CERN Open Source Collaborative tools: Digital Library Software

Tim Smith CERN/IT

Libraries…

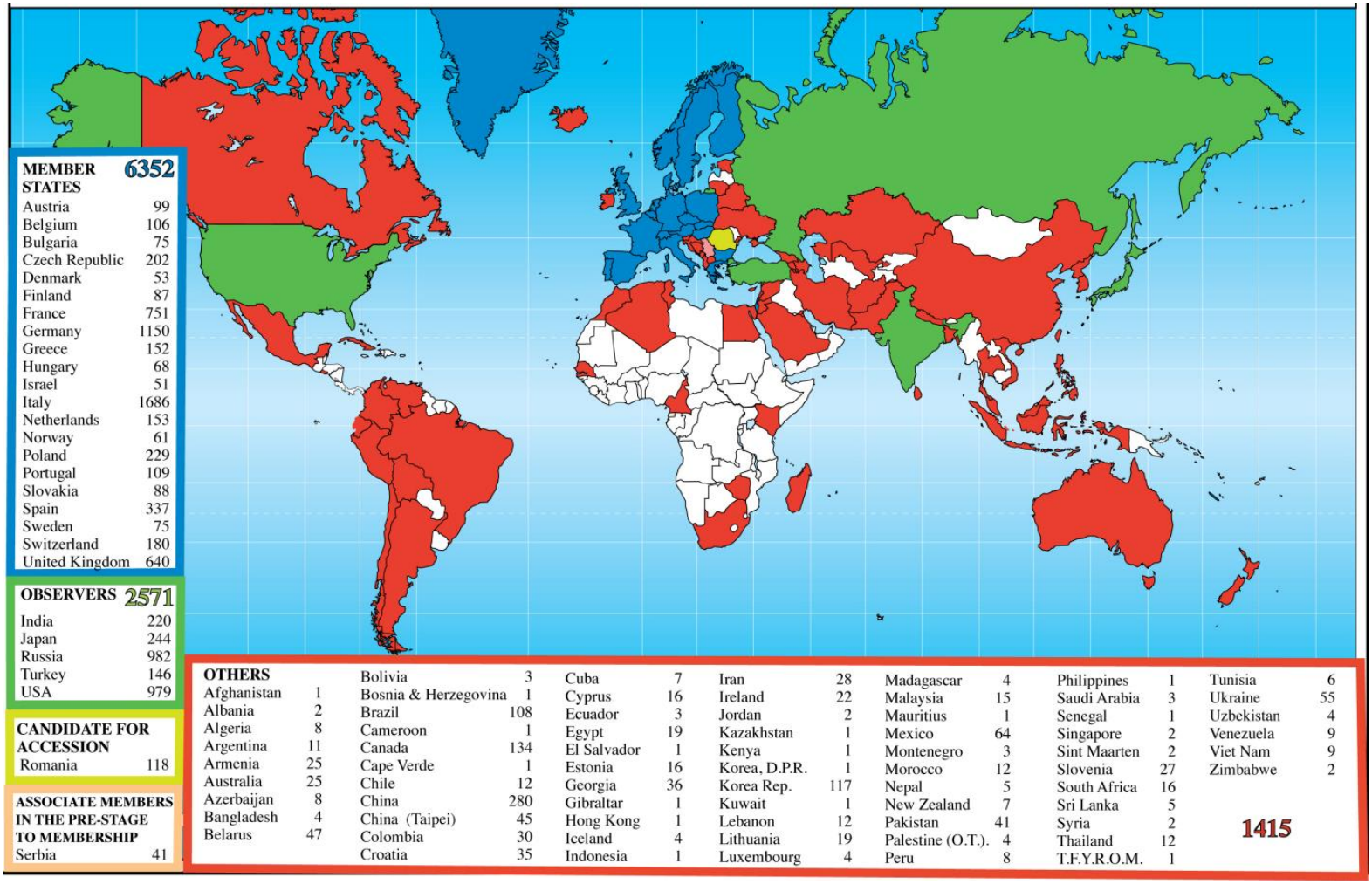

Quantum Electronics
Quantenphysik
Mehra
Mecanica Quantica
quantum mechanics
Miller
Quantum Mechanics for Scientists and Engineers
Relativistic Quantum

A Visionary Perspective

Preprint Culture

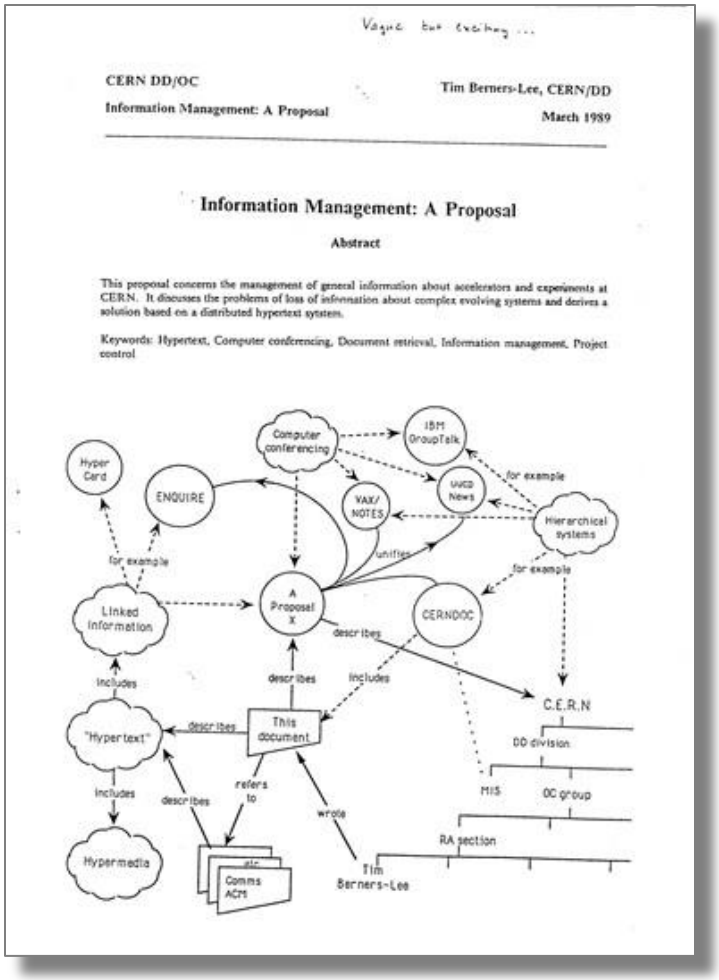
Dissemination

CERN Users around the World



10,000 scientists and engineers, 98 countries

Dawn of Internet Age



Vague but exciting ...

