



A Modern Digital Library System for Grey Literature

Jérôme Caffaro
Samuele Kaplun

<http://invenio-software.org>

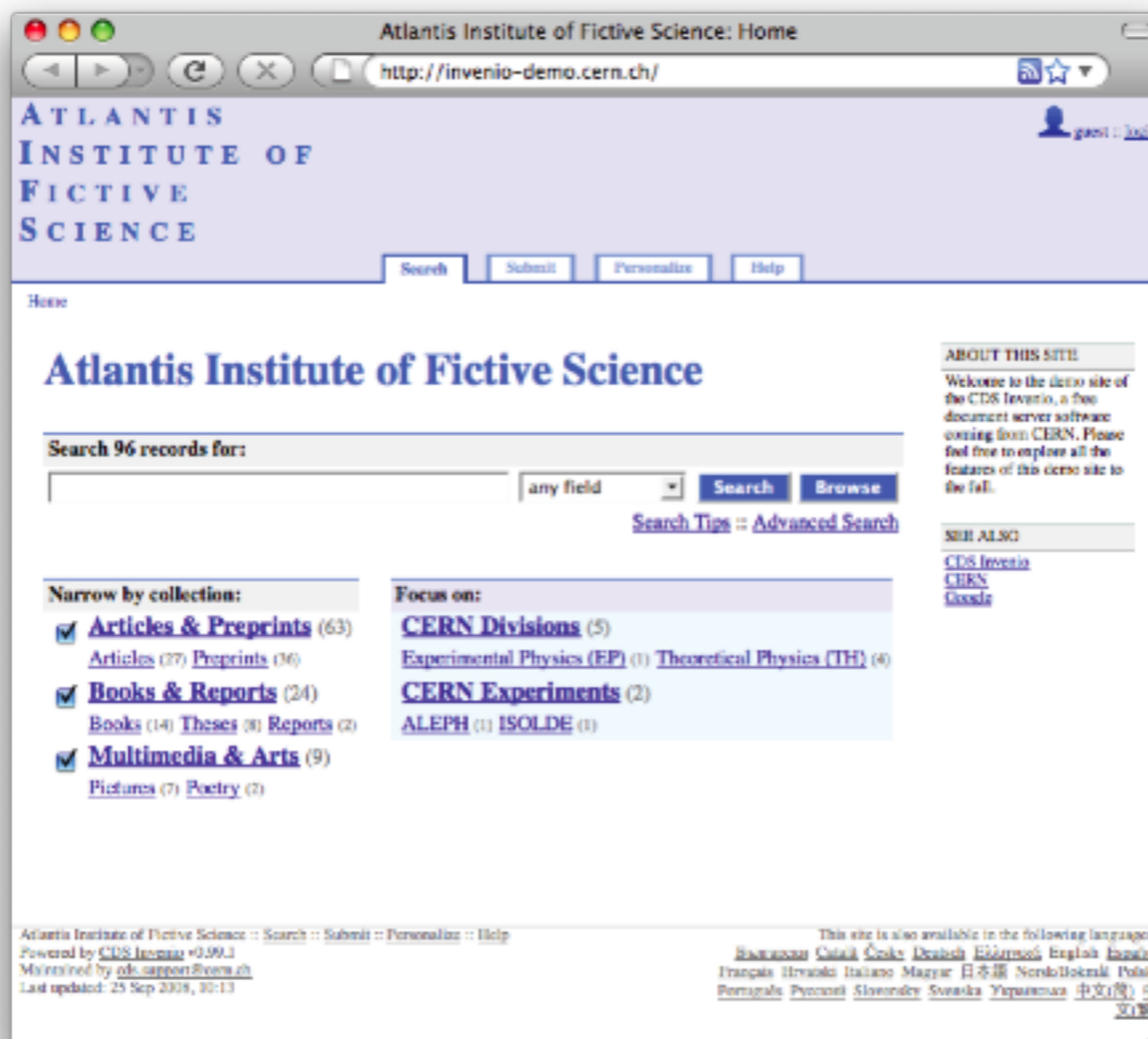


What is Invenio?



