

INVENIO



Nové funkce a technologie  
v současných a budoucích  
verzích Invenia

Jiří Kunčar  
[jiri.kuncar@cern.ch](mailto:jiri.kuncar@cern.ch)







# Obsah

---

- ▶ úvod
- ▶ vývojový proces
- ▶ verze
- ▶ novinky
- ▶ nové technologie
- ▶ závěr




# CERN

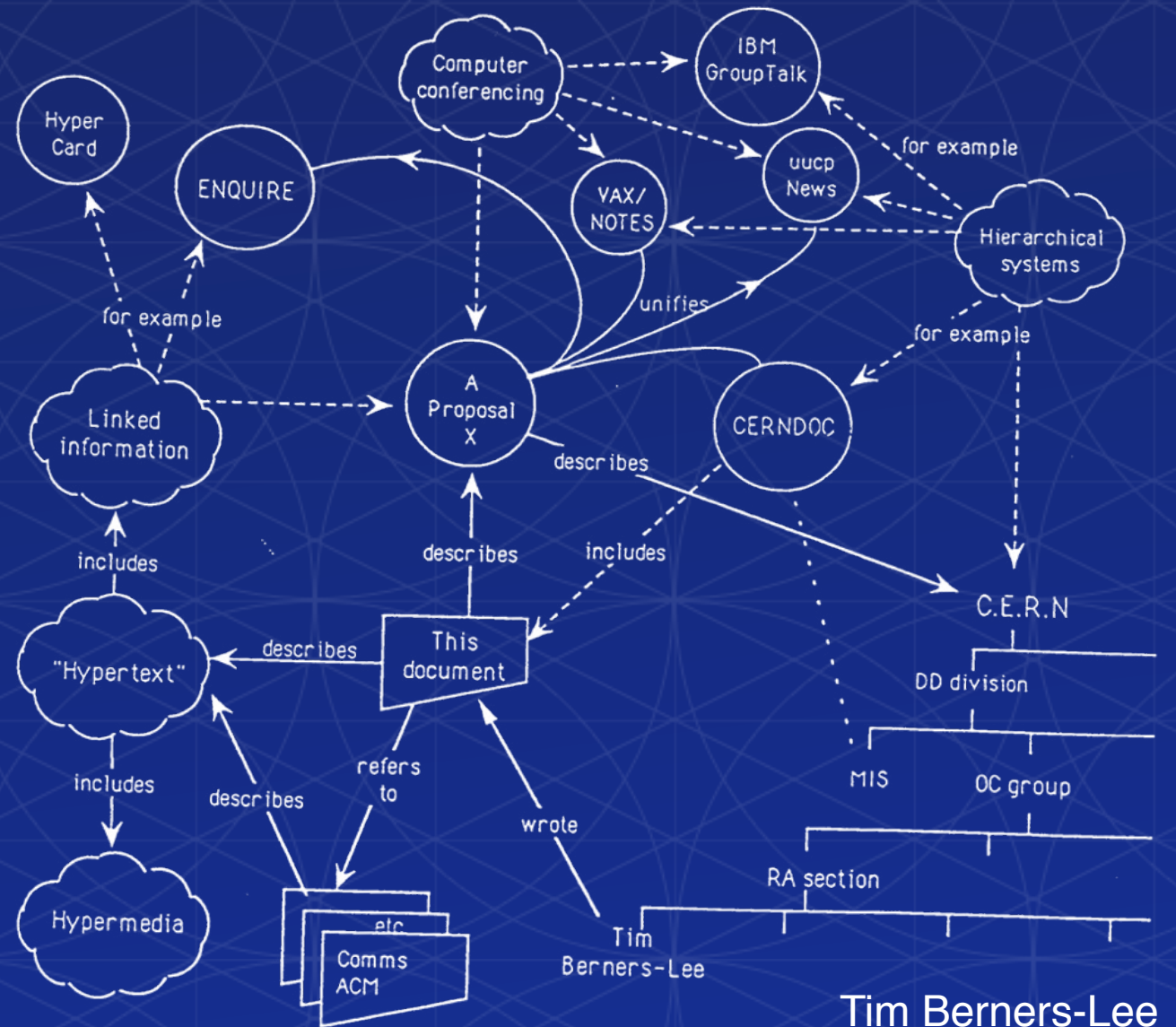
European Organization for Nuclear Research