CERN DD/OC

Tim Berners-Lee, CERN/DD

Information Management: A Proposal

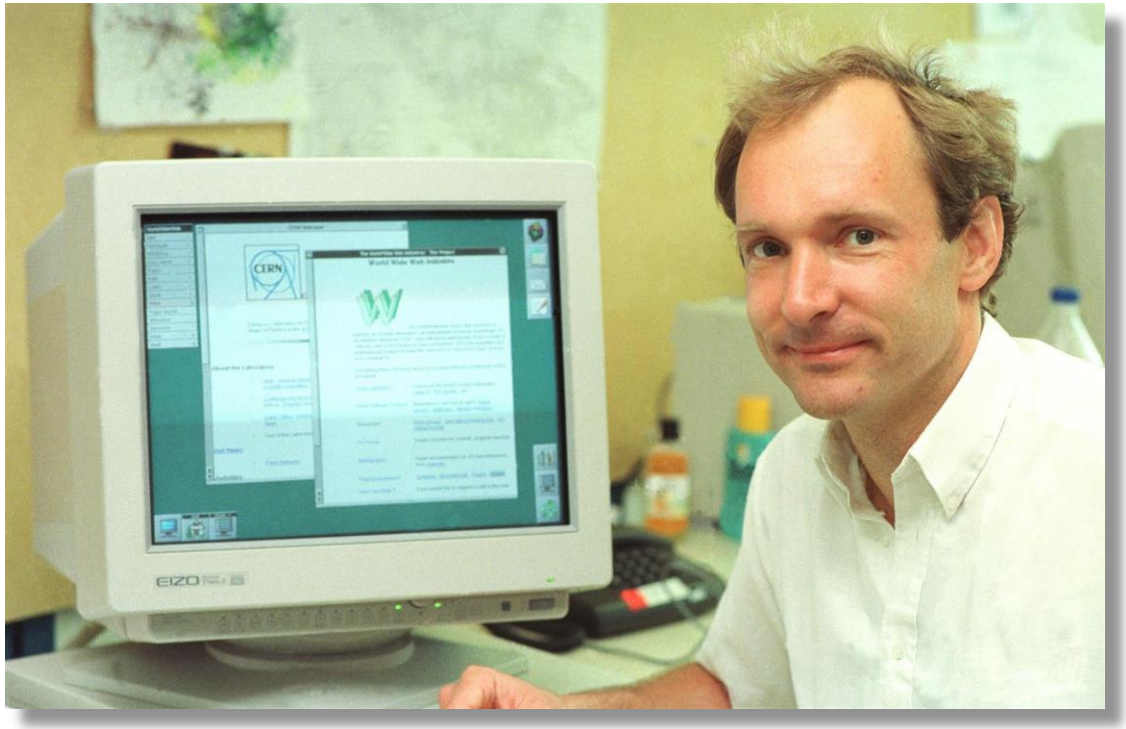
March 1989

Information Management: A Proposal

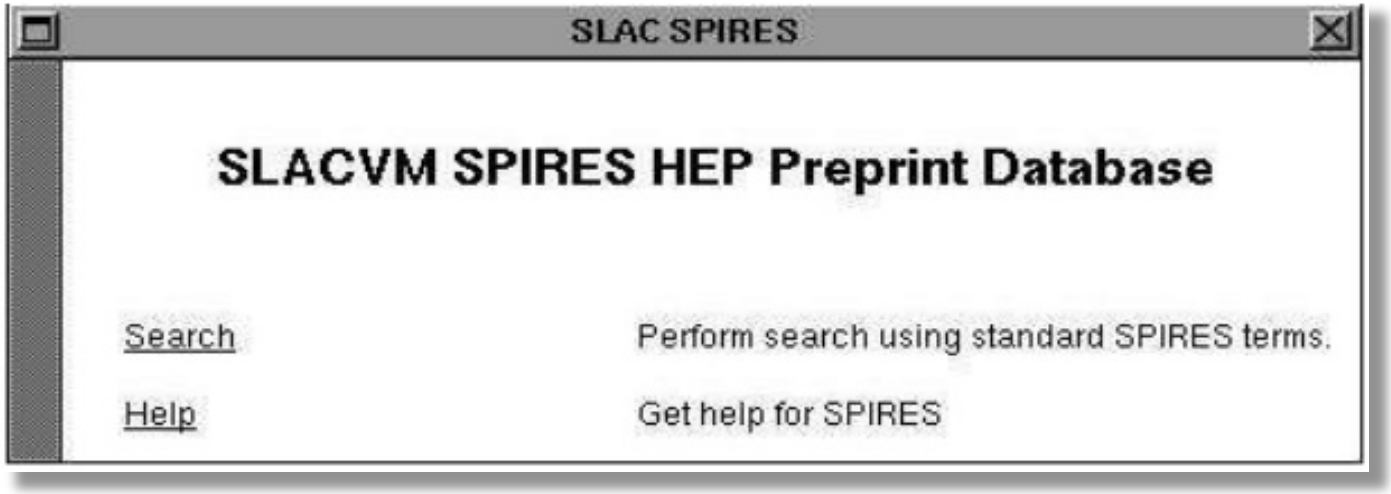
Abstract

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control

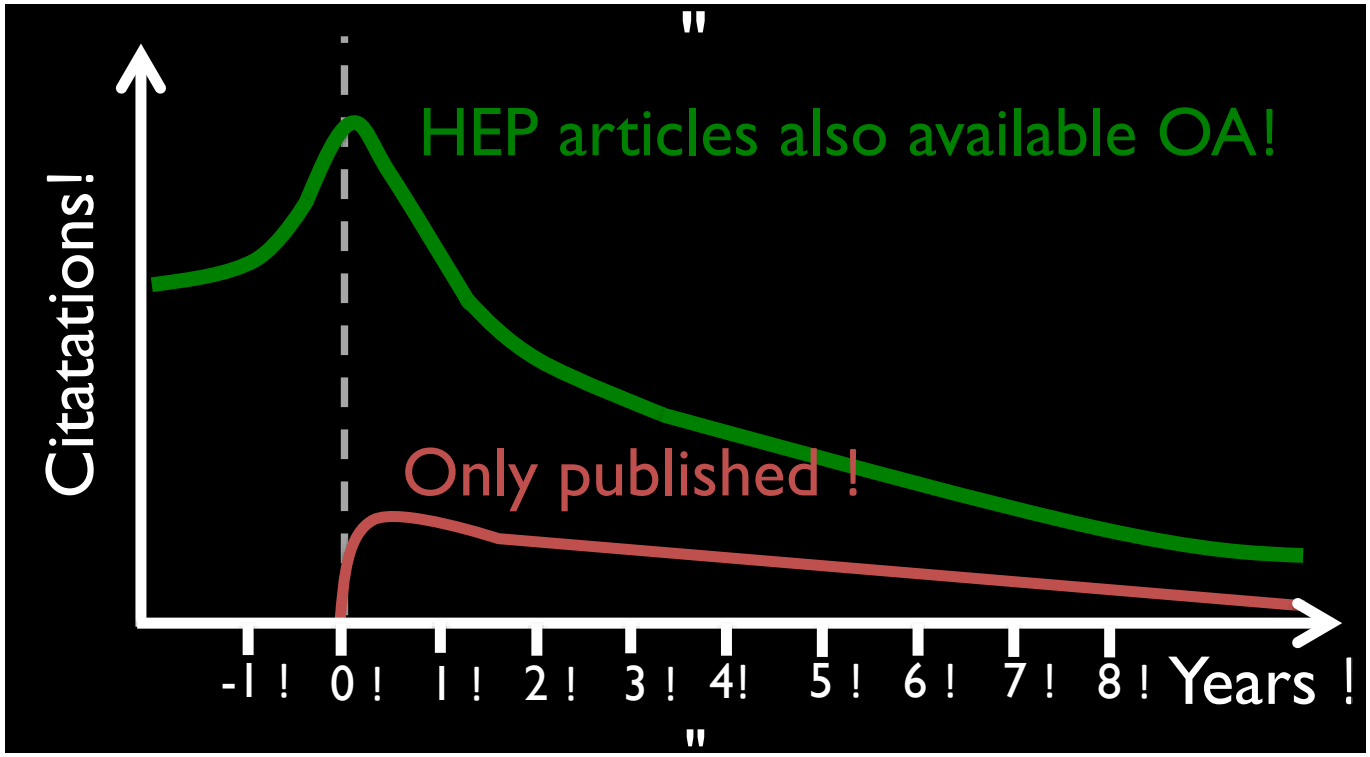


SPIRES: first web site in the USA



And the first DataBase on the web

Accelerating Science



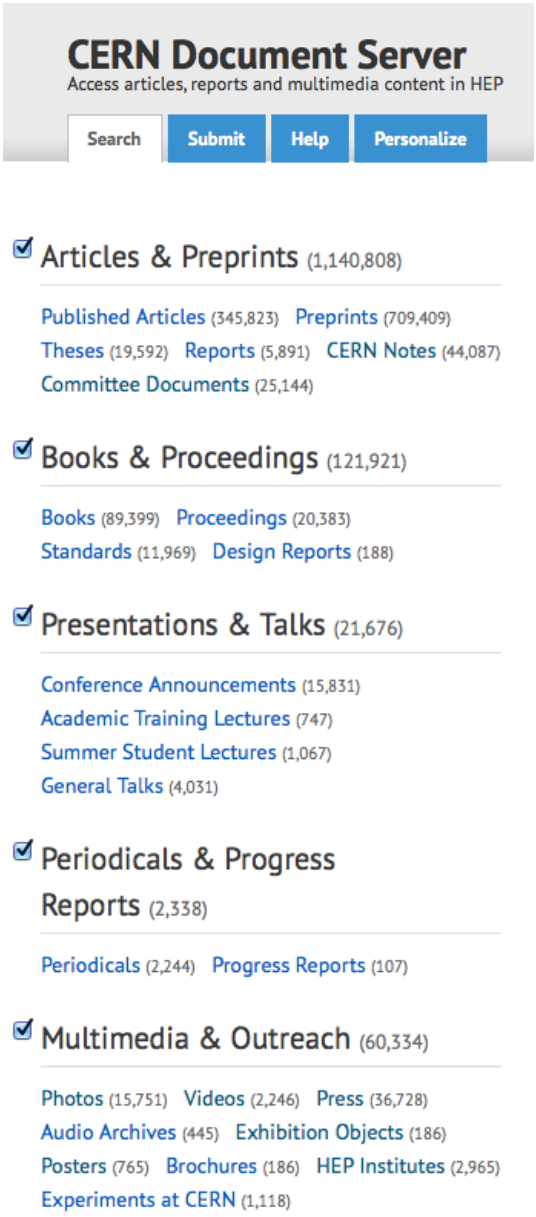
Scientific dialogue on repositories

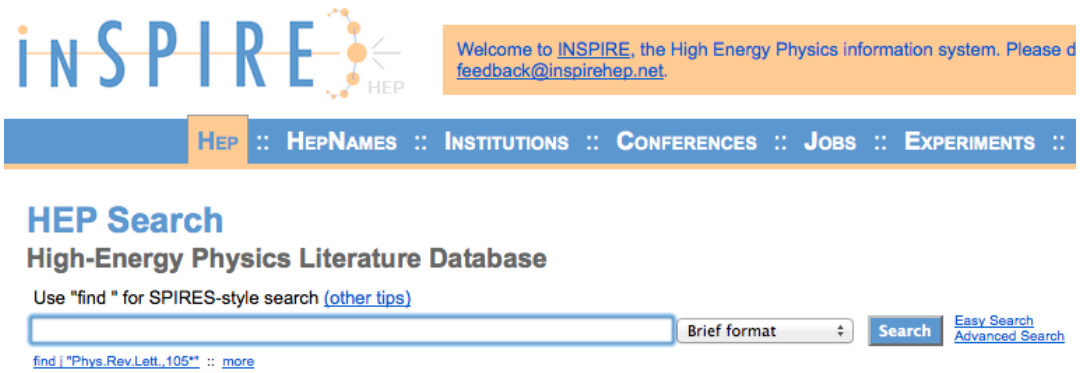
Gentil-Beccot, Mele, Brooks arXiv:0906.5418

Towards Digital Libraries

- 1993:
  - CERN Preprint Server serves HEP & CERN preprints
- 1996:
  - CERN Library Server provides access to Library Catalog
- 2000:
  - CERN Document Server includes multimedia, restricted notes
- 2002:
  - CDSWare SW is released open source
- 2006:
  - CDSWare becomes Invenio; start of I18N collaborations
- 2010:
  - Invenio 1.0 released and adopted world-wide

