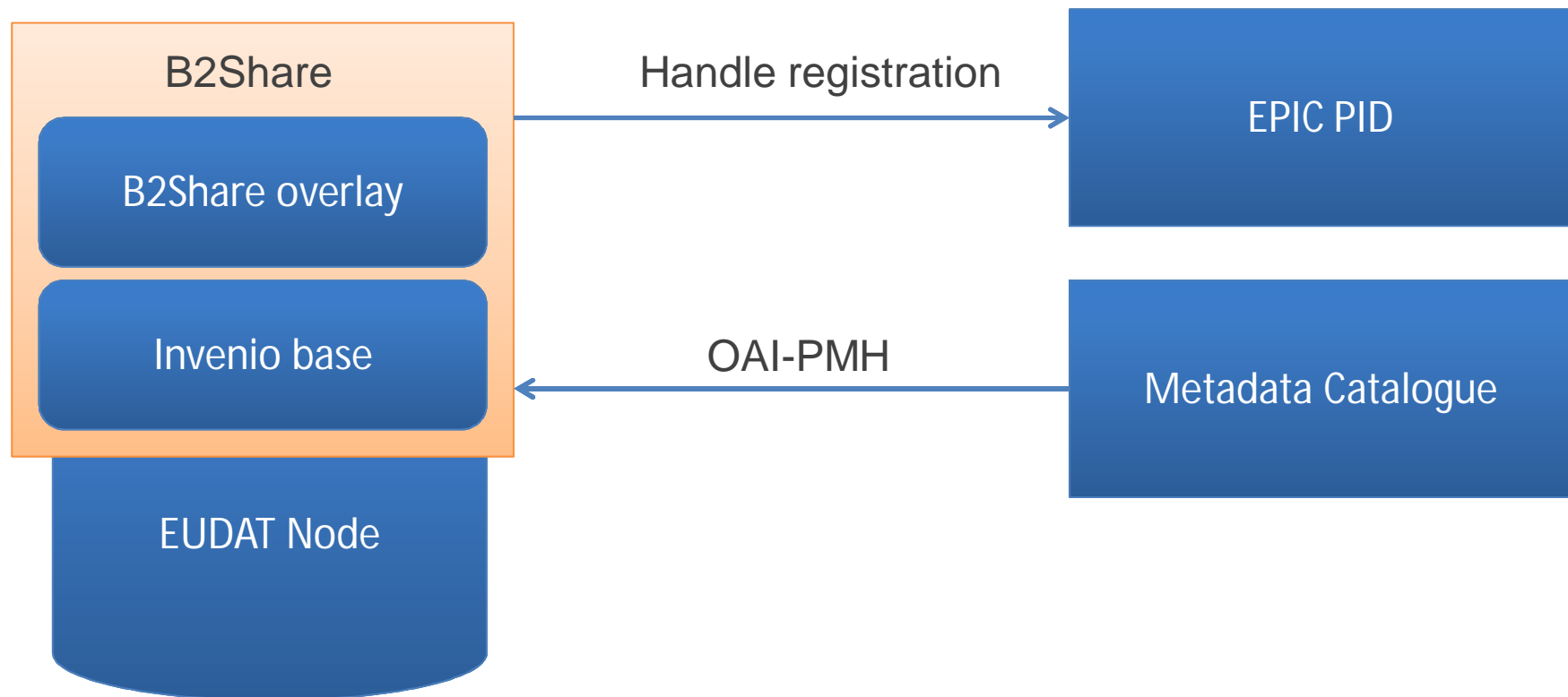


B2Share functionality

- Data upload
 - Metadata tagging
 - Parameterized by the scientific domain
- Search and retrieval
- Persistent references
- Metadata export



Architecture





Himalayan Document Centre
FOR MOUNTAINS AND PEOPLE

ICIMOD

About ICIMOD Programmes News & Events Information Resources HOME CONTACT US SEARCH

Search Collections any field Search Search Tips | Advanced Search

BROWSE BY TYPE

- Articles, Reports, Theses (0,183)
- Books, Proceedings (4,629)
- Multimedia (3,085)
- Maps, Atlases (3)
- Posters, Banners (3)
- Periodicals (213)