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



## Large Hadron Collider

- ▶ částicový urychlovač
- ▶ obvod 27 km
- ▶ 100 metrů podzemí
- ▶ umístění  



# INVENIO

Integrovaný systém  
pro digitální knihovny

2002-2012



- ▶ **free** software
- ▶ 10 let vývoje
- ▶ 35 vývojářů, přispěvatelů a překladatelů v roce 2011
- ▶ ~30 instalací, ~4 miliony dokumentů





# 1. workshop pro uživatele Invenia

7.-9. květen 2012  
Ženeva, Švýcarsko

“

Workshop je cílen pro administrátory Invenia a skládá se ze série přednášek, praktických ukázek a diskuzí s vývojáři Invenia.

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



# Vývojářská komunita

další funkce, moduly, témata, konfigurace





# Schéma číslování verzí

0.99.5  
1.0.1

## major

- ▶ významné změny a nové vlastnosti
- ▶ větší změny v API

## minor

- ▶ nové vlastnosti
- ▶ evoluce API a schéma databáze

## patch-level

- ▶ pouze opravy chyb
- ▶ žádné změny v API a schématu



# Vývojové větve

## *next*

- ▶ testování nových funkcí
- ▶ experimentální

## *master*

- ▶ stabilizování funkcí
- ▶ dosažení požadované kvality

## *maint-X.X*

- ▶ pouze opravy chyb
- ▶ vhodné pro ostré nasazení



**INVENIO)**

**1.1**





## Novinky ve verzi 1.1

---

**INVENIO**

- ▶ Nástroj pro upgrade
  - ▶ Vypůjčky a sbírky
  - ▶ Zpracování videa
  - ▶ Identifikace autorů
  - ▶ Předtřízené segmenty
  - ▶ Database slaves
  - ▶ Výkonnostní optimalizace
- ▶ 20+ nových spoluautorů a přispěvatelů
  - ▶ 1220 commitů od řady 1.0