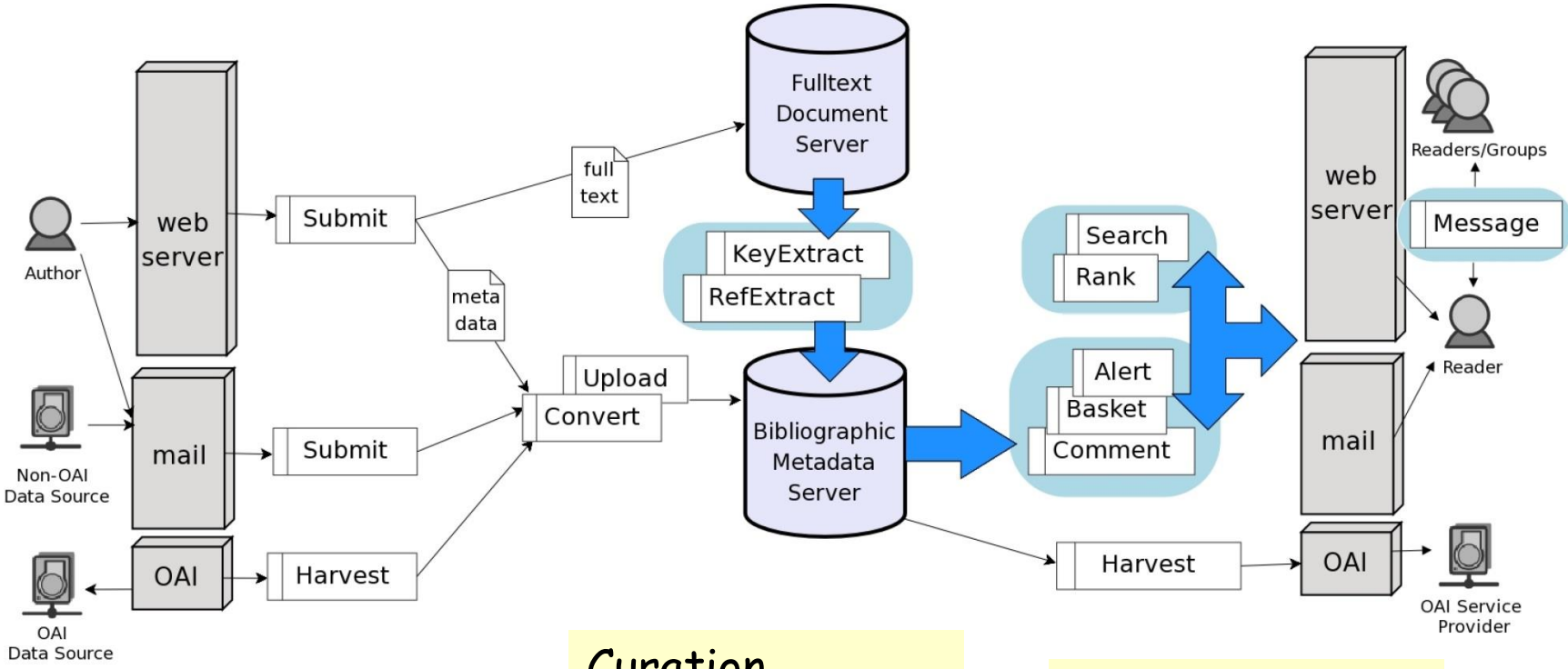
“One Stop Shop”





> 1 million records

Digital Library Services



Collection
Aggregation
Conversion
Stamping
Watermarking

Curation
Cataloguing
Organisation
Enrichment
Preservation

Access
Indexing
Ranking
Clustering
Classifying

Plot Extraction

The ridge in proton-proton collisi...

Information | References (41) | Citations (3) | Plots

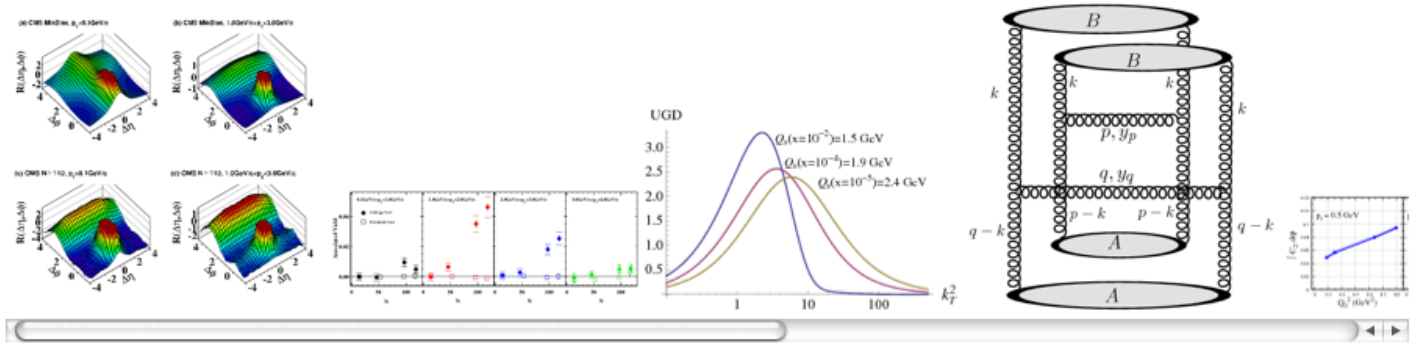
The ridge in proton-proton collisions at the LHC.

Adrian Dumitru, Kevin Dusling, Francois Gelis, Jamal Jalilian-Marian, Tuomas Lappi, Raju Venugopalan.

Sep 2010
e-Print: **arXiv:1009.5295 [hep-ph]**

Abstract: We show that the key features of the CMS result on the ridge correlation seen for high multiplicity events in sqrt(s)=7TeV proton-proton collisions at the LHC can be understood in the Color Glass Condensate framework of high energy QCD. The same formalism underlies the explanation of the ridge events seen in A+A collisions at RHIC, albeit it is likely that flow effects may enhance the magnitude of the signal in the latter.

Keyword(s): INSPIRE: CERN LHC Coll | p p: scattering | quantum chromodynamics | color glass condensate | Brookhaven RHIC Coll | flow | CMS
Note: * Temporary entry *



Record created 2010-10-01, last modified 2010-10-10

Similar records

Visualizing Patterns of Connection

A Model of Leptons – HEP

http://inspirebeta.net/record/51188/citations

Google

A Model of Leptons – HEP

Cited by: 7265 records

(6121) **CP Violation in the Renormalizable Theory of Weak Interaction** - Kobayashi, Makoto *et al.* Prog.Theor.Phys. 49 (1973) 652-657 . KUNS-242

(3641) **Supersymmetry, Supergravity and Particle Physics** - Nilles, Hans Peter Phys.Rept. 110 (1984) 1-162 . UGVA-DPT-1983-12-412

(3529) **The Search for Supersymmetry: Probing Physics Beyond the Standard Model** - Haber, Howard E. *et al.* Phys.Rept. 117 (1985) 75-263 . UM-HE-TH-83-17, SCIPP-85-47

(3255) **Unity of All Elementary Particle Forces** - Georgi, H. *et al.* Phys.Rev.Lett. 32 (1974) 438-441

(3196) **Review of particle physics. Particle Data Group** - Particle Data Group Collaboration (Barnett, R.Michael *et al.*) Phys.Rev. D54 (1996) 1-720

more

Co-cited with: 61796 records

(3762) **Partial Symmetries of Weak Interactions** - Glashow, S.L. Nucl.Phys. 22 (1961) 579-588

(2110) **Weak Interactions with Lepton-Hadron Symmetry** - Glashow, S.L. *et al.* Phys.Rev. D2 (1970) 1285-1292

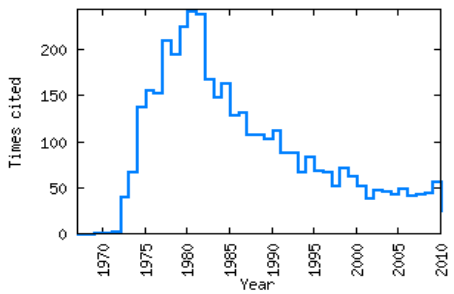
(1464) **CP Violation in the Renormalizable Theory of Weak Interaction** - Kobayashi, Makoto *et al.* Prog.Theor.Phys. 49 (1973) 652-657 . KUNS-242

(1070) **Broken symmetries, massless particles and gauge fields** - Higgs, Peter W. Phys.Lett. 12 (1964) 132-133

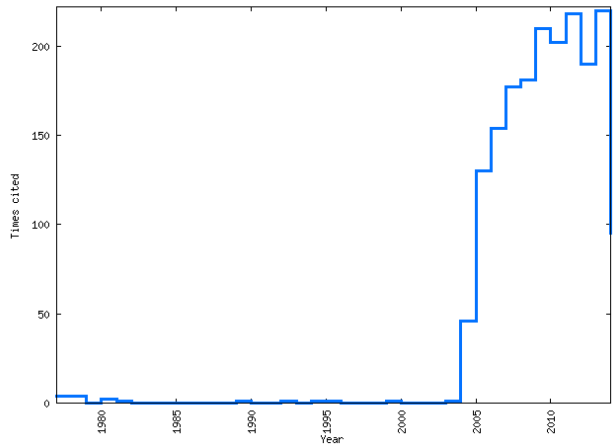
(876) **Unity of All Elementary Particle Forces** - Georgi, H. *et al.* Phys.Rev.Lett. 32 (1974) 438-441

more

Citation history:

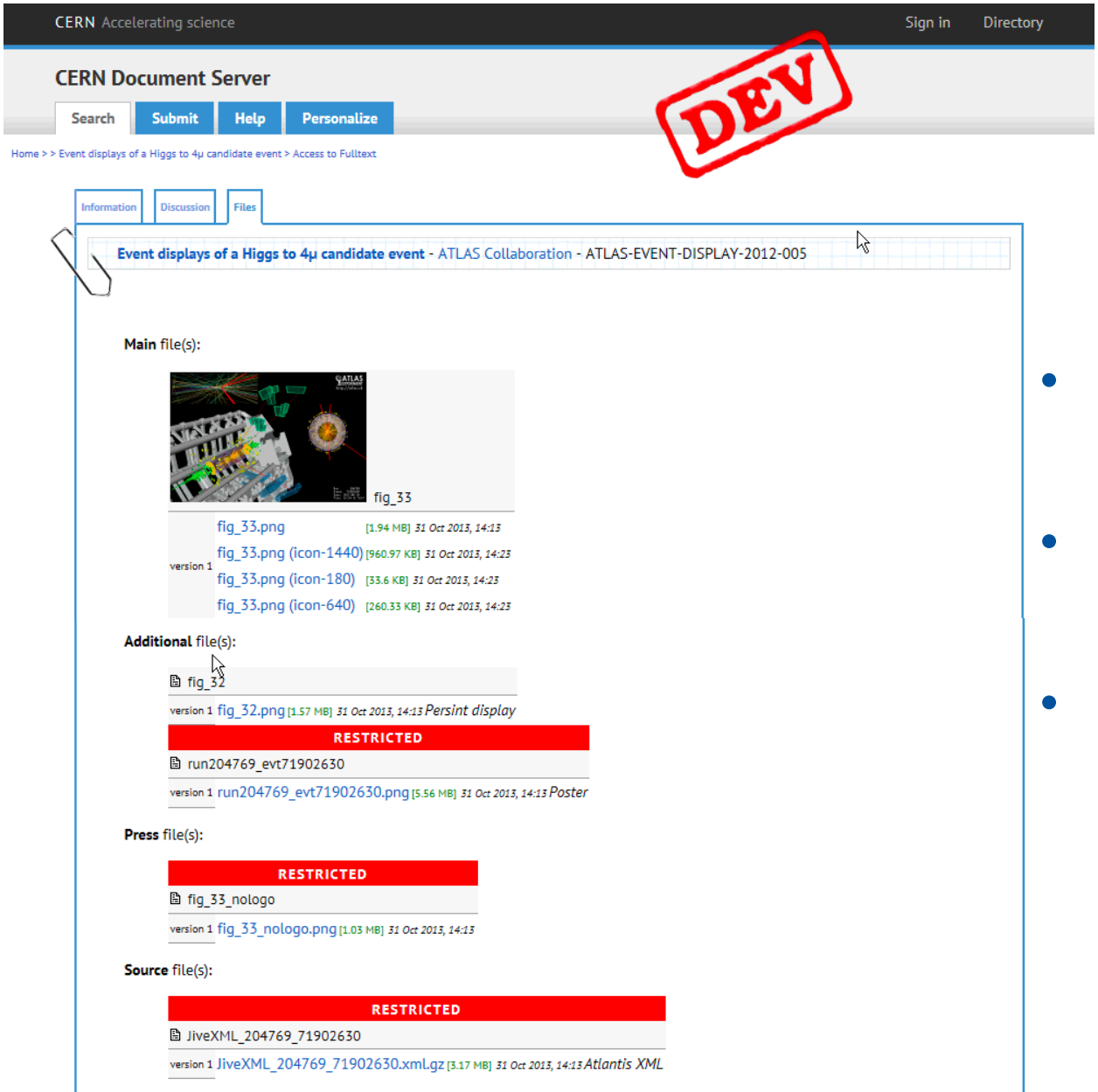


Citation history:



Done

Open and Closed Data !

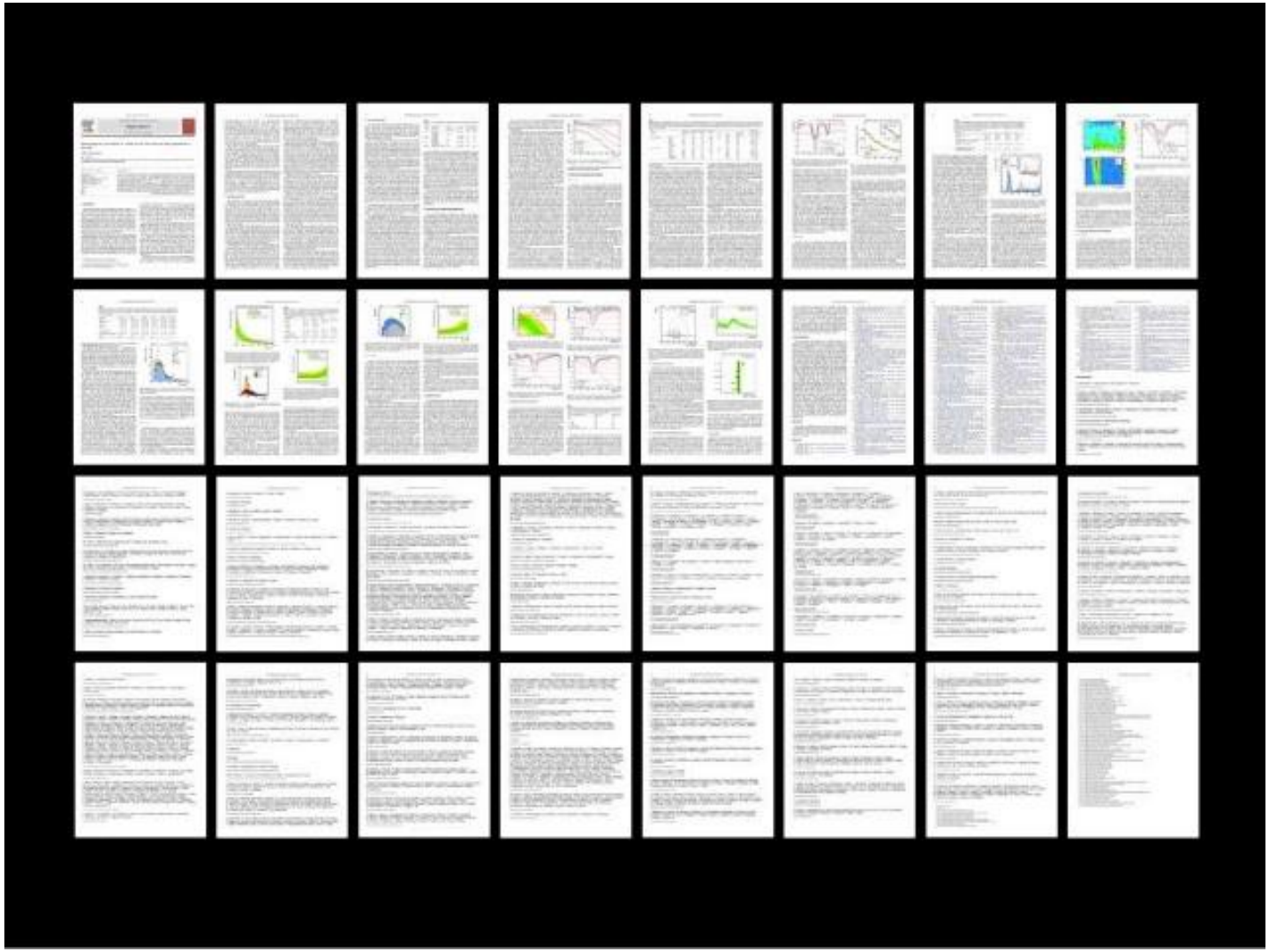


- Workflows
- Transformations
- Restrictions

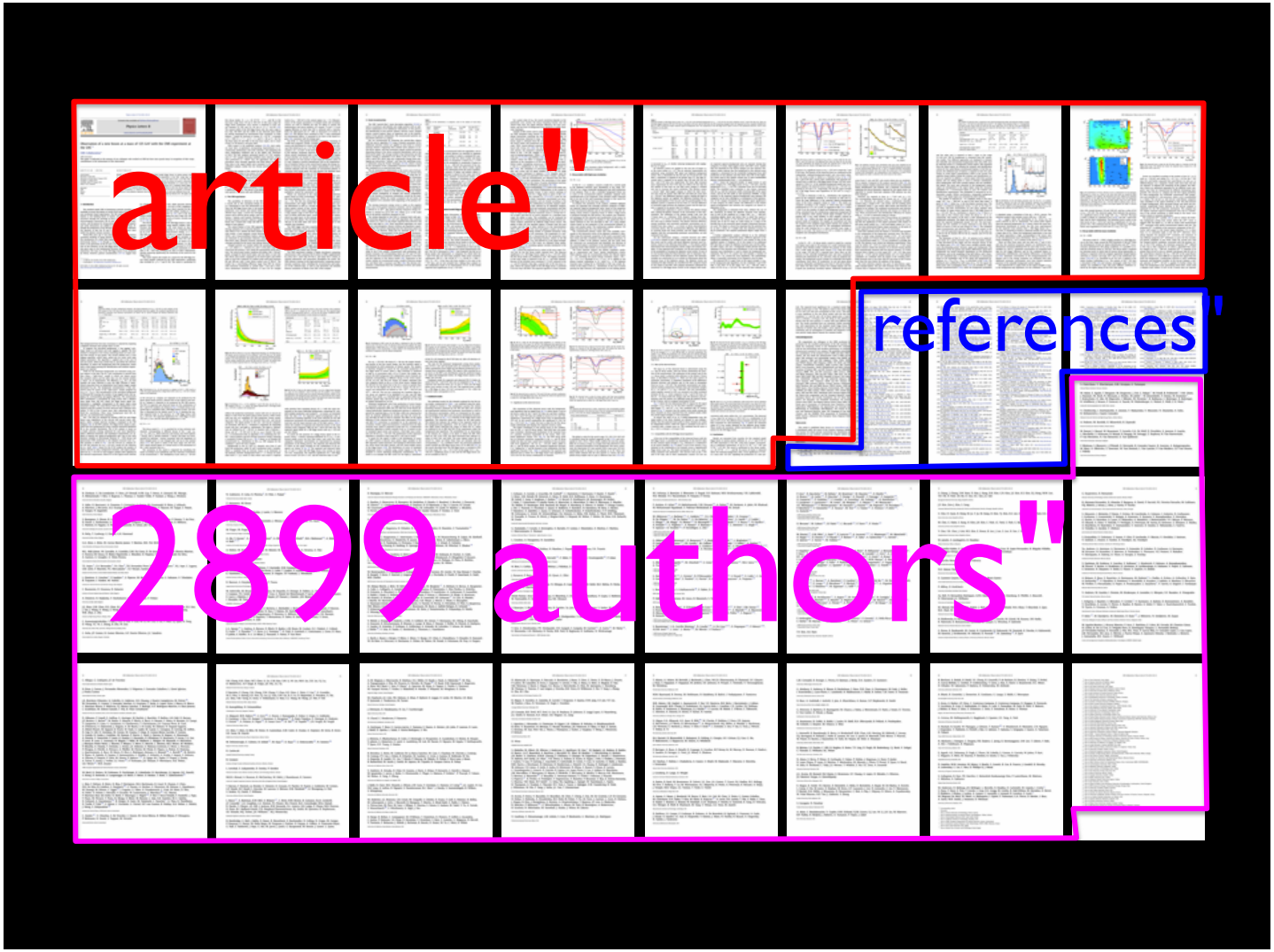
Digital Age Services

- Collaboration “Web2.0”
  - Comments, reviews, baskets
- Immediacy
  - Email alerts, RSS feeds
- Intensive tasks
  - Keyword & reference extraction
  - Citation analysis
  - Full text indexing & ranking
  - Conversion services: multiple download formats
- Flexible formats
  - Remove constraints of print versions
  - Internationalisation

