



národní
úložiště
šedé
literatury

Nové funkce a technologie v současných a budoucích verzích Invenia

Kunčar, Jiří
2012

Dostupný z <http://www.nusl.cz/ntk/nusl-126793>

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 03.05.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní [nusl.cz](http://www.nusl.cz) .

INVENIO



New features and technologies
in current and future Invenio versions

Jiří Kunčar
jiri.kuncar@cern.ch





Outline

- ▶ Introduction
- ▶ Development process
- ▶ Releases
- ▶ Novelties
- ▶ New development stack
- ▶ Conclusion




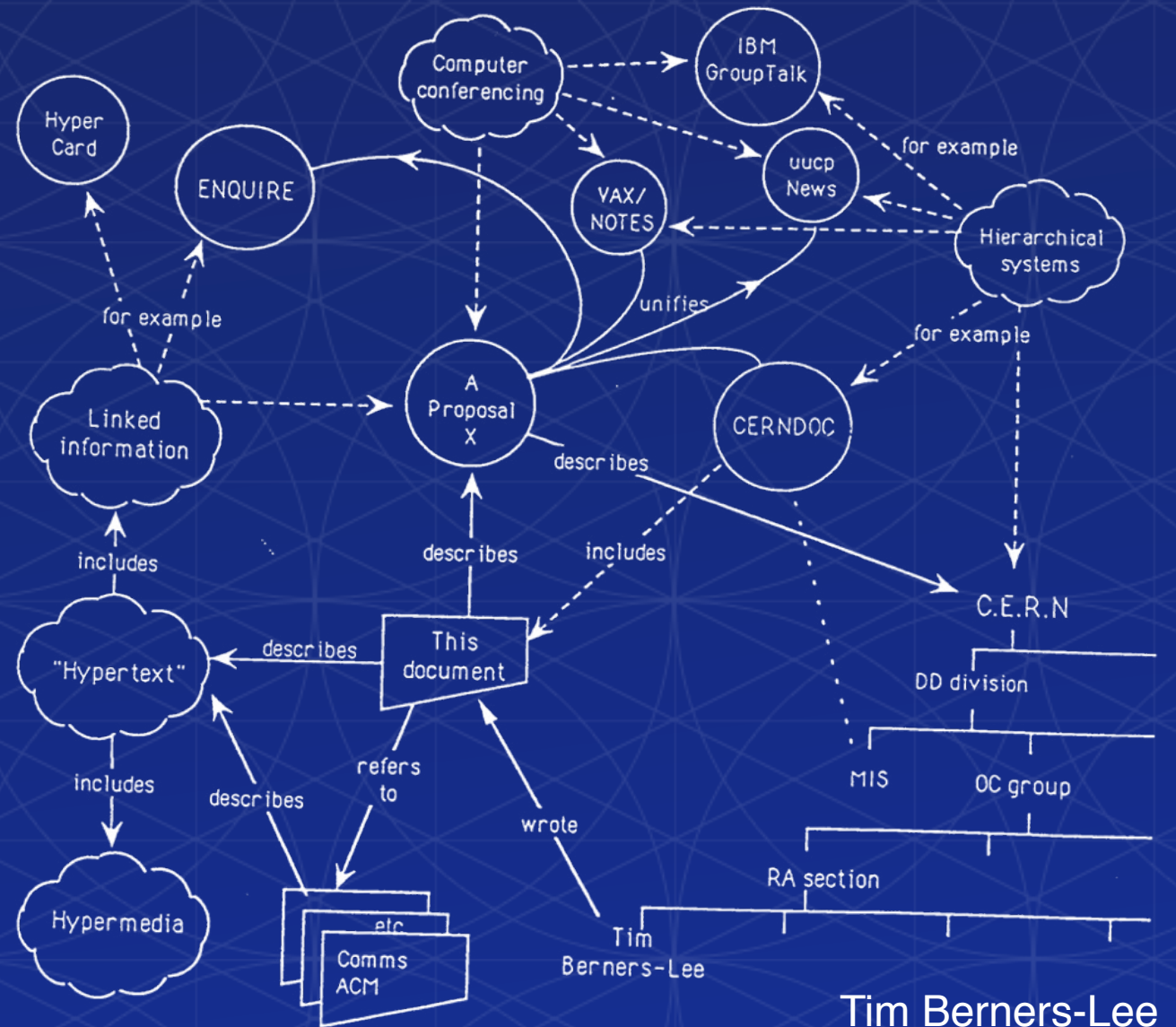
CERN

European Organization for Nuclear Research

- ▶ 1954 foundation
- ▶ 1989 World Wide Web
- ▶ 2002 Invenio Digital Library
- ▶ 2004 Indico Digital Conference

Large Hadron Collider

- ▶ particle accelerator
- ▶ 27 km circumference
- ▶ 100 meters underground
- ▶ located 



INVENIO

Integrated Digital Library System

2002-2012



- ▶ **free** software
- ▶ 10 years of development
- ▶ 35 developers, contributors and translators in 2011
- ▶ ~30 installations, ~4M of documents



1st Invenio User Group Workshop

7-9 May 2012
Geneva, Switzerland

“

The workshop is intended for Invenio administrators and consists of a series of lectures, practical exercises, and discussions with Invenio developers.