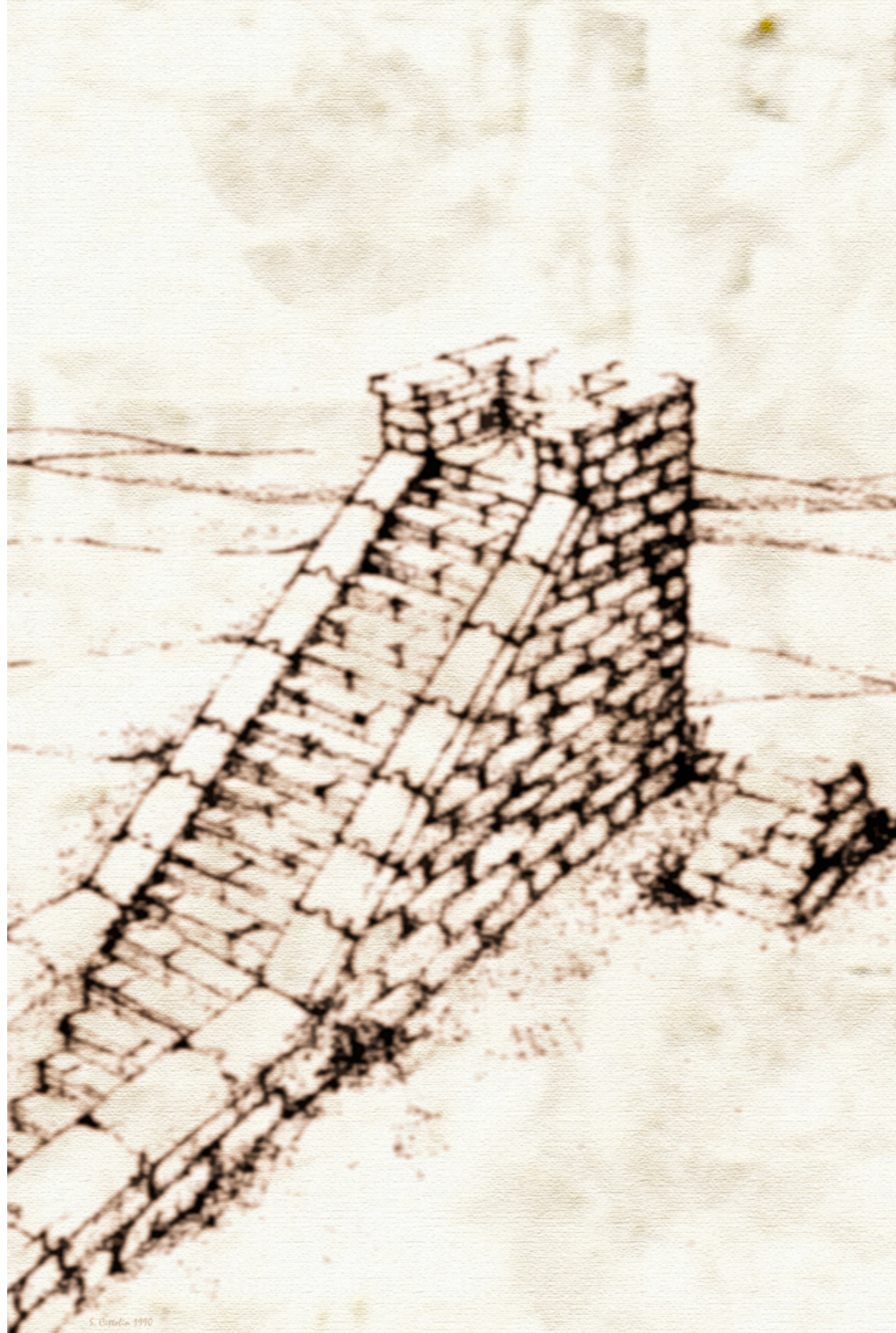


# Nástroj pro upgrade

---

- ▶ automatické nasazení
- ▶ graf závislosti
- ▶ více nezávislých grafů
- ▶ ověření před aktualizací
- ▶ odhad potřebného času

inveniocfg --upgrade







## Vypůjčky a sbírky

tradiční knihovna

- ▶ snížení nákladů na údržbu infrastruktury sloučením několika knihovných služeb na jeden server
- ▶ správa institucionálních položek, výpůjček (včetně meziknihovných výpůjček) a žádostí uživatelů