BROWSE BY TOPIC

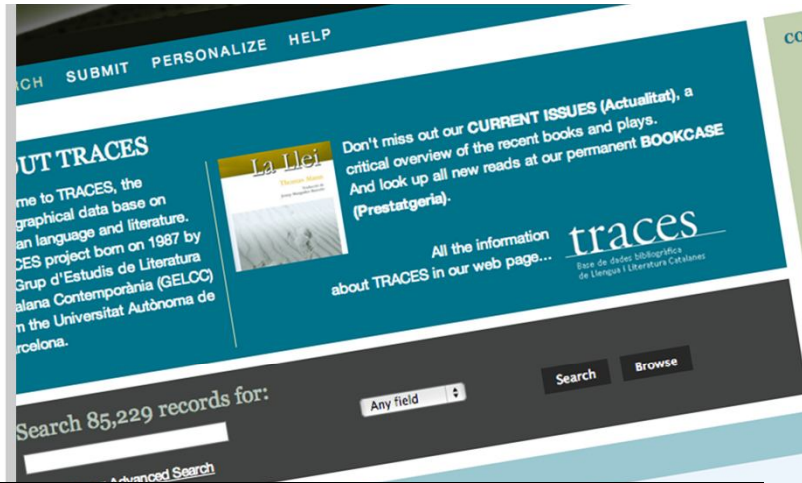
- Climate change (923)
- Cryosphere (343)
- Hazards and disaster (383)
- Knowledge management (628)
- Mountain livelihoods (372)
- Natural resource management (1,372)
- Policies and governance (2,011)

SEARCH COLLECTIONS

- Teri

SEARCH BY COLLECTIONS

- ICIMOD Publications
- ICIMOD Photo Gallery
- Other Publications by ICIMOD Staff
- PhD Theses
- SANDEE Publications
- HICAP
- Dr. NS Jodha Publications



SEARCH SUBMIT PERSONALIZE HELP

WELCOME TO TRACES, the geographical data base on Catalan language and literature. TRACES project born on 1987 by Grup d'Estudis de Literatura Catalana Contemporània (GELCC) in the Universitat Autònoma de Barcelona.

Don't miss out our **CURRENT ISSUES (Actualitat)**, a critical overview of the recent books and plays. And look up all new reads at our permanent **BOOKCASE (Prestatgeria)**.

All the information about TRACES in our web page... **traces** Base de dades bibliogràfiques de Llengua i Literatura Catalanes

Search 85,229 records for: Any field Search Browse



COLLABORATE

Obra Social "la Caixa"