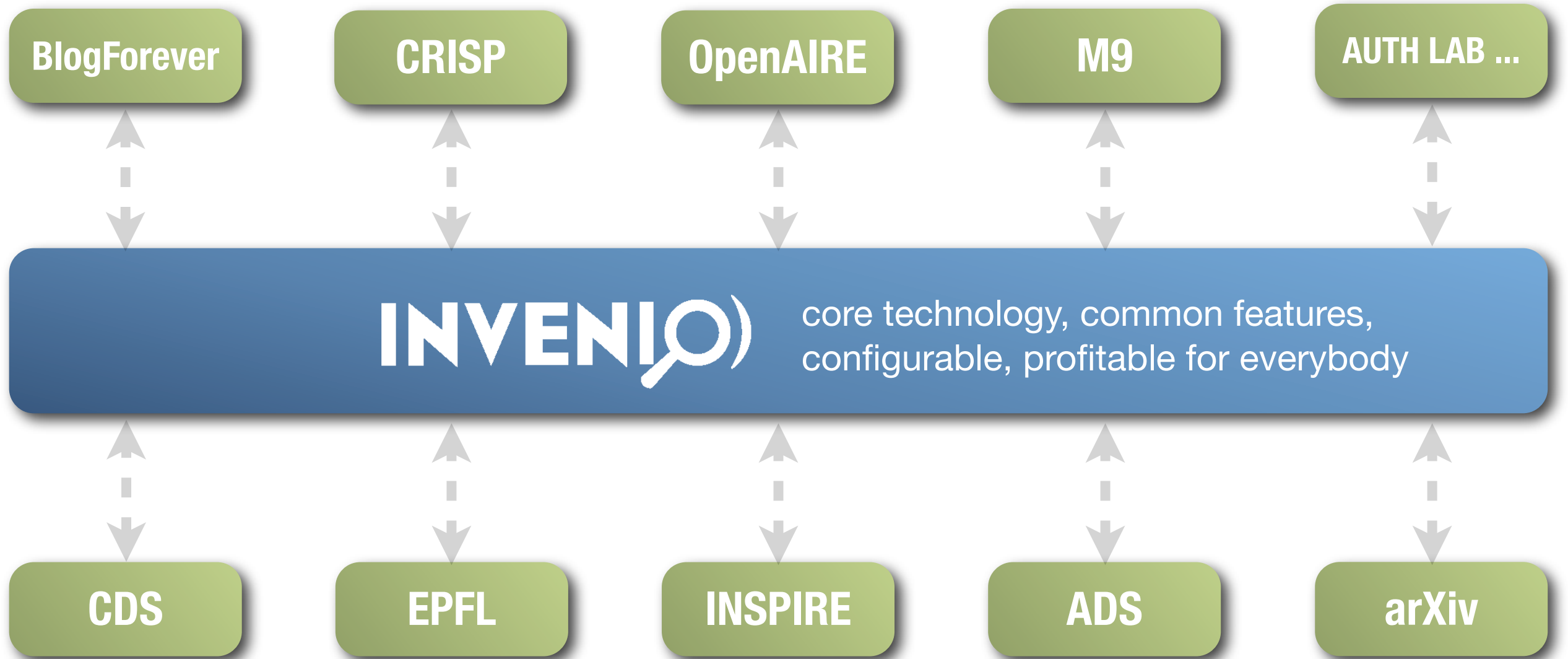


<http://indico.cern.ch/event/invenio-user-group-workshop-2012>



-----> Development Community

additional features, modules, themes, configurations



Release Numbering Scheme Explanation

0.99.5
1.0.1

major

- ▶ important rewrites and new functionality
- ▶ major API changes

minor

- ▶ new features
- ▶ APIs evolution
- ▶ DB schema evolution

patch-level

- ▶ bug fixes only
- ▶ no change to APIs or DB schema



Development Repository Branches

next

- ▶ testing new features
- ▶ experimental

master

- ▶ stabilizing features
- ▶ reaching desirable quality

maint-X.X

- ▶ bug fixes only
- ▶ production ready



INVENIO)

1.1



Novelties in version 1.1



- ▶ Upgrade Tool
- ▶ Circulation and Holdings
- ▶ Multimedia Encoder
- ▶ Author Identifier
- ▶ Sorting Buckets
- ▶ Database slaves
- ▶ Performance improvement
- ▶ 20+ new co-authors and contributors
- ▶ 1220 commits over 1.0 series