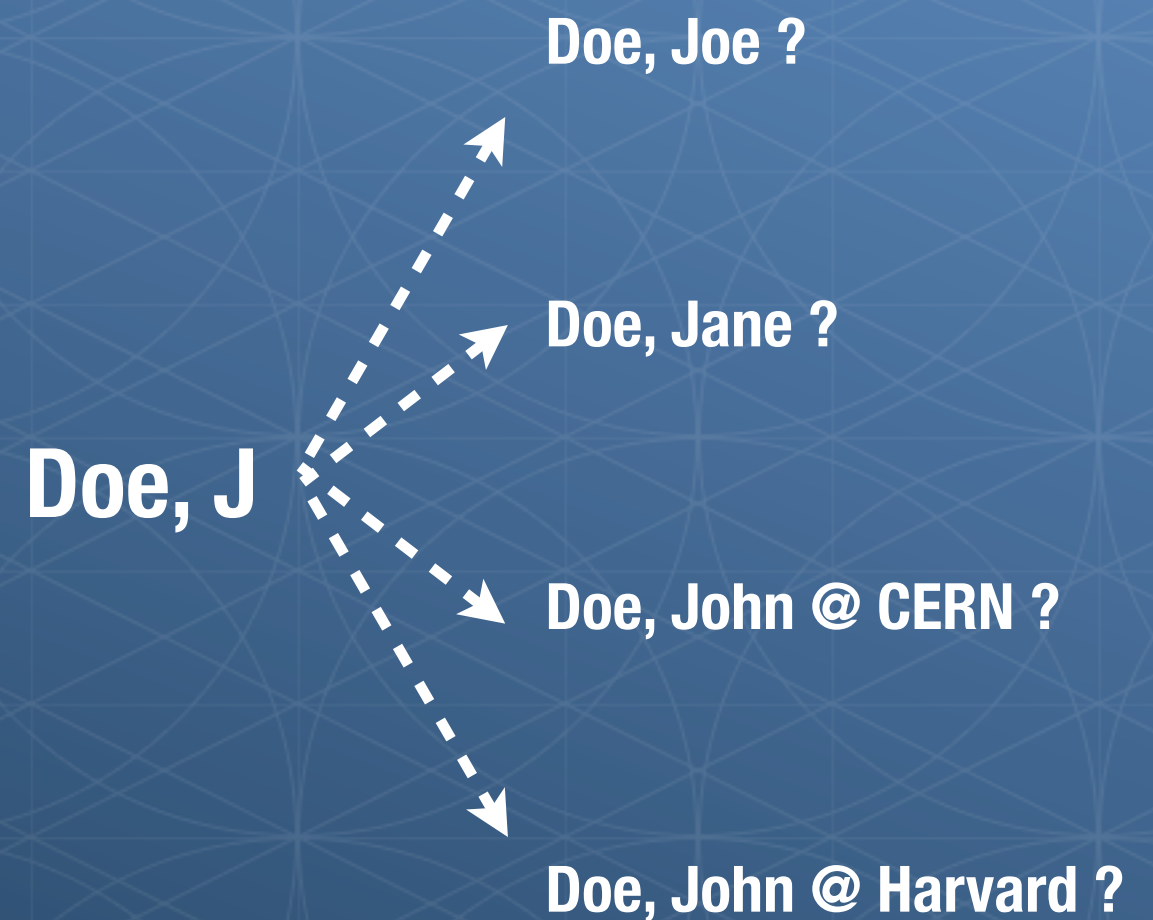
## Identifikace autora

---

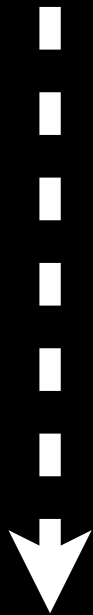
- ▶ disambiguace a nárokování
- ▶ algoritmus používá:
  - ▶ textová podobnost
  - ▶ historie afiliace
  - ▶ témata publikací
  - ...



# Kdo je kdo?







# Zpracování videa

- ▶ nový video přehrávač
- ▶ naplánované úlohy
  - video transcoding
  - extrakce klíčových snímků
  - správa metadat
- ▶ používá FFMPEG



480p



0:11



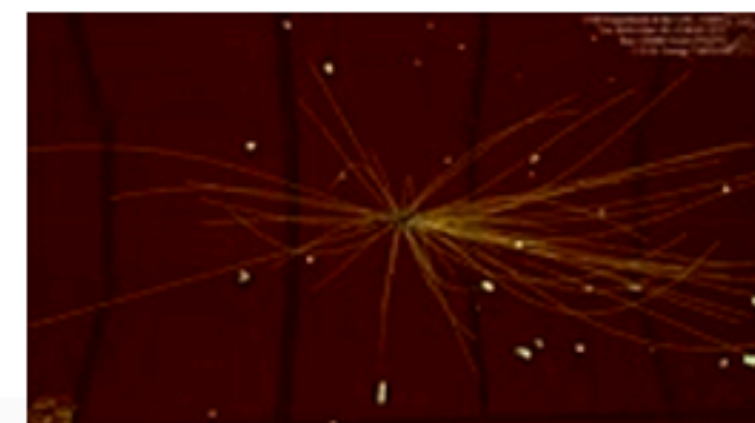
CMS animation of the high-energy collisions at 7 TeV on 30th March 2010 2010