User

approve loan items
 share
 "bookmark"
 submit
 search
 comment, review
 subscribe



send overdue letters
 "check"
 edit



Librarian

import, harvest
 merge
 convert
 extract keywords
 export
 monitor
 script
 customize



Admin

▶ Free, open-source (GNU GPL) repository software





Context

- 1954** CERN (European Center for Nuclear Research) is born. 12 initial member states (20 + 15 observers today; About 10'000 persons; > 500 universities)
- 1989** Tim Berners-Lee invents the Web at CERN.
- 1991** SPIRES (SLAC) first database available on the web.
- 1991** arXiv repository for preprints in physics.
- 1993** CERN Preprint Server on the Web: institutional repository.
- 1996** CERN Library Server (weblib): added books, periodicals and "other material".



Context

2000 CERN Document Server: multimedia material, internal notes)

2002 First release (CDSware). Worldwide installations. International collaborations.

2006 changed name to CDS Invenio

2007 started collaboration with SPIRES

2008-2010 many new features, optimizations, change of development model and code stabilization for Invenio 1.0

started collaboration with NASA-ADS

- Grey literature is a major channel of communication in the High-Energy physics community^[1]
- CERN is the world's largest particle physics laboratory, and has developed Invenio as a mean to archive produced material and make it accessible to the largest possible audience

[1] Anne-Gentil Beccot, How do High Energy Physics scholars search their information?, Grey J. 4, 1 (2008)



Feature-Full Digital Library



Librarian & Admin tools



Collaborative tools



Multi-language



Fast search + Browse

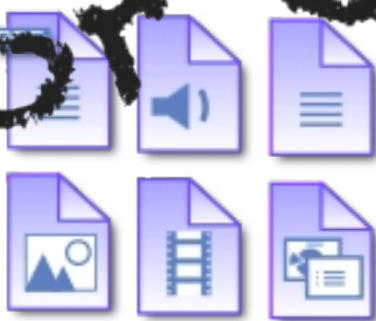


Alerts (emails, RSS)



Commenting + reviewing

For Grey Literature too!



Multiple files formats support



Circulation

OAI-PMH MARCXML
OpenURL
XSLT EndNote
BibTeX OpenSearch

Standards support

Web journals,
OCR,
and more...