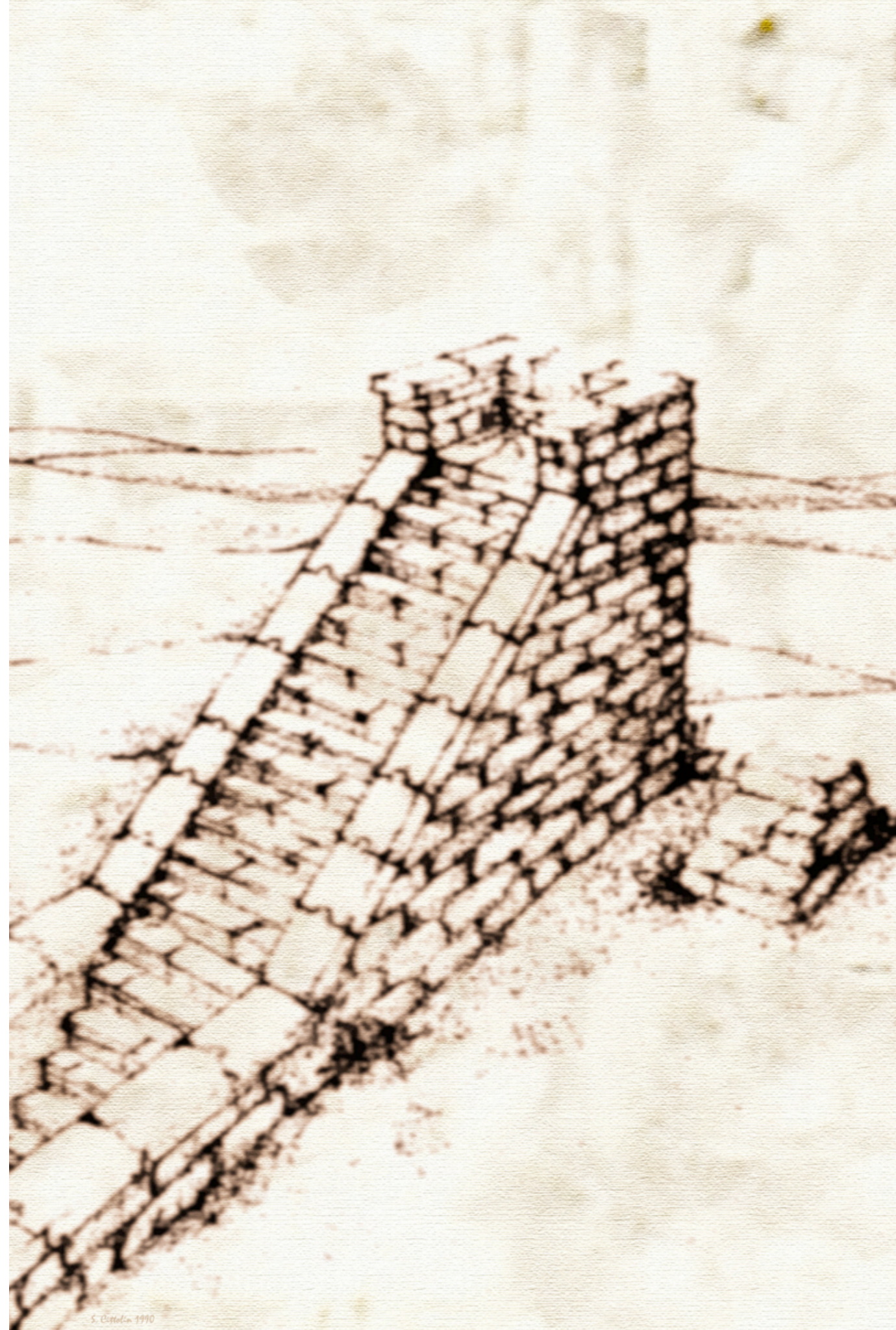




Upgrade tool

- ▶ automated deployment
- ▶ dependency graph
- ▶ multiple independent graphs
- ▶ upgrade validation
- ▶ time estimate calculation