Generalitat de Catalunya Institut de les Lletres Catalanes

SEE ALSO

TRACES UN POBLER LITERATURA CATALANA

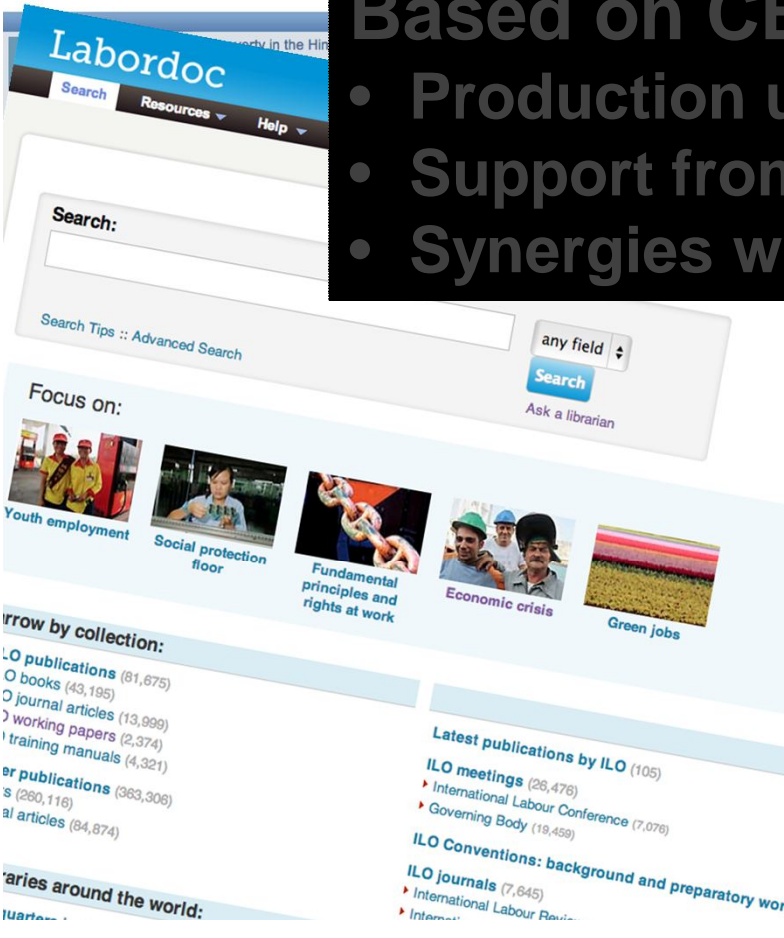
gelo

Universitat Autònoma de Barcelona

Josep M. Bayesa Athilde Pernaü (21) Manuel (72) Alexandre

Based on CERN's Invenio technology

- Production use in many different contexts
- Support from CERN's core team
- Synergies with OpenAIRE+



Labordoc

Search Resources Help

Search:

Search Tips :: Advanced Search any field Search Ask a librarian

Focus on:

- Youth employment
- Social protection floor
- Fundamental principles and rights at work
- Economic crisis
- Green jobs

Narrow by collection:

- ILO publications (81,675)
- Books (43,195)
- Journal articles (13,999)
- Working papers (2,374)
- Training manuals (4,321)
- Other publications (363,306)
- Articles (260,116)
- Journal articles (84,874)

Latest publications by ILO (105)

- ILO meetings (26,476)
- International Labour Conference (7,076)
- Governing Body (19,459)
- ILO Conventions: background and preparatory work
- ILO journals (7,645)
- International Labour Review

Articles around the world:



UAB Universitat Autònoma de Barcelona

dipòsit digital de documents de la UAB

Servei de Biblioteques Sobre el DDD Help Català English Español

Home

Muere Jordi Castellanos Vila, catedràtic del Departament de Filologia Catalana de la UAB.

Castellanos va llicenciar en Filologia Romànica en la UB i es doctorà en filologia catalana en 1970 amb la tesi Raimon Casellas i el Modernisme, dirigida per Joaquim Molas. Impartí classes de llengua espanyola i catalán en la Universitat de Durham (Reino Unido) y, desde 1988, era catedrático de literatura catalana contemporánea de la UAB. Trabajó con la Gran Enciclopedia Catalana y estuvo vinculado a las revistas Els Marges, Recerques, El Pou de les Lletres y Hemenueus, además de colaborar con La Vanguardia, TeleXpress, Avui, El Punt y El País. Era miembro del Institut d'Estudis Catalans, de la Associació Internacional de Llengua i Literatura Catalanes y de la Associació d'Escriptors en Llengua Catalana. Por otra parte, su compromiso político lo vinculó a diferentes organizaciones: al Sindicat Democràtic d'Estudiants, al Front Nacional de Catalunya y al Partit Socialista d'Alliberament Nacional (PSAN). You can find in DDD 46 documents in full text

Search 88,131 records for: any field SEARCH BROWSE

Search Tips :: Advanced Search

Narrow by collection:

- Course materials (36,174)
- Exams (5,502) Student guides (18) Course programs (Study plans) (30,572) Teaching resources (82)
- Books and collections (2,115)
- Books (1,849) Book chapters (89) Institutional documents (51) UAB Libraries guides (104) Radio scripts (22)
- Research literature (6,899)
- Studies (541) Patents (73) Preprints (450) Doctoral theses (2,902) Dissertations (1,958) Working papers (975)
- Periodical publications (1,788)
- Annual reports (547) Newsletters (1,039) Journals (202)
- Articles (24,777)

Focus on:

- UAB course materials (36,174)
- Exams (5,502) Student guides (18) Course programs (Study plans) (30,572) Teaching resources (82)
- UAB Books and series (293)
- UAB Books (49) UAB chapters of books (89) UAB Institutional documents (51) UAB Libraries guides (104)
- UAB research literature (6,818)
- UAB studies (541) UAB Preprints (450) UAB doctoral theses (2,894) UAB dissertations (1,958) UAB working papers (975)
- UAB periodical publications (149)



Invenio features

- navigable collection tree (regular, virtual)
- powerful search engine
 - combined metadata, reference and fulltext search
- flexible metadata (MARC, OA)
- handling any kind of document (multimedia)
- customizable input, formatting and linking
- personalization and collaborative features:
 - alerts, baskets, groups, reviews, comments
- internationalisation (28 languages)

<http://simko.home.cern.ch/simko/talks/invenio-technology-openlab-aug-2011.pdf>



B2Share deposit flow



Upload

- Select one or more files
- Click to start uploading



Domain

- Pick a scientific domain of submission
- Modifies the metadata enrichment form



Metadata

- Domain dependent
- Configurable to simplify adding new domains



B2Share security model

- All metadata is public
- Data can be marked as private – in this case accessing the data files will result in a 403 code
- More elaborate models will follow
 - E.g. sharing only among collaborators



Metadata Model

- All submission have a generic metadata associated with them
 - Title, description, tags, etc
 - DataCite-compliant
- Domain specific metadata is requested only if a specific domain is selected
 - Areas of science: e.g. ecology, linguistics
 - Projects: e.g. DRIHM, CLARIN
- Input fields + validation are automatically generated from the metadata description



Integration with EUDAT services (existing)

- EPIC PID
 - On deposit EPIC service is contacted to register a UUID
 - UUID is embedded into the metadata record of the entry
- Metadata catalogue
 - B2Share exposes metadata via OAI-PMH protocol
 - Metadata catalogue consumes it and maps to its standard model



B2Share vs Zenodo

- Are based on the same framework (Invenio)
- B2Share is more 'lightweight'
 - Leverages on the EUDAT services
- Zenodo targets 'high-quality' data, i.e. publications and associated data
- B2Share aims at offering a more generic submission mechanism for multiple scientific domains + EUDAT service integration



Technology

- Basis for the implementation: Invenio 1
 - Library software
 - Mostly CERN-based development
- A separate B2Share overlay
 - Installs on top of a working Invenio instance
 - Brings models, views and deposit logic
 - flask + jinja2 + sqlalchemy
 - <https://github.com/SimpleStore/simplestore>

Diapositiva 14

1

Ilja Livenson; 21/10/2012



Try out

- B2Share instance for trying out

<https://ss-dev.pdc.kth.se>

- Central instance will be available at:

<https://b2share.eudat.eu>



Adding a new domain

- Generic domain of submission is often not enough
- B2Share supports easy way for enabling additional scientific domains
- Hopefully engaged by a community



Defining metadata

- Define a domain description based on a template
- Add domain-specific fields

```
domain = "Cats"  
table_name = 'cats_imgs'  
icon = 'icon-cat'  
fields = [  
    {  
        'name': 'funlevel',  
        'display_text': "Fun level",  
        'col_type': db.String(256),  
        'mandatory': True,  
        'extra': False,  
        'description': 'Increase in the happiness  
                        level caused by the image'  
    }  
]
```



Enabling the domain

```
$ # check enabled domains
```

```
$ grep CFG_SIMPLESTORE_DOMAINS invenio-local.conf
```

```
CFG_SIMPLESTORE_DOMAINS = generic, drihm, eudat,  
ecology
```

```
$ # append 'cats' to the list
```

```
$ sudo -u www-data /opt/invenio/bin/ \ inveniocfg --update-  
all
```



Future Work

- Integration with further EUDAT services
 - Ingestion into Safe replication
- Increased security/privacy model
- Clean API for batch uploads



THANK YOU!

ON BEHALF OF B2SHARE CONTRIBUTORS

ADRIAN MOUAT

HUW MORRIS

ILJA LIVENOSN

livenson@kth.se



Feedback

- We'd like to use this session also for getting immediate feedback 😊
1. Do you think that B2Share makes sense for your needs?
 2. What would need to be changed to make the service more attractive?