Authors



Authors



Author Disambiguation

Ellis, Jonathan Richard

View Profile | Manage Profile | Manage Publications | Search Profiles | Help

Publications list

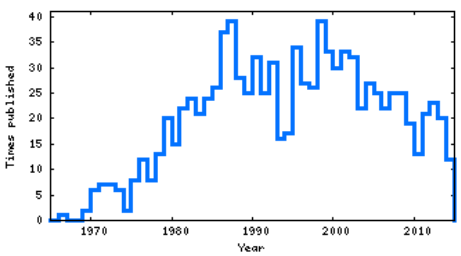
Publications in INSPIRE:

1. A spinor approach to quaternion methods in relativity
2. D-Flation
3. The Physics Programme Of The MoEDAL Experiment At The LHC
4. Astrophysical Shrapnel: Discriminating Among Extra-solar Sources of Live Radioactive Isotopes
5. A No-Scale Inflationary Model to Fit Them All
6. The Extent of the Stop Coannihilation Strip
7. Generalized Skyrmions and The Mass of the Lightest Electroweak Baryon*
8. Closing in on the Tip of the CMSSM Stau Coannihilation Strip
9. Complete Higgs Sector Constraints on Dimension-6 Operators
10. Resurrecting Quadratic Inflation in No-Scale Supergravity in Light of BICEP2
11. Starobinsky-Like Inflation in Dilaton-Brane Cosmology
12. The CTA Sensitivity to Lorentz-Violating Effects on the Gamma-Ray Horizon
13. Exploring Two-Field Inflation in the Wess-Zumino Model

Papers

	All papers	Single authored
All papers	982	239
Book	2	0
ConferencePaper	240	161
Introductory	21	19
Lectures	46	36
Published	643	55
Review	105	72
Thesis	1	1
Proceedings	16	0

Publications per year



Subject categories

Frequent keywords

Citations (from papers in INSPIRE)

Citations summary

Generated on 2014-06-13

982 papers found, 806 of them citeable (published or arXiv)

Citation summary results	Citeable papers	Published only
Total number of papers analyzed:	806	643
Total number of citations:	60,506	57,364
Average citations per paper:	75.1	89.2
Breakdown of papers by citations:		
Renowned papers (500+)	12	12
Famous papers (250-499)	43	41
Very well-known papers (100-249)	116	111
Well-known papers (50-99)	135	128
Known papers (10-49)	278	238
Less known papers (1-9)	172	99
Unknown papers (0)	50	14
h_{HEP} index [?]	126	124

See additional metrics

Exclude self-citations or RPP

Warning: The citation search should be used and interpreted with great care. Read the fine print

HepNames data

John Richard Ellis (King's Coll. London)
[Author Profile] [Google] [Students] [arXiv] [ADS]

PhD Institution: Cambridge U.
Undergrad: Cambridge U.
Email: john.ellis@cern.ch
URL: http://www.kcl.ac.uk/news/news_details.php?news_id=1324&year=2010
Field: HEP-TH, HEP-PH, ASTRO-PH
Author Profile: J.R.Ellis.1
Inspire ID: INSPIRE-00146525

Institutional History:

Institution	Rank	Start Date	End Date
King's Coll. London	SENIOR	2010	
CERN	SENIOR	1973	2010
Caltech	PD		
SLAC	PD		
Cambridge U.	PHD	1967	1971
Cambridge U.	UG	1964	1967

Name variants

Ellis, Jonathan Richard (2)
Ellis, Jonathan R. (2)
Ellis, John.R. (1)
Ellis, John R. (843)
Ellis, John (91)
Ellis, J.R. (16)
Ellis, J. R. (1)
Ellis, J. (25)
Ellis, J (1)

Affiliations

CERN (892)
King's Coll. Lond
SLAC (37)
Caltech (9)
Cambridge U., D
LBL, Berkeley (
Sussex U. (3)
Cambridge U. (
Karlsruhe U., TT
UC, Santa Cruz

The Invenio Platform

- Mature **digital library** platform
  - Articles, books, notes, photos, videos, software, data
  - **OAIS**-inspired **preservation** practices
- Typical use cases:
  - **Institutional** document **repositories**, e.g. CERN, EPFL, GSI
    - Internal collections, pre-publication workflows with approval
  - **Subject-based** information **systems**, e.g. INSPIRE, ILC
    - Public collections, worldwide data with citation analysis
  - **Large libraries** and **library networks**, e.g. ILO, RERO, FZ
- Co-developed by international collaboration

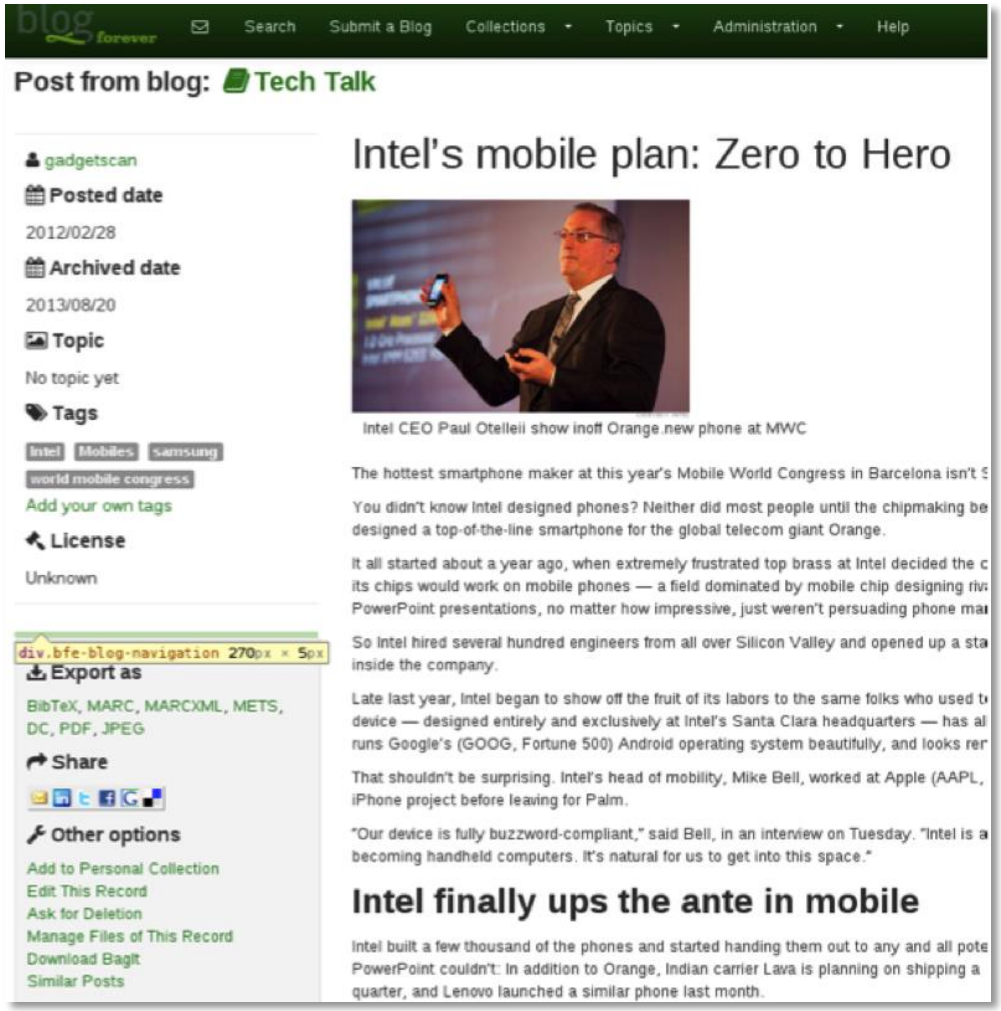
Invenio @ M9

- “Museum of the 20th Century” in Mestre, Italy
- privately funded, Fondazione di Venezia, 2011–2014
- collaborating with CILEA
- using Invenio for repository
  - native support for formats ICCD, MAG, UNIMARC
  - native support of **archival formats** EAD, ISAD(G)
  - **data model abstraction** and **logical fields**
  - **rich API** to expose objects to apps