inveniocfg --upgrade





Circulations and holdings

traditional library

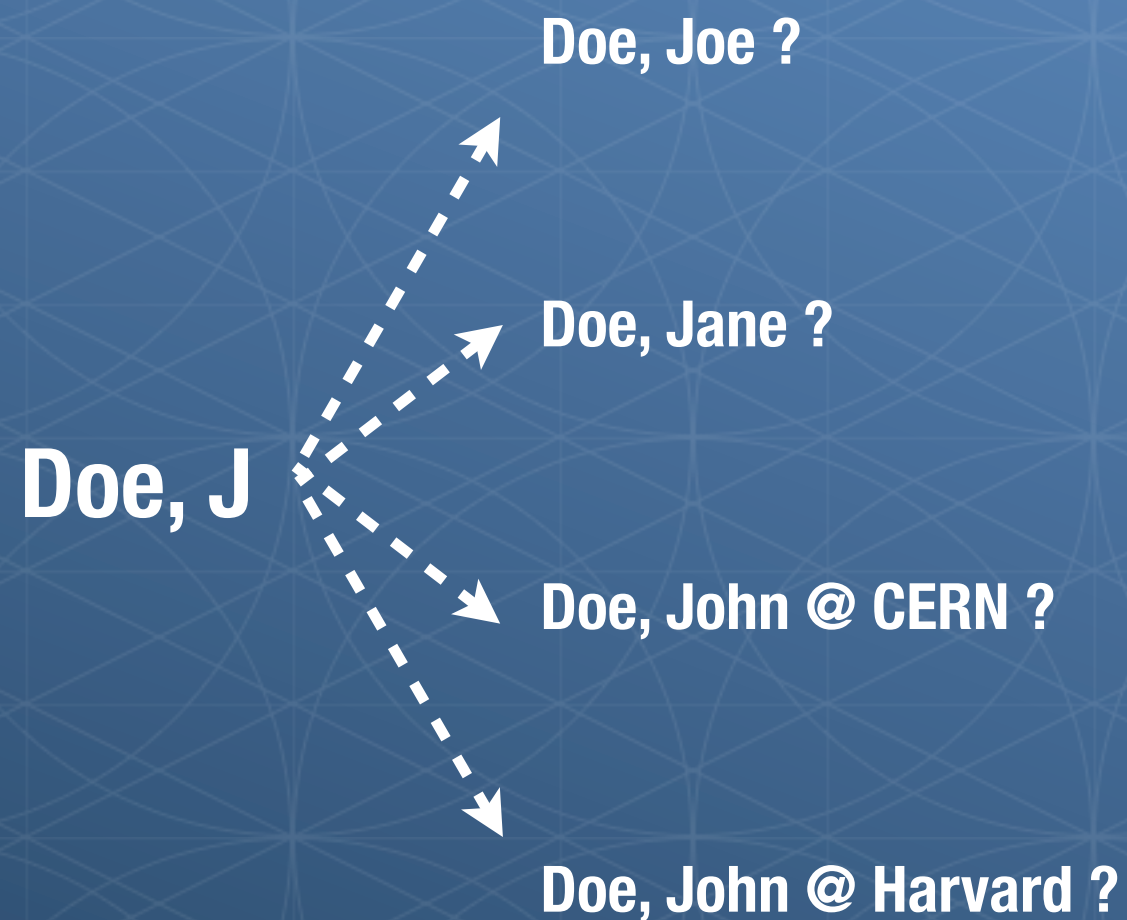
- ▶ lowering infrastructure maintenance costs by grouping several library services on a single server
- ▶ manage acquisition of institutional items, loans (including ILL) and user requests

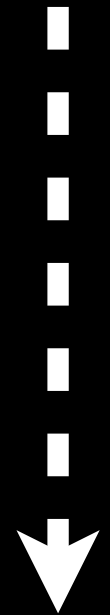


Author Identification

- ▶ disambiguation and claiming
- ▶ matching algorithm uses:
 - ▶ textual similarity
 - ▶ affiliation history
 - ▶ topics of papers
 - ...

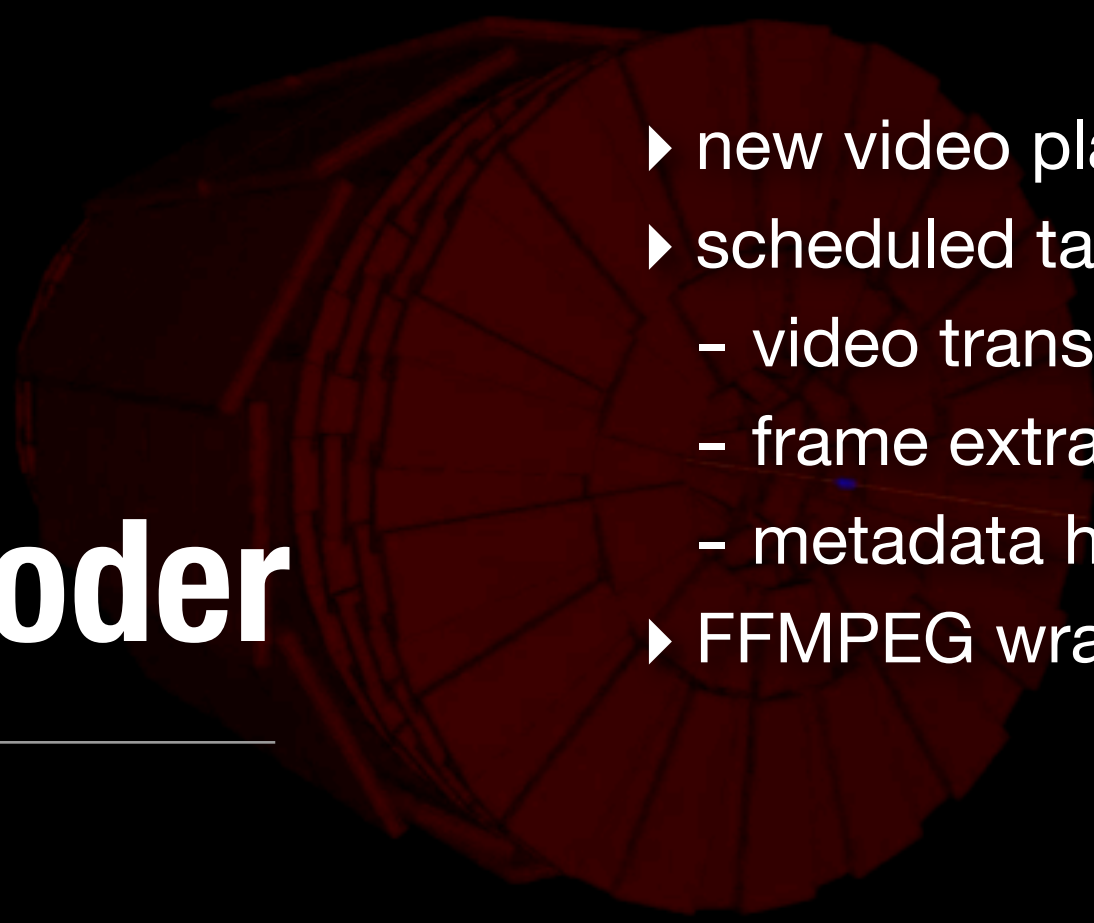
Who is who?



