



Targeted Annotations of Scientific Papers

Tibor Šimko
CERN

AAHEP7
Stony Brook University, New York
3 April 2014

- 1 Use cases
- 2 Technology
- 3 Demo
- 4 Conclusions

1

Use cases

- internal commenting rounds in LHC experiments
 - e.g. 1 conference paper, 180 comments, 1000+ annotations

CERN Document Server

[Search](#)[Submit](#)[Help](#)[Personalize](#)[Administration](#)

[Home](#) > [CERN Experiments](#) > [LHC Experiments](#) > [LHCb](#) > [LHCb Internal](#)

LHCb Internal

Search 952 records for:

[Search](#)

[Search Tips](#)
[Advanced Search](#)


[Add to Search](#) [+](#)

Narrow by collection:

[LHCb Publication Drafts \(11\)](#) [\[restricted\]](#)

—

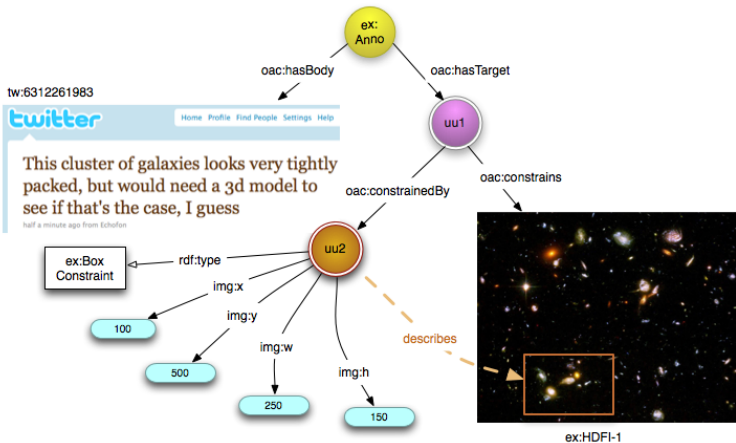
- targeted annotations: referring to locations e.g. “L.181”
- online & offline annotation style

▼  user <NAME WITHDRAWN> <DATE WITHDRAWN>

>>- L65: "using all invisible particles" -> "using all neutrinos" (or are you using others?)
Fixed

>>- L181: $|m_{ee} - m_Z|$ -> $|m_{ee} - M_Z|$ (pole mass of Z)
Fixed

>>- L188: "Monte Carlo (MC) simulation" -> "simulation". Not sure if it's still advised to replace all MC
by simulation.



Open Annotation RDF data model

2

Technology

```
{
  _id: UUID, // annotation identifier
  who: Number, // user identifier
  where: *, // annotation target (e.g. URI, recID)
  what: *, // content and its properties
  when: Timestamp, // creation date and time
  perm: Object // permissions (public, private,
                // group-restricted)
}
```

General JSON model for annotations


```
{  
  where: {  
    record: Number, // record identifier  
    marker: String, // location e.g. "P.2-T.3-L.4"  
  },  
  what: String,  
  comment: Number // comment ID of this annotation  
}
```

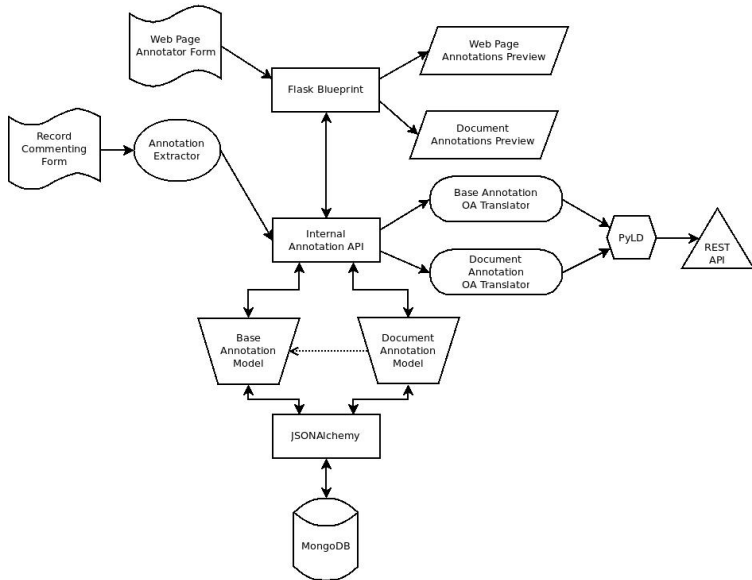
JSON fields specific for targeted annotations

■ location markers:

- S.7 refers to section 7
- P.7 refers to page 7
- PP.7 refers to paragraph 7
- L.7 refers to line 7
- F.7 refers to figure 7
- T.7 refers to table 7
- E.7 refers to equation 7
- R.[7] refers to reference [7]
- G refers to general paper aspects

■ nested markers, no depth limit:

- P.2: T.3: L.4 refers to line 4 of table 3 on page 2



3

Demo



Invenio Labs

"Playground for experimental features not quite ready for prime-time"

Search records

Search

Annotation Demo (1)

Click to add an annotation on the current page

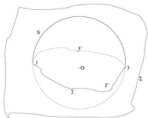


Figure 2:



Figure 3:

Σ : the curve Γ . (Fig. 2).

