The CERN Multimedia Repository

Ludmila Marian
CERN

Invenio Interest Group
OR2014 - Helsinki, Finland
Outline

- Provenience of multimedia
- Additional processing
- Outreach
CERN Multimedia
- ~2’000 videos
- ~5’500 lectures
- ~15’500 photo records (~150’000 photos)
- ~500 audios
Where is multimedia coming from?
- CERN VideoLab
- CERN PhotoLab
- Lecture Recordings
- Experiments (Movies, Masterclasses, Google hangouts, Photos)
- Users
External Provenience

- VideoLab Masterclasses (Final Cut Server)
- Lectures (Micala)
- Google Hangouts (YouTube)
- PhotoLab (MediaArchive)

Connections:
- VideoLab Masterclasses to CDS via webupload
- Lectures to CDS via webupload
- Google Hangouts to CDS via custom script, bibupload
- PhotoLab to CDS via custom synchronization script, bibupload
Images

- submitted via WebSubmit;
- possibility of creating crop versions of the images.
Additional processing

- Only for images:
  - asynchronous subformat creation (3 icon sizes: 180, 640, 1440 width) - via a bibtasklet;
  - Album creation (under testing)
CDS = common store for CERN multimedia material

Easy to export
Easy to reuse
Photos - Current state
- one record - several images (one photo shoot)
  (even 300 images per record)
- Photos stored either locally (BibDocFile) or externally (MARC) or both

Videos - Current state
- one record - one video (in most of the cases)
- Videos stored externally (MARC)
New API for referencing one image:

- CFG_SITE_URL/images/{image-id} # image-id = reportnumber + counter
New API for exporting (the API used for communicating with Drupal sites):

- Using oEmbed format (used also by YouTube, Vimeo, Flickr, Instagram, ..)
- Endpoint: CFG_SITE_URL/api/oembed
- URL Scheme: CFG_SITE_URL/images/.* ; CFG_SITE_URL/video/.*
- Additional parameters: maxwidth, maxheight, format={xml, json}

```json

{  thumbnail_width: 180,
  provider_url: "http://cds.cern.ch",
  copyright: "&copy; 2007 CERN",
  title: "The first ATLAS Inner Detector End-Cap after complete insertion within the Liquid Argon Cryostat. Le premier détecteur interne bouchon d'ATLAS après son insertion à l'intérieur de son cryostat d'argon liquide",
  url: "https://mediastream.cern.ch/MediaArchive/Photo/Public/2007/0705021/0705021_01/0705021_01_a4_at-144-dpi.jpg",
  author_name: "Claudia Marcelloni",
  height: 1078,
  copyright_url: "http://copyright.cern.ch/",
  width: 1615,
  thumbnail_url: "https://mediastream.cern.ch/MediaArchive/Photo/Public/2007/0705021/0705021_01/0705021_01_icon.jpg",
  version: "1.0",
  provider_name: "CERN Document Server",
  cache_age: "10800",
  type: "photo",
  thumbnail_height: 120
}
Artistic view of the Higgs Field Image d'artiste du champs de Higgs

<author_name>Dominguez, Daniel</author_name>
<copyright>©copy; 2013 CERN</copyright>

<thumbnail_url>
<provider_url>http://cds.cern.ch</provider_url>
<cache_age>10800</cache_age>
<type>photo</type>
<thumbnail_height>120</thumbnail_height>
</oembed>
Custom exporting API (for CERN consumers) - also JSON based

- covers all multimedia material
- Endpoint: CFG_SITE_URL/api/mediaexport?id={persistent-identifier}
IBL installation into the inner detector of the ATLAS Experiment side C

Condition of Use: ATLAS Experiment © 2014 CERN

Photos: Marcelloni De Oliveira, Claudia
Date: 15-03-2014
A new subdetector for ATLAS

The problem given was that with higher luminosity in the LHC's next run, significant radiation damage to the inner layers of the detector could occur, which meant ATLAS would lose tracking efficiency, especially in tagging the decays of the beauty quark — crucial for physics analysis. The idea was to substitute a thick brick-on-invertable layer instead of replacing the existing B-layer in the Paul Detector. The IB was born and it was the only way to integrate it into the detector by shrinking the diameter of the beam pipe and inserting it into the gap between the Paul Detector and the pipe. This was a new feature in the inner detector region of ATLAS, an additional pair for tracking particles. More points mean better precision which is always good for physics.

Making space wasn’t the only challenge for the IB project. Much of the technology did not exist. Increased luminosity in the LHC meant the IB has to cope with high radiation and high particle occupancy because of its proximity to the particle interaction point in the beam pipe. This also meant the number of bits on the detector and the amount of data collected will increase substantially. Faster readout chips and two different silicon sensor technologies were developed.

Paul saw the need to get rid of split modules, and a COTS-based readout system was introduced as opposed to the O2P. New active-scan structures were invented to support the modules that make up the IB. These scans had to be just fine enough to serve as mechanical support but flexible enough to be inserted.

As remarkable as the developments were, even more remarkable was the collaborative nature of the project. Projects like ATLAS require the involvement of everyone, and the IB was no exception. The science, as does everything else in ATLAS, depended on the members.

"The atmosphere in the cavern during the insertion was quite lively," says Martin Hildal, who together with Raphael Vallone, coordinated the engineering and construction. "There was a lot of excitement because of the many practice sessions, but more importantly, there was a lot of trust."

Posted by Dan O'Loughlin on 04 May 2014. Last updated 21 May 2014, 18.19.
The problem was that with higher luminosity in the LHC's first run, significant radiation damage to the inner layers of the detector could occur, which meant ATLAS would lose tracking efficiency, especially in tagging the decay of the beauty quark – crucial for physics analyses. The idea was to introduce thin X-strip sensors in an additive layer instead of replacing the existing B-layer in the Pixel Detector. The IBL has been shown to be the only way to integrate it by shrinking the diameter of the beam pipe and inserting it into the gap between the Pixel Detector and the pipe.

The IBL is now the new fourth layer in the inner detector region of ATLAS, an additional layer for tracking particles. More points mean better precision which is always good for physics.

Making space wasn’t the only challenge for the IBL project. Tools of the technology did not exist. Increased luminosity in the LHC meant the IBL had to cope with high radiation and higher particle occupancy because of its proximity to the particle interaction point in the beam pipe. This also meant the number of hits on the detector and the amount of data collected will increase substantially. Faster read