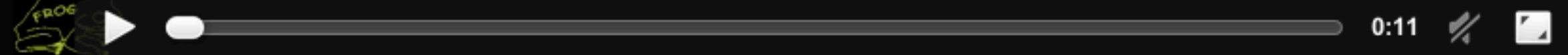


- ▶ new video player
- ▶ scheduled tasks
 - video transcoding
 - frame extraction
 - metadata handling
- ▶ FFMPEG wrapper

Media Encoder



480p



0:11



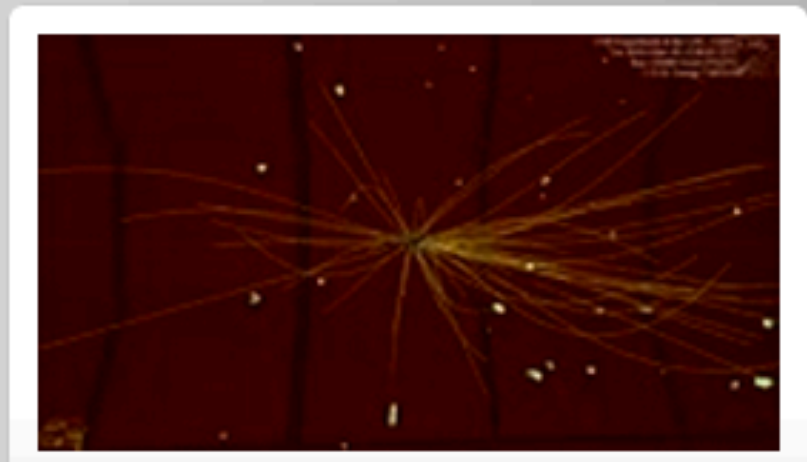
CMS animation of the high-energy collisions at 7 TeV on 30th March 2010 2010

by [CMS team](#) © 2010 CERN

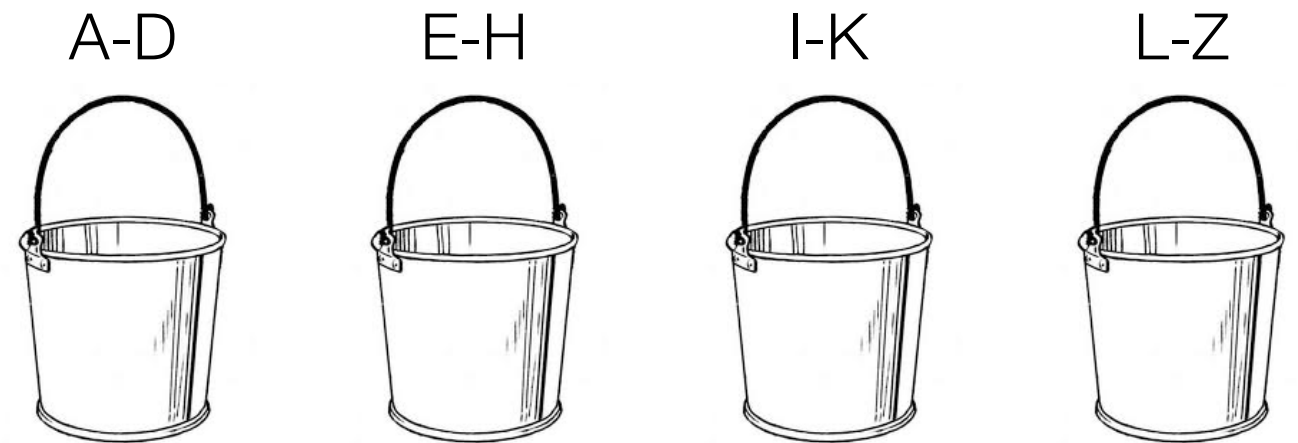
Run 132440 - Event 2732271

Write a comment

Download Video



CMS animation of the high-energy collisions at 7 TeV on 30th March 2010
by CMS team



Sorting Buckets

more efficient sorting algorithm for large repositories

- ▶ multiple sorting options
- ▶ configurable number of buckets
- ▶ in memory for fast processing