by [CMS team](#) © 2010 CERN

Run 132440 - Event 2732271

Write a comment

Download Video

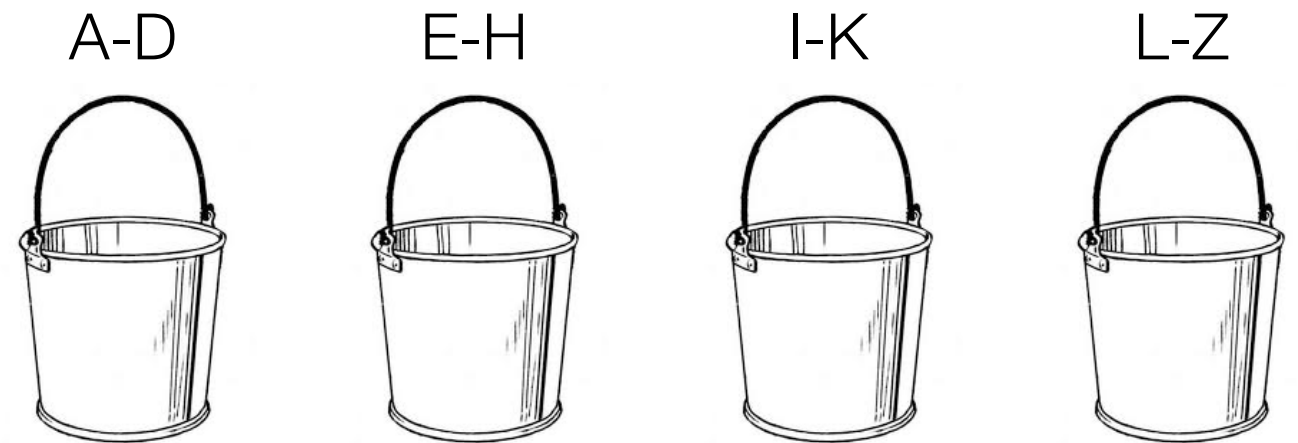


CMS animation of the high-energy collisions at 7 TeV on 30th March 2010

by CMS team



# Předtřízené segmenty



efektivnější třídící algoritmus pro velká úložiště

- ▶ více možností třídění
- ▶ konfigurovatelný počet segmentů
- ▶ uloženo v hlavní paměti pro lepší výkon