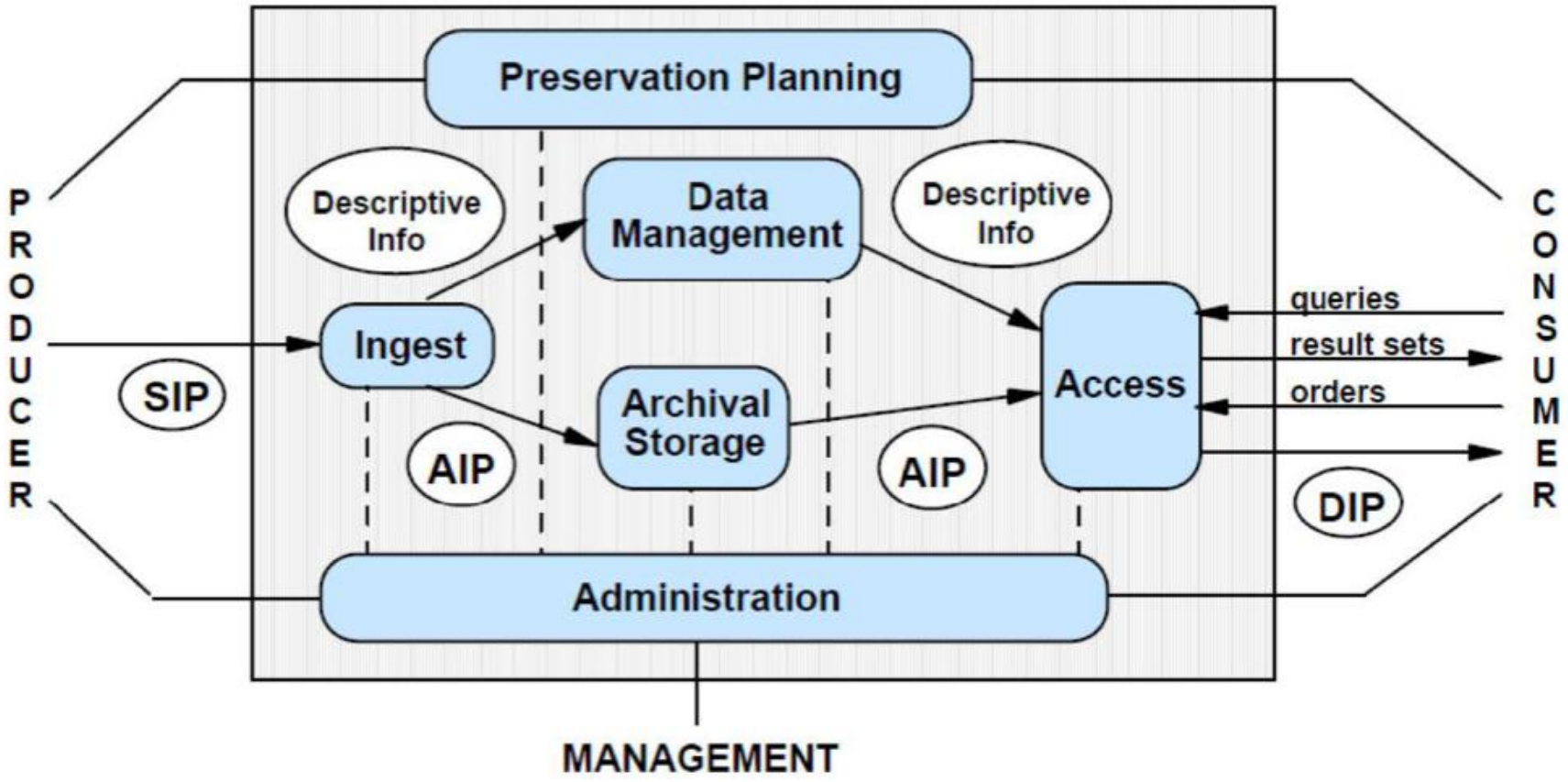
Scientific dialogue 2.0

BlogForever - Preservation

- EC funded project, 2011–2013 (Invenio based)
  - Platform to **harvest**, **manage**, **preserve** and **disseminate** blog content
  - Blog **posts**, **comments**, **embedded material** (images, videos)
  - Ensure **authenticity**, integrity, completeness, **long-term usability**
  - **OAIS AIP**



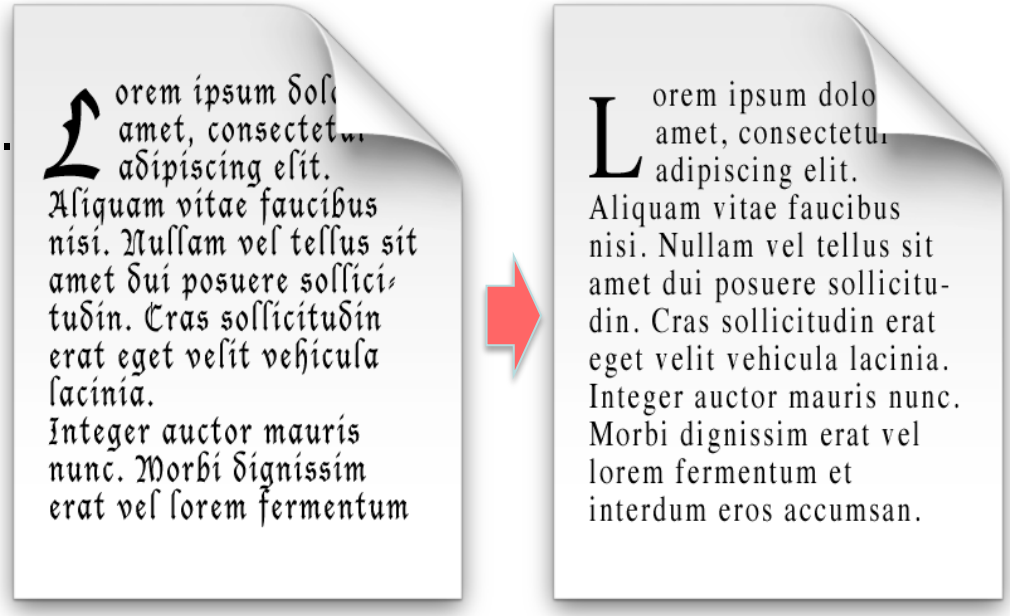
Open Archival Information System



SIP = Submission Information Package · AIP = Archival Information Package · DIP = Dissemination Information Package

Open Access *…always*

- DOI
  - 10.1103/PhysRevLett.105.
- Citation networks
- Format



$$\mathscr{L} \propto H_D \frac{n_\gamma^{3/2}}{\sqrt{\varepsilon_y}} \frac{P_b}{E_b} \propto H_D \frac{n_\gamma^{3/2}}{\sqrt{\sigma_z}} \frac{1}{\sqrt{\varepsilon_y}} \eta \frac{P_{AC}}{E_b}$$

$$\mathscr{L} \square H_D \frac{n_\gamma^{3/2}}{\sqrt{\varepsilon_y}} \frac{P_b}{E_b} \square H_D \frac{n_\gamma^{3/2}}{\sqrt{\sigma_z}} \frac{1}{\sqrt{\varepsilon_y}} \square \frac{P_{AC}}{E_b}$$

- Transformation: PDF/A
- OAIS (ISO 14721:2012)
  - Preservation meta data: provenance, context, usage

Data Intensive Science

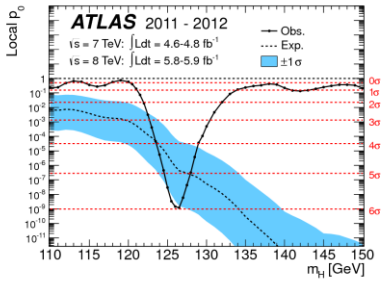
Data Analysis and Preservation

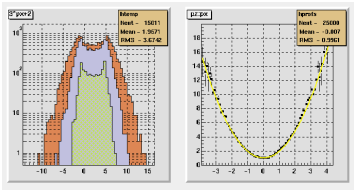
- Papers
- Tabular Data
- Correlation Matrices

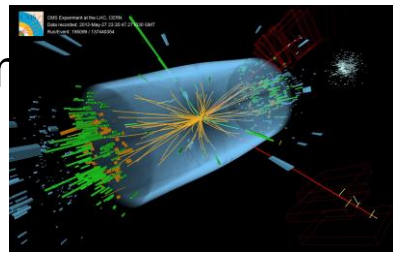
- Internal Notes
- Wikis
- Presentations

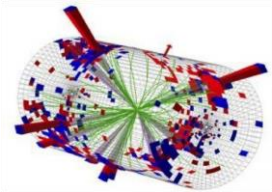
- Quality monitoring data
- Filter / selection algorithm
- Formatters