INVENIO

master branch
developments



Record Field Model

- ▶ logical fields
- ▶ data model description
- ▶ centralized definition
- ▶ non-MARC master formats
- ▶ easy API to access information

```
<collection xmlns="http://www.loc.gov/MARC21/slim">
  <record>
    <controlfield tag="001">1407699</controlfield>
    <controlfield tag="003">SzGeCERN</controlfield>
    <datafield tag="035" ind1=" " ind2=" ">
      <subfield code="9">arXiv</subfield>
      <subfield code="a">oai:arXiv.org:1203.6893</subfield>
    </datafield>
    <datafield tag="037" ind1=" " ind2=" ">
      <subfield code="a">arXiv:1203.6893</subfield>
    </datafield>
    <datafield tag="041" ind1=" " ind2=" ">
      <subfield code="a">eng</subfield>
    </datafield>
    <datafield tag="100" ind1=" " ind2=" ">
      <subfield code="a">Diaz-Cruz, L</subfield>
    </datafield>
    <datafield tag="700" ind1=" " ind2=" ">
      <subfield code="a">Diaz-Furlong, A</subfield>
    </datafield>
    <datafield tag="700" ind1=" " ind2=" ">
      <subfield code="a">Gaitan-Lozano, R</subfield>
    </datafield>
    <datafield tag="700" ind1=" " ind2=" ">
      <subfield code="a">Y, J H Montes de Oca</subfield>
    </datafield>
    <datafield tag="856" ind1="4" ind2=" ">
      <subfield code="u">http://arxiv.org/pdf/1203.6893.pdf</subfield>
      <subfield code="y">Preprint</subfield>
    </datafield>
    <datafield tag="916" ind1=" " ind2=" ">
      <subfield code="s">n</subfield>
      <subfield code="w">201214</subfield>
    </datafield>
    <datafield tag="245" ind1=" " ind2=" ">
      <subfield code="a">
        Probing flavor changing neutral currents and CP violation through the decay  $h \rightarrow c-b+W^-$  in the
        two Higgs doublet model
      </subfield>
    </datafield>
    <datafield tag="269" ind1=" " ind2=" ">
      <subfield code="c">02 Apr 2012</subfield>
    </datafield>
    <datafield tag="300" ind1=" " ind2=" "></datafield>
    <datafield tag="520" ind1=" " ind2=" ">
      <subfield code="a">
        We discuss the formulation of the general two-Higgs doublet model type III, which incorporates
        flavor changing neutral scalar interactions (FCNSI) and CP violation from several sources. CP
        violation can arise either from Yukawa terms or from the Higgs potential, and it can explicit or
        spontaneous. We discuss the case that includes CP violation with Yukawa textures to control FCNSI
        and evaluate the CP asymmetry for the decay  $h \rightarrow c-b+W^-$ , which may allow to test the patterns of
        FCNSI and CP violation, that arises in these models.
      </subfield>
    </datafield>
    <datafield tag="500" ind1=" " ind2=" ">
      <subfield code="a">Comments: 16 pages, 9 figures</subfield>
    </datafield>
    <datafield tag="650" ind1="1" ind2="7">
      <subfield code="a">Particle Physics - Phenomenology</subfield>
      <subfield code="2">arXiv</subfield>
    </datafield>
    <datafield tag="695" ind1=" " ind2=" ">
      <subfield code="a">hep-ph</subfield>
      <subfield code="9">LANL EDS</subfield>
    </datafield>
    <datafield tag="260" ind1=" " ind2=" ">
      <subfield code="c">2012</subfield>
    </datafield>
    <datafield tag="690" ind1="C" ind2=" ">
      <subfield code="a">PREPRINT</subfield>
    </datafield>
    <datafield tag="980" ind1=" " ind2=" ">
      <subfield code="a">PREPRINT</subfield>
    </datafield>
    <datafield tag="960" ind1=" " ind2=" ">
      <subfield code="a">11</subfield>
    </datafield>
  </record>
</collection>
```



Search 97 records for: Search Tips

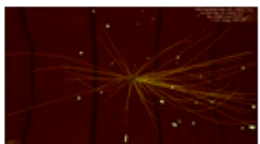
Showing records 1 to 2 out of 2 results.

Any Collection
Multimédia a Umění (2)

Any Author
Charles Darwin (1)
CMS team (1)

Any Year
2010 (1)

1 **The order Rodentia in South America / Charles Darwin**
The order Rodentia is here very numerous in species: of mice alone I obtained no less than eight kinds. 1The largest gnawing animal in the world, the Hydrochaerus capybara (the water-hog), is here also common. One which I shot at Monte Video weighed ninety-eight pounds: its length, from the end of the snout to the stump-like tail, was three feet two inches; and its girth three feet eight [...] test fr
Published in: Atlantis Times 02/2009, 03/2009
[Detailed record - Similar records](#)

2 
CMS animation of the high-energy collisions at 7 TeV on 30th March 2010
by CMS team
Run 132440 - Event 2732271
[Detailed record - Similar records](#)

HTML brief Export

before

Search 97 records for: Search Tips

Showing records 1 to 3 out of 3 results.