**INVENIO**

vývoj v *master* větvi





# Formát záznamu

- ▶ logická pole
- ▶ popis datového modelu
- ▶ centralizovaná definice
- ▶ možnost jiného hlavního formátu (non-MARC)
- ▶ jednoduché API pro přístup k uloženým informacím

```
<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  $h \rightarrow c\text{-}b^+ W^-$  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  $h \rightarrow c\text{-}b^+ W^-$ , 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

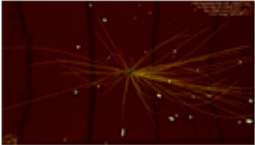
Showing records 1 to 2 out of 2 results.

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

dříve

Search 97 records for:   Search Tips

Showing records 1 to 3 out of 3 results.

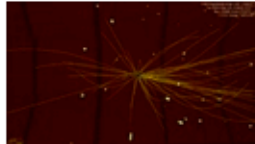
Any Collection  
Multimedia & Arts (2)

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

nyní

Vylepšený výstup hledání  
pro přihlášené uživatele

automaticky zobrazuje výsledky  
z soukromých kolekcí





# Externí klasifikace a indexování

---

- ▶ propojení s externími nástroji
- ▶ zdokonalené klasifikační techniky
- ▶ vylepšuje škálovatelnost klasifikace podobnosti slov
- ▶ testováno na datech CDS





**INVENIO**

vývoj v *next* větvi





## Invenio: *next* větev

---



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



Twitter



# Bootstrap

- ▶ front-end framework
- ▶ rozložení stránky reagující na změnu velikosti
- ▶ jednoduše přizpůsobitelný
- ▶ velká škála UI komponent
- ▶ zásuvné JavaScriptové moduly



zdroj: <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

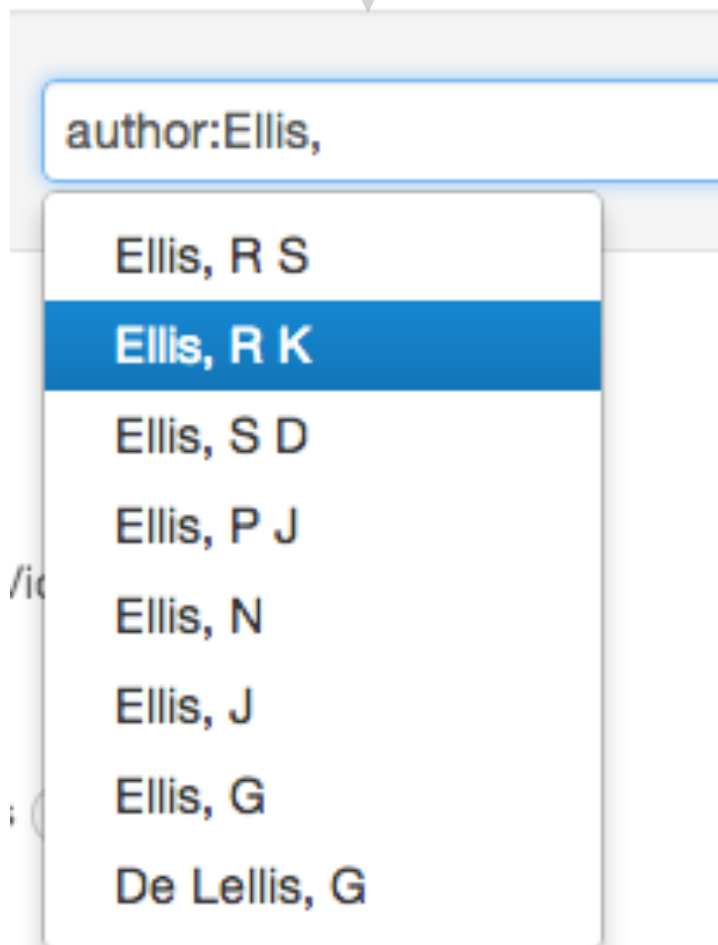
[Invenio](#)  
[CERN](#)

# Demo

nový design hlavní stránky



# Demo

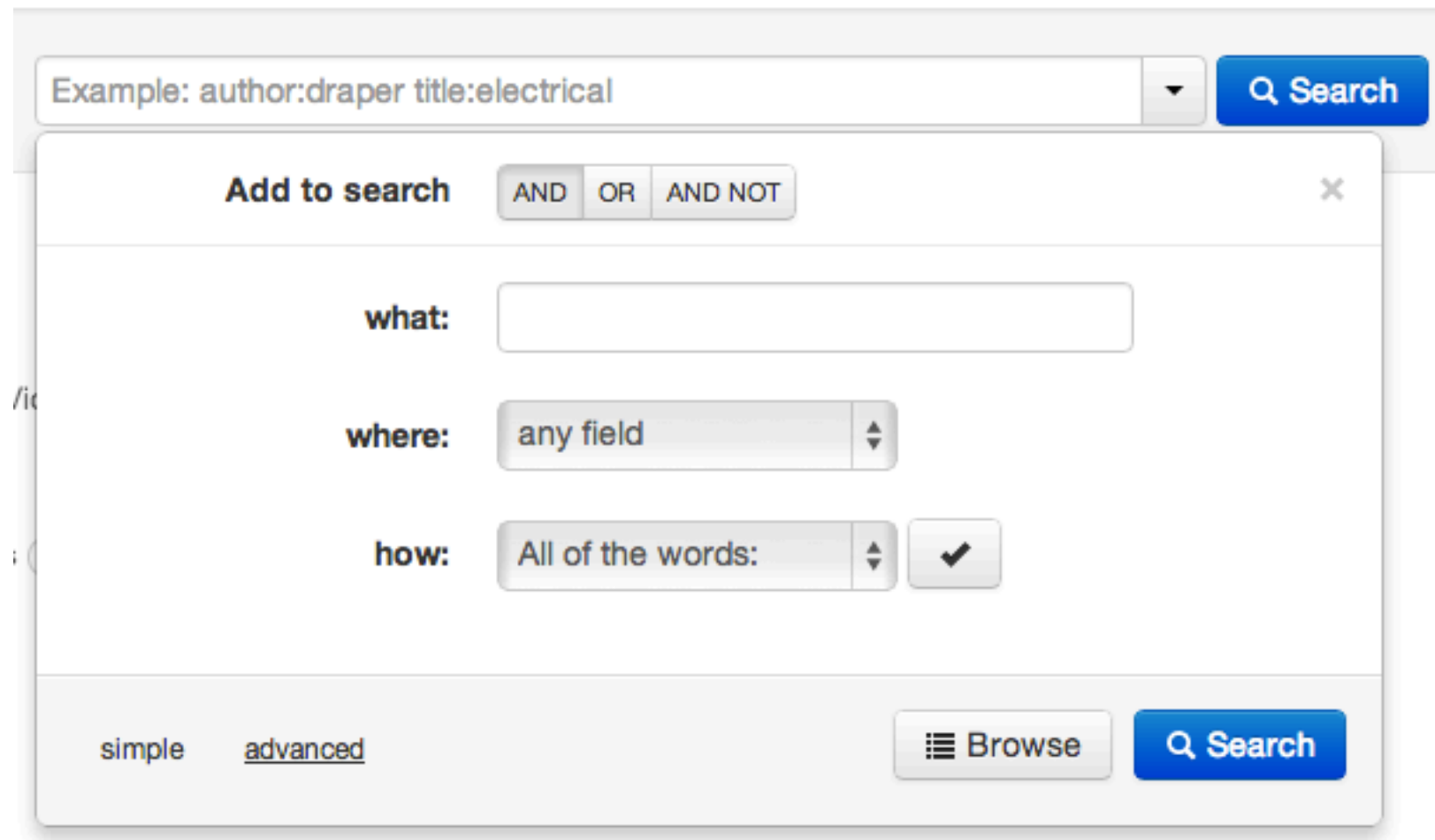