- Calibration Data
- Conditions Data
- Log Books









Researchers
T2s, T1s

Analysis Coordinators
T1s

Production Managers
T0, T1s

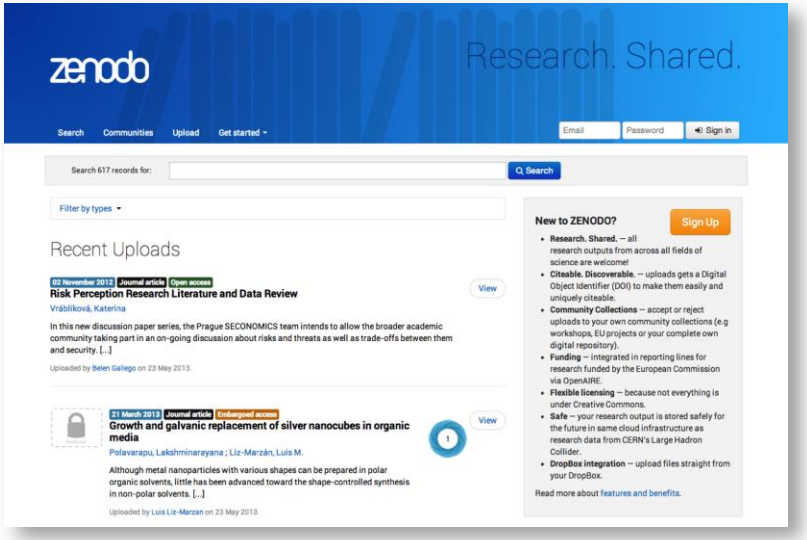
SW: 10M LoC

Contextual metadata

Workflows

Big Data … *in small pieces*

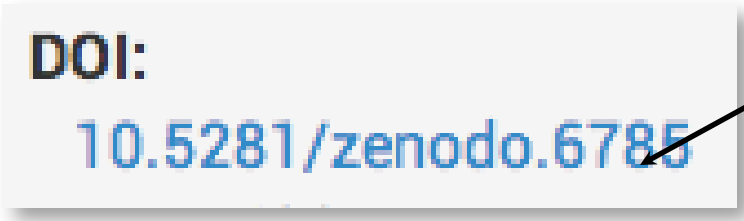
zenodo



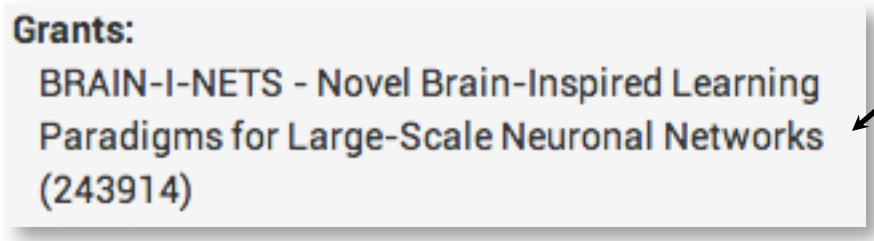
http://zenodo.org

Features

http://www.altmetric.com

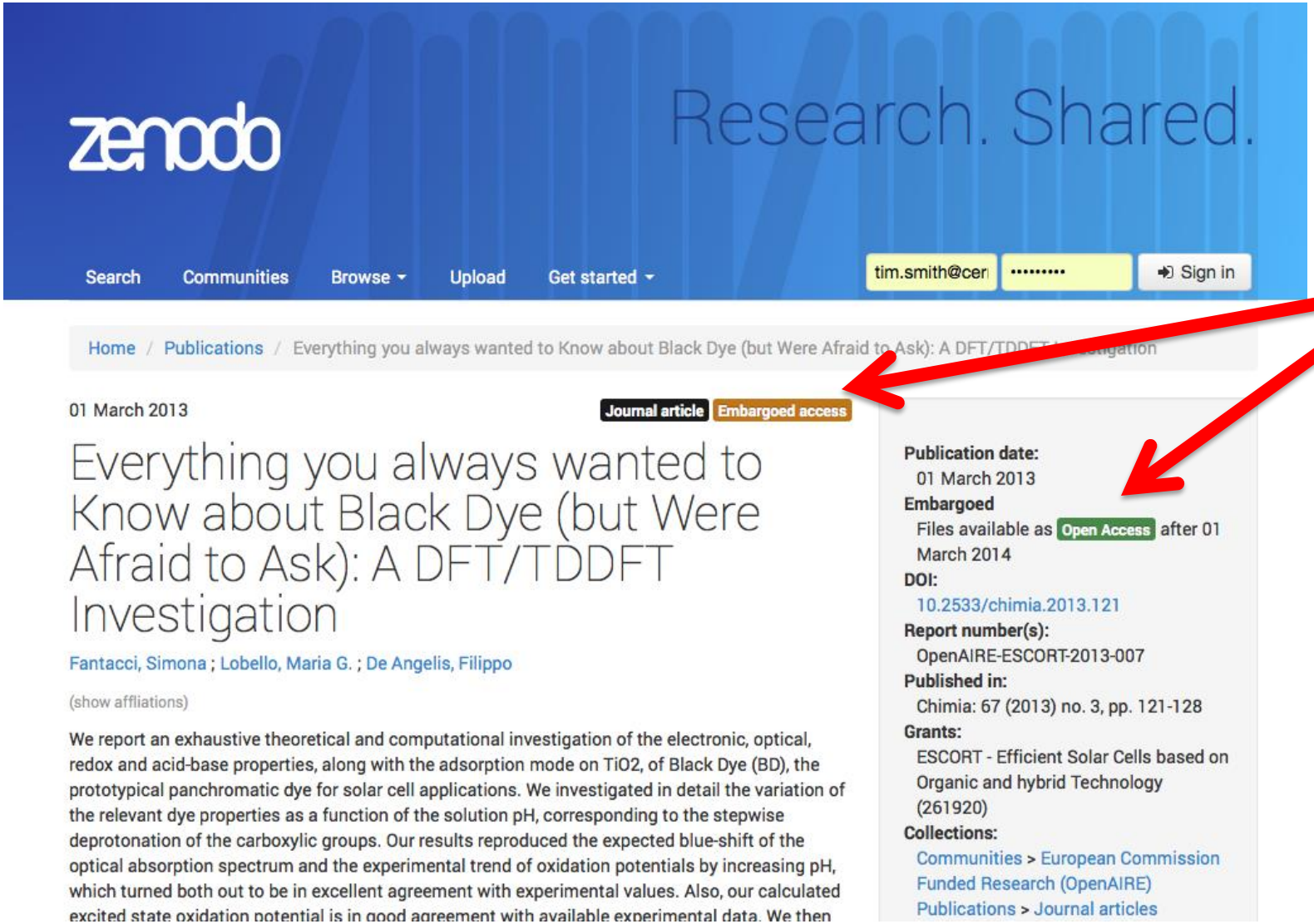


http://www.datacite.org

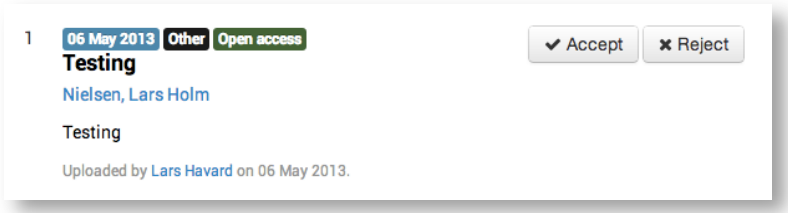


http://www.openaire.eu

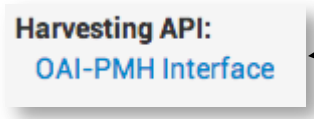
Research Repository



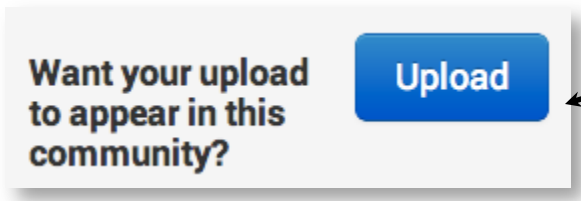
Communities



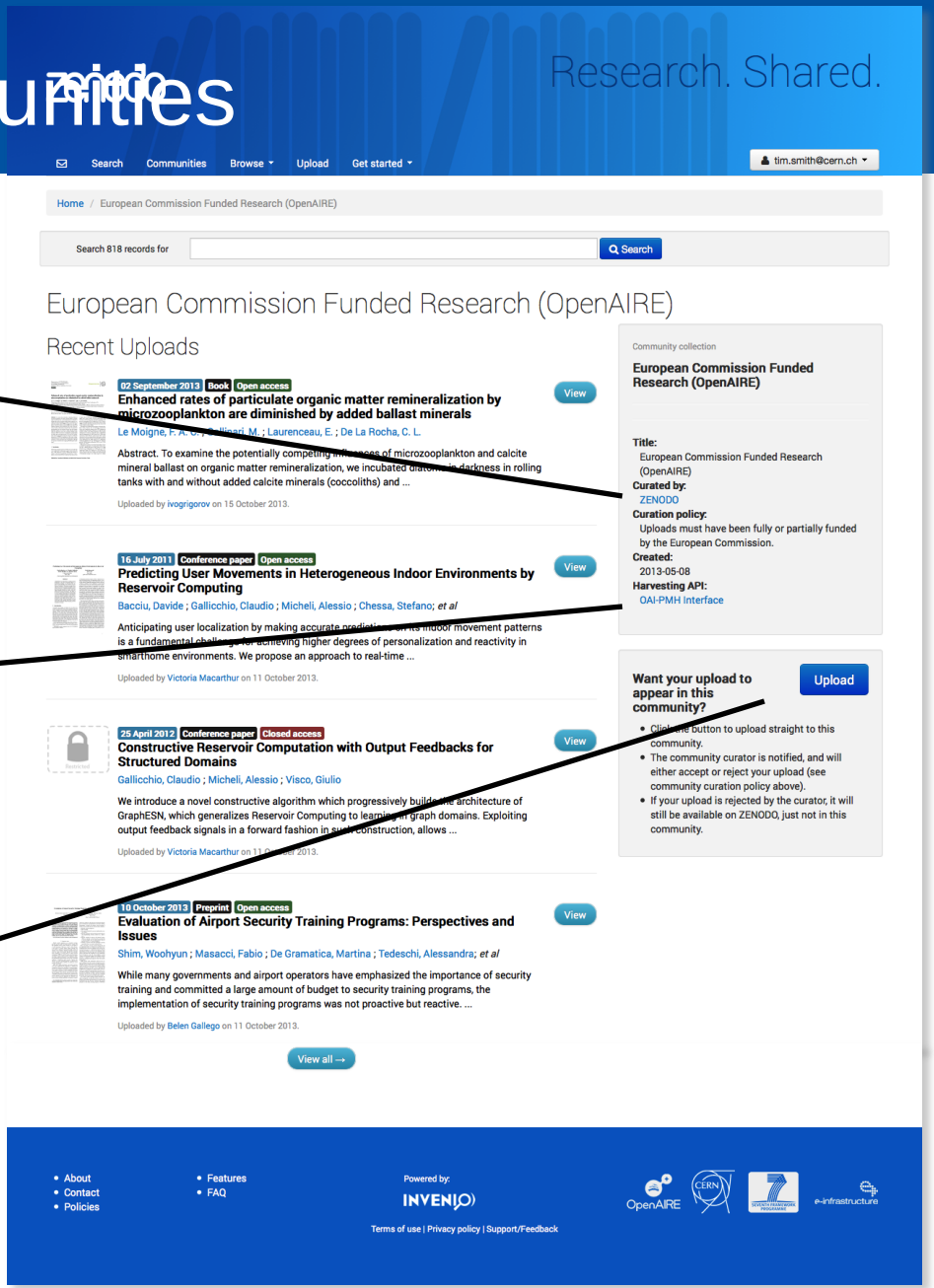
Accept/reject uploads



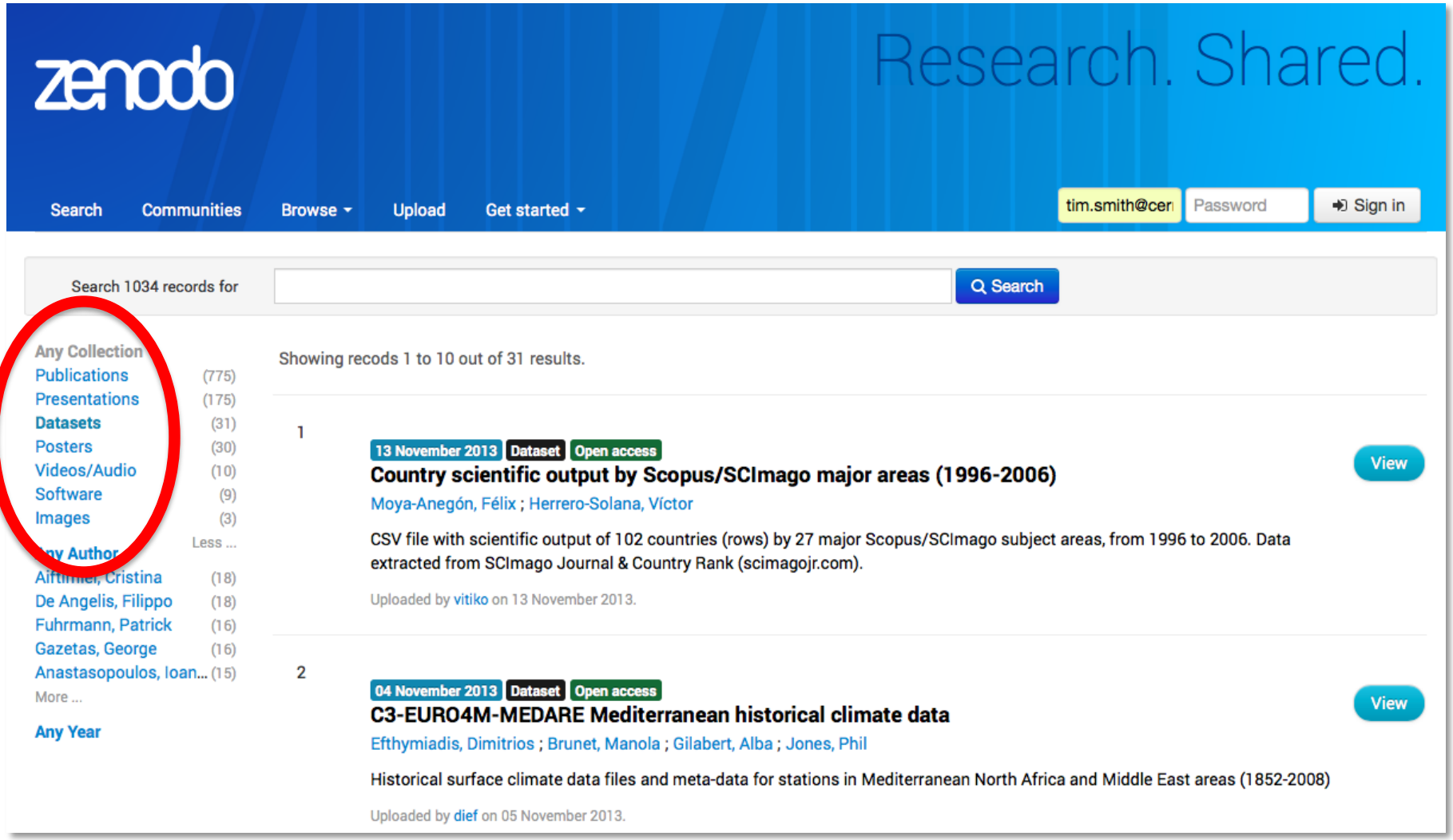
Export



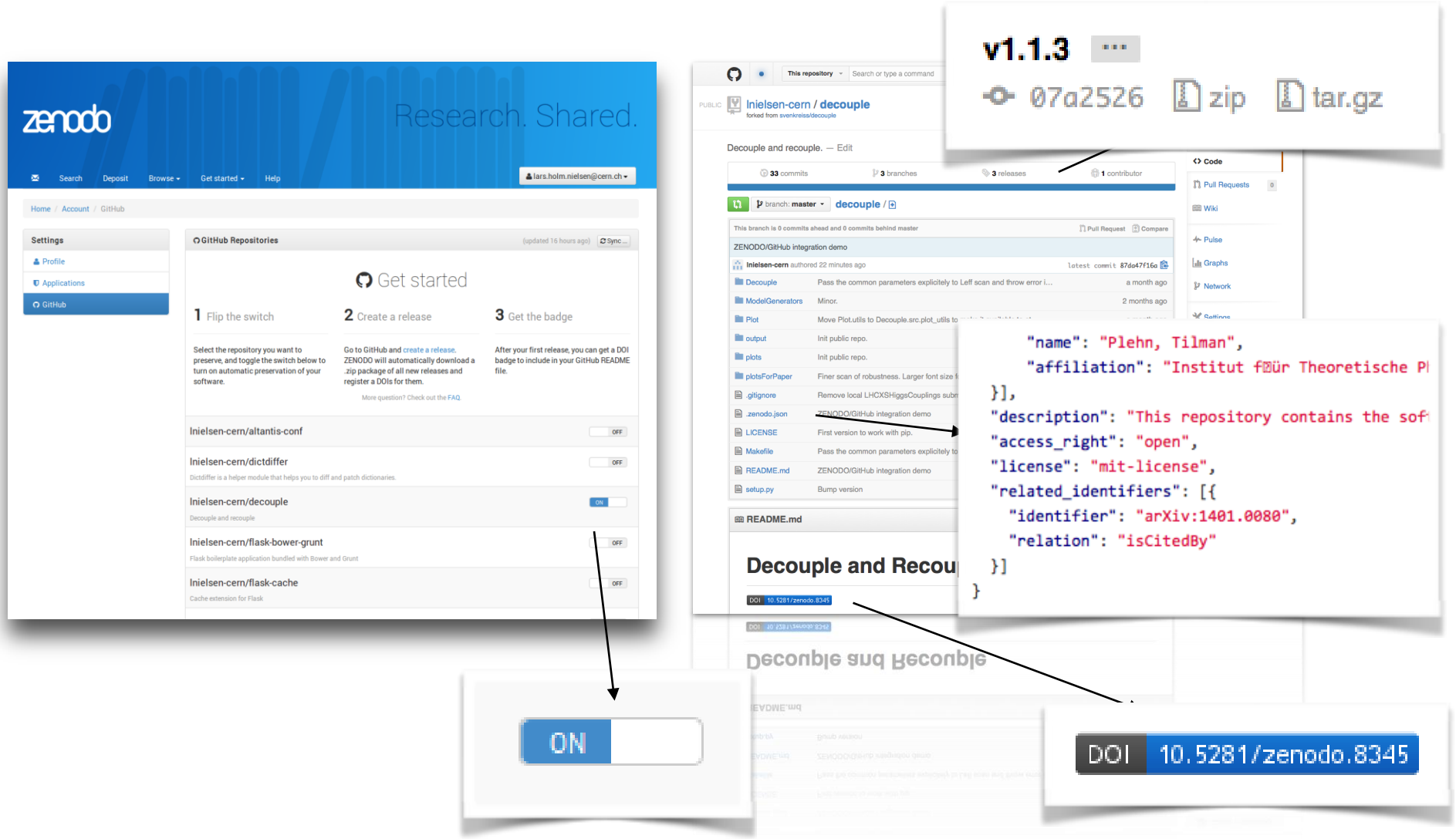
Direct community upload



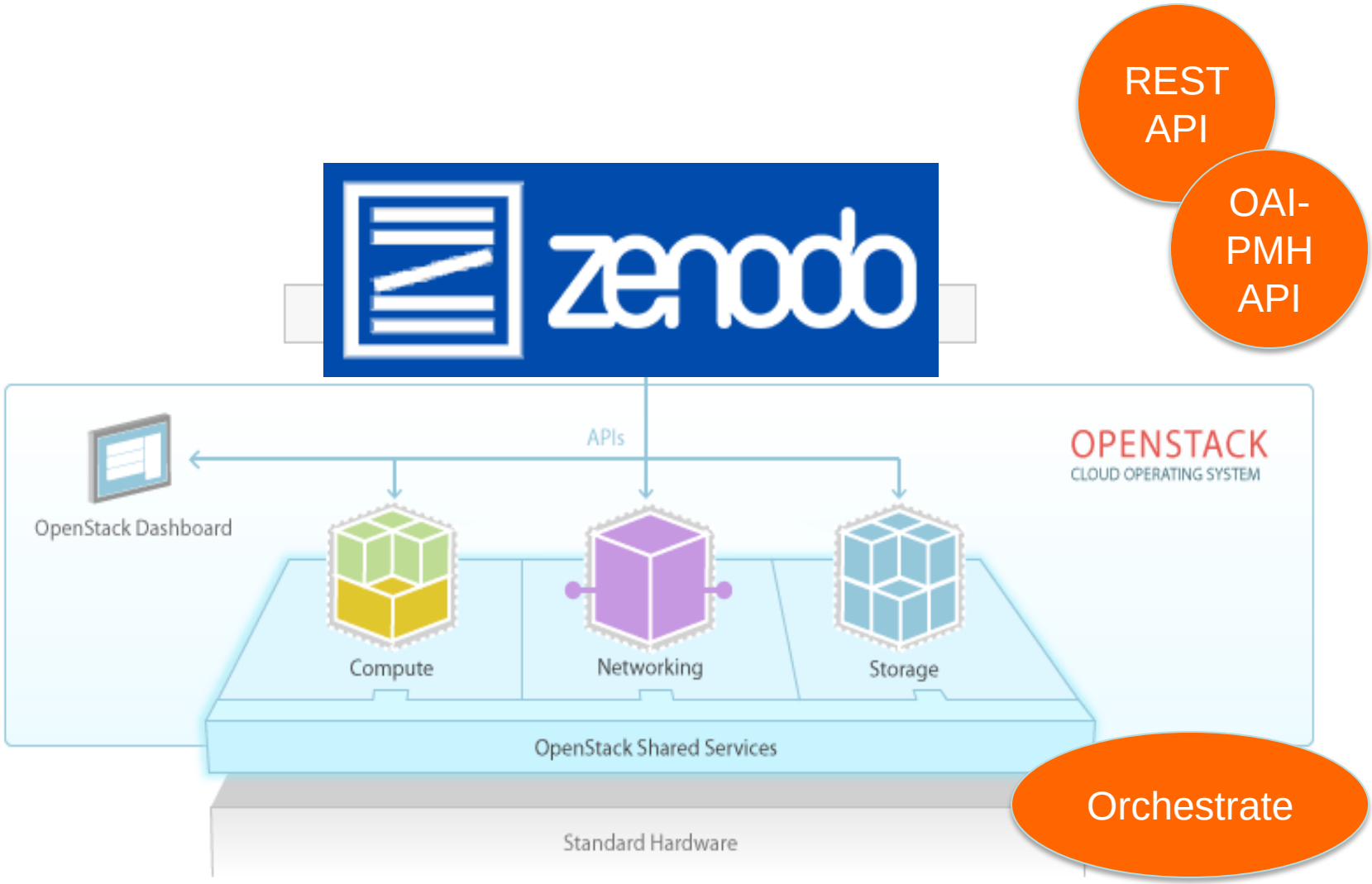
Research Repository



Reusability: Software Preservation



Open Data as a Service



Conclusions

- *Information is a valuable asset that is multiplied when it is shared*
- Mandates and policies
  - Openness, preservation
- Open Data
  - Discoverable, Accessible, Intelligible, Assessable, Useable
- Digital Libraries make this possible !

www.cern.ch