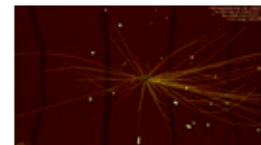
Any Collection
Multimedia & Arts (2)

Any Author
Charles Darwin (1)
CMS team (1)
Tavanapong, W (1)

Any Year
1999 (1)
2010 (1)

1 **A High-performance Video Browsing System / Tavanapong, W**
Recent advances in multimedia processing technologies, internetworking technologies, and the World Wide Web phenomenon have resulted in a vast creation and use of digital videos in all kinds of applications ranging from entertainment, business solutions, to education. [...] [Detailed record - Similar records](#)

2 **The order Rodentia in South America / Charles Darwin**
The order Rodentia is here very numerous in species: of mice alone I obtained no less than eight kinds. 1The largest gnawing animal in the world, the Hydrochaerus capybara (the water-hog), is here also common. One which I shot at Monte Video weighed ninety-eight pounds: its length, from the end of the snout to the stump-like tail, was three feet two inches; and its girth three feet eight [...] test fr
Published in: Atlantis Times 02/2009, 03/2009
[Detailed record - Similar records](#)

3 
CMS animation of the high-energy collisions at 7 TeV on 30th March 2010
by CMS team
Run 132440 - Event 2732271
[Detailed record - Similar records](#)

HTML brief Export

after

Improved search experience for logged in users

automatically showing records from restricted collections





External Ranking and Indexing

- ▶ generic bridge for external tools
- ▶ enhanced ranking techniques
- ▶ improves scalability
of the word similarity ranking
- ▶ tested on CDS production data





INVENIO

next branch
developments



Invenio **Next** Branch



DO SA SIAMO?
WHERE ARE WE GOING?
WOHIN GEHEN WIR?
¿QUÉ SOMOS?
WE?



Twitter



Bootstrap

- ▶ front-end framework
- ▶ responsive layout
- ▶ easily customizable
- ▶ variety of UI components
- ▶ JavaScript based plugins



sources: <http://twitter.github.com/bootstrap/index.html>

Search 97 records for:

[Search Tips](#)

Narrow by collections

[Multimedia & Arts](#) (9)[Atlantis Times](#) (5) [Poetry](#) (3) [Videos](#) (1) [Pictures](#) (0)[Books & Reports](#) (24)[Reports](#) (2) [Theses](#) (8) [Books](#) (14)[Articles & Preprints](#) (64)[Preprints](#) (36) [Articles](#) (28)

Focus on

[CERN Experiments](#) (2)[ISOLDE](#) (1) [ALEPH](#) (1)[CERN Divisions](#) (5)[Theoretical Physics \(TH\)](#) (4) [Experimental Physics \(EP\)](#) (1)

ABOUT THIS SITE

Welcome to the demo site of the Invenio, a free document server software coming from CERN. Please feel free to explore all the features of this demo site to the full.

SEE ALSO

[Invenio](#)
[CERN](#)