author:Ellis,

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

This screenshot shows a search input field with the text "author:Ellis,". Below the input field, a dropdown menu is open, displaying a list of author names. The name "Ellis, R K" is highlighted in blue, indicating it is the selected or most relevant suggestion.

Automatické  
dokončení



Example: author:draper title:electrical

Q Search

Add to search AND OR AND NOT

what:

where: any field

how: All of the words:

simple advanced Browse Search

This screenshot shows a more complex search interface. At the top, there is a search bar containing the example query "author:draper title:electrical" and a "Q Search" button. Below this is a modal window titled "Add to search" with a close button (X). The modal contains three sections: "what:" with a text input field, "where:" with a dropdown menu set to "any field", and "how:" with a dropdown menu set to "All of the words:" and a checked checkbox. At the bottom of the modal, there are links for "simple" and "advanced" search, a "Browse" button, and another "Q Search" button.

Interaktivní formuláře





# Demo

## Facetová klasifikace

---

nezahrnout  
omezit na

### 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**

## Prizpůsobitelné widgety

- ▶ zásuvné moduly
- ▶ možnost přetáhnout a změnit pořadí
- ▶ rychlé hledání

2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012



**INVENIO**

- ▶ navigovatelný strom kolekcí
- ▶ výkonný vyhledávací engine
- ▶ flexibilní metadata
- ▶ nástroje pro spolupráci
- ▶ personalizace

<http://invenio-software.org>

Jiří Kunčar  
[jiri.kuncar@cern.ch](mailto:jiri.kuncar@cern.ch)