From now on, we shall assume that, with respect to some reference "horizontal" plane the slope of the surface does not exceed Θ_M , i.e. given P and Q on Σ :

$$\left(\vec{n}_Q, \vec{n} \right) \geq \cos \Theta_M, \quad (1)$$

where \vec{n} is the perpendicular to the horizontal plane. We assume further that at any given point there is a tangent plane to the ground.

It turns out to be convenient to impose a condition which guarantees that O , the center of the sphere S , is inside the tetrahedron $1\ 2\ 3\ 3'$.

This tetrahedron has right angles at 2 and 3. If we project it onto the plane perpendicular to 23 we get the view shown in Fig. 3.

O the image of O , is on the bisector of $1\ 23$ and $3'\ 23$, and also on the mediatrice of $1\ 23$. O will be inside the tetrahedron if the angle $(23\ 1, 23\ 3')$

Annotations on Page 5

Praesent velit velit, consectetur nec mattis ut, dapibus id metus.

— [juliet](#) · 2014-03-10 09:50:00 · [reply to original comment](#)

Phasellus mattis sodales nisl, quis euismod diam consectetur sit amet.

— [bervolio](#) · 2014-03-10 09:50:00 · [reply to original comment](#)

Annotations on Equation 1

Quisque vehicula tempor tempor. Vestibulum tempor posuere felis in interdum. Donec quis lorem id felis mattis fermentum.

— [romeo](#) · 2014-03-10 09:50:00 · [reply to original comment](#)

Genius!

— [cranmer](#) · 2014-03-14 17:07:59 · [reply to original comment](#)

INVENIO) » REST API: Export All

```
$ curl http://labs.invenio-software.org/api/annotations/export/
```

```
[
  {
    "@context": {},
    "_id": "91a134f1-7ae2-4054-aac3-c38767469464",
    "comment": 1,
    "perm": {
      "groups": [],
      "public": true
    },
    "what": "Lorem ipsum dolor sit amet, consectetur adipiscing
           elit. Donec interdum nisi ut nibh porta, non viverra
           neque faucibus.",
    "when": "2014-03-10T08:50:01+00:00",
    "where": {
      "marker": "P.1-S.Abstract",
      "record": 1
    },
    "who": 2
  },

```

```
$ curl http://labs.invenio-software.org/api/annotations/export/ \
-H "Content-Type: application/json" \
--data '{"query": {"where.record": 1}, "ldexport": "full"}'
```

```
[
  {
    "@context": { ... },
    "@id": "http://labs.invenio-software.org/api/annotations/
      export/?_id=91a134f1-7ae2-4054-aac3-c38767469464",
    "@type": "oa:Annotation",
    "annotatedAt": "2014-03-10T08:50:01+00:00",
    "annotatedBy": { ... },
    "hasBody": {
      "@id": "oa:hasBody",
      "@type": [ "cnt:ContentAsText", "dctypes:Text" ],
      "chars": "Lorem ipsum dolor sit amet, consectetur
        adipiscing elit. Donec interdum nisi ut nibh porta,
        non viverra neque faucibus.",
      "cnt:characterEncoding": "utf-8",
      "format": "text/plain"
    },
  },
]
```


INVENIO) » REST API: RDF Quads

```
curl http://labs.invenio-software.org/api/annotations/export/ \  
-H "Content-Type: application/json" \  
--data '{"query": {"_id": "829c23e8-2c00-4bf8-b057-e8ea1d3c6b58"},  
      "new_context": {"format": "application/nquads"},  
      "ldexport": "normalized"}'
```

```
<http://labs.invenio-software.org/api/accounts/account/?id=8> <  
  http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://xmlns.  
  com/foaf/0.1/Person> .
```

```
<http://labs.invenio-software.org/api/accounts/account/?id=8> <  
  http://xmlns.com/foaf/0.1/mbox> <mailto:balthasar.montague@cds.  
  cern.ch> .
```

```
<http://labs.invenio-software.org/api/accounts/account/?id=8> <  
  http://xmlns.com/foaf/0.1/name> \"balthasar\" .
```

```
<http://labs.invenio-software.org/api/annotations/export/?_id=829  
c23e8-2c00-4bf8-b057-e8ea1d3c6b58> <http://www.w3.org  
/1999/02/22-rdf-syntax-ns#type> <http://www.w3.org/ns/oa#  
Annotation> .
```

```
<http://labs.invenio-software.org/api/annotations/export/?_id=829  
c23e8-2c00-4bf8-b057-e8ea1d3c6b58> <http://www.w3.org/ns/oa#  
annotatedAt> \"2014-03-10T08:50:01+00:00\" .
```

```
<http://labs.invenio-software.org/api/annotations/export/?_id=829  
c23e8-2c00-4bf8-b057-e8ea1d3c6b58> <http://www.w3.org/ns/oa#
```

4

Conclusions

- commenting facility extended
 - annotate any URI
 - annotate targeted document location (page, table, equation, etc)
- standard compliant
 - JSON-LD, RDF
 - W3C Open Annotation
 - REST API
- prototype tested by experimental collaborations
 - created simple page-by-page document viewer
 - testing Mozilla PDF.js; visual highlight of fragments