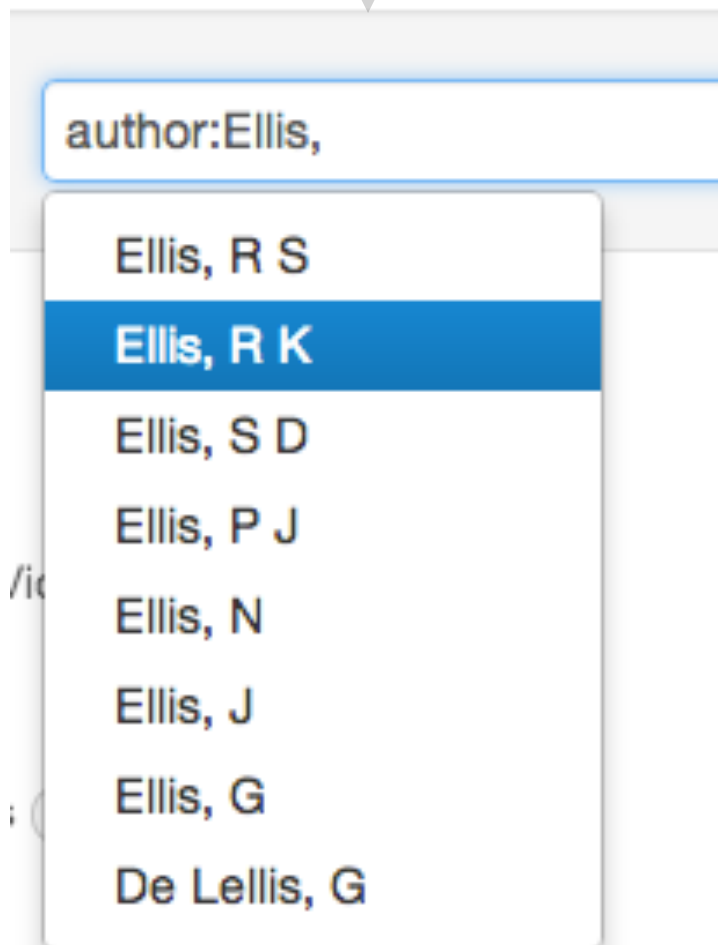


Demo

new page layout



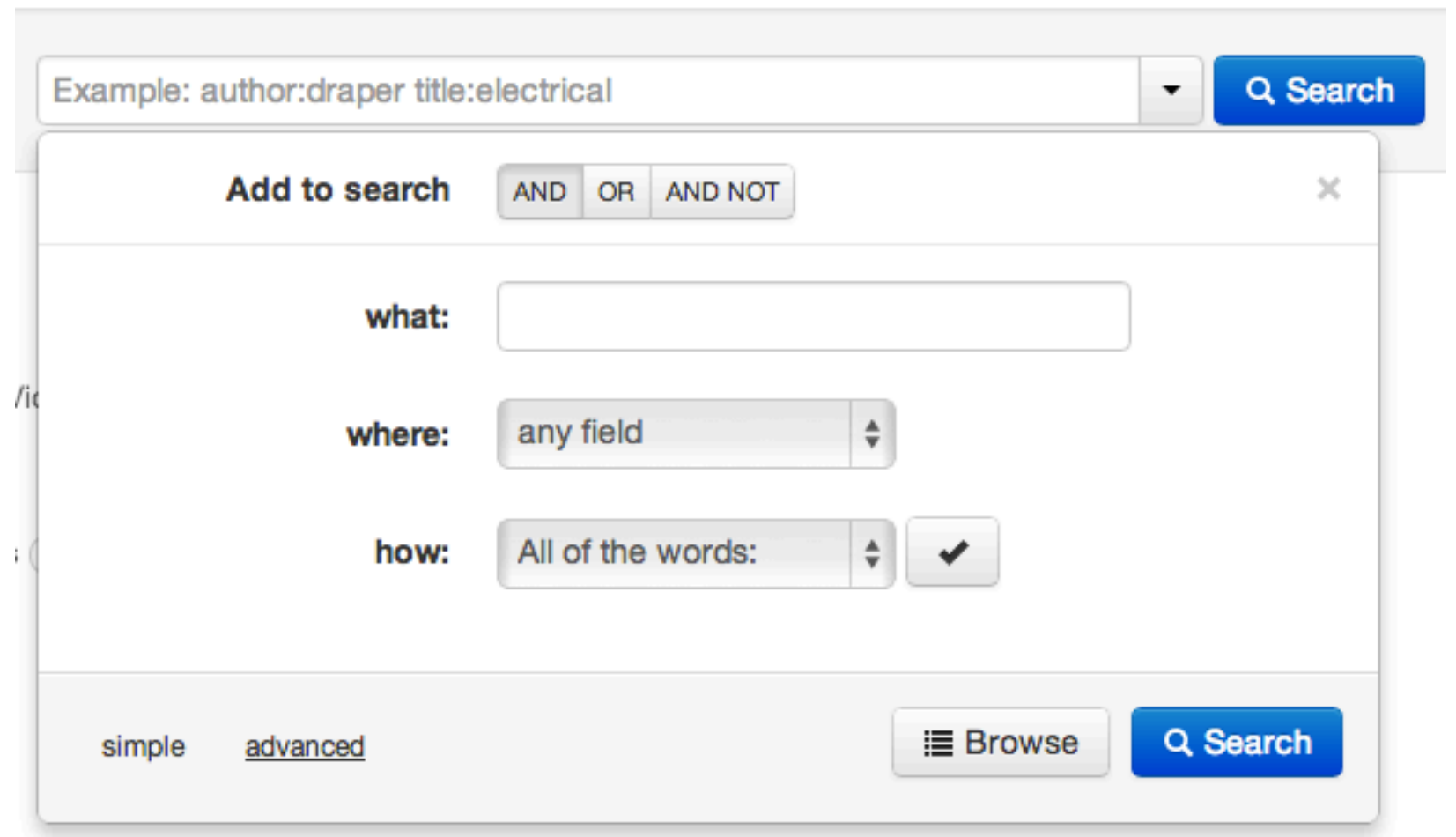
Demo



A search input field contains the text "author:Ellis,". Below the input, a dropdown menu is open, displaying a list of author names. The second item, "Ellis, R K", is highlighted with a blue background. The list includes:

- Ellis, R S
- Ellis, R K**
- Ellis, S D
- Ellis, P J
- Ellis, N
- Ellis, J
- Ellis, G
- De Lellis, G

Auto-completion



An advanced search interface with the following elements:

- Search bar: "Example: author:draper title:electrical" with a "Q Search" button.
- Modal window: "Add to search" with a close button (X) and logical operators: "AND", "OR", "AND NOT".
- Form fields:
 - "what:" followed by an empty text input field.
 - "where:" followed by a dropdown menu showing "any field".
 - "how:" followed by a dropdown menu showing "All of the words:" and a checkmark icon.
- Footer: "simple" and "advanced" (underlined) tabs, a "Browse" button with a list icon, and a "Q Search" button.

Interactive forms



Demo

Faceted search

limit to
exclude

Any Collection

Articles & Preprints (11)

CERN Divisions (4)

CERN Experiments (1)

Any Author

Ellis, J (4)

Bohrer, A (2)

Markou, C (2)

Abbaneo, D (1)

Abbott, R B (1)

More ...

Any Year

2001 (3)

1982 (2)

1985 (2)

2002 (2)

1987 (1)

More ...

Demo

Your Account

Filter widgets

Searches

You have made 32 queries. A detailed list is available with a possibility to (a) view search results and (b) subscribe to an automatic email alerting service for these queries.

Messages

You have 0 new messages out of 0 messages.

Group

You are involved in following groups:
montague-family

Customizable
widgets

- ▶ plugin based
- ▶ drag-able & sortable
- ▶ fast search

2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012



INVENIO

- ▶ navigable collection tree
- ▶ powerful search engine
- ▶ flexible metadata
- ▶ collaborative features
- ▶ personalization

<http://invenio-software.org>

Jiří Kunčar
jiri.kuncar@cern.ch