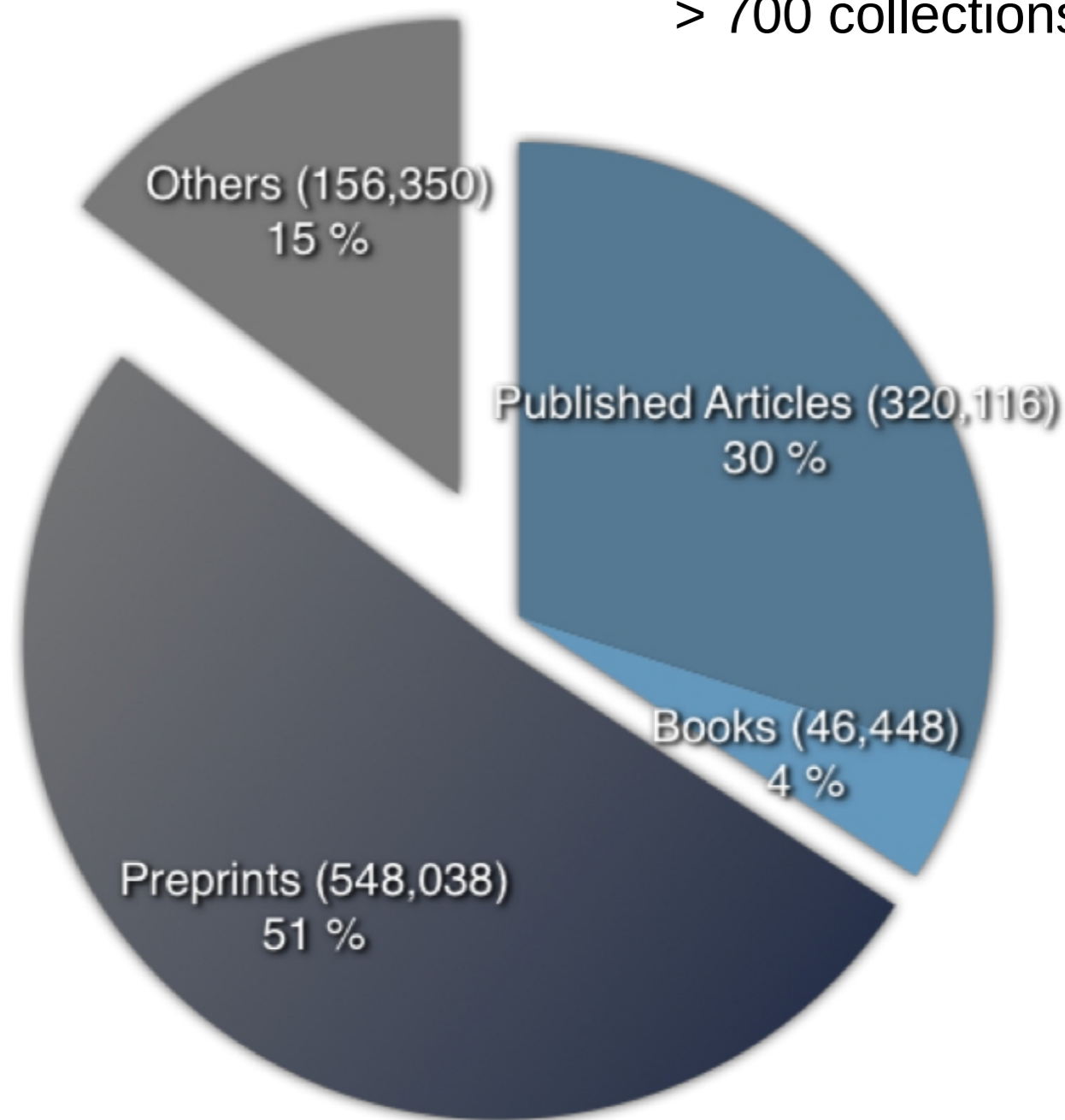
A case study:

CERN Document Server

(powered by Invenio)

- Reports, notes
- Slides, posters
- Photos, videos (movies, rushes, lectures, etc.) & audio archives
- Museum objects
- Agreements, contracts
- etc.

> 1'000'000 records
> 700 collections



Data: 2nd november 2010

Invenio for grey literature

- **MARCXML as core metadata format**
 - Flexible to adapt to about any material
 - Understood and loved by our librarians
 - Same powerful cataloguing tool to curate grey and non-grey material
- **Flexible metadata-formatting layer**
 - Map your material to any HTML/XML representation

- **Customizable workflows engine**
 - Handle complex workflows
 - Living documents are immediately searchable and made available in their dedicated collection
 - Support for authors or community-based revision of files and metadata
 - Possibility to integrate with commenting system

Invenio for grey literature

- **Powerful search engine**
 - Support large repositories
 - "Google-like" search speed
 - Simple and advanced search syntax.
- **Collaborative tools**
 - Commenting and reviewing system
 - Notifications (email, RSS) of new material or new comments
 - Support for groups and sharing of material

- **Interoperability**
 - Harvest and disseminate using OAI-PMH
 - Multiple/customizable exports
 - SWORD export (coming soon)
- **Integrated digital library**
 - One system: IT infrastructure simplified; easier maintenance
 - Does not cost more to actively support grey material in addition to other library material

- Invenio is designed from the ground for a community heavily relying on grey literature.
- Invenio is an established, open-source software for diverse medium to large sized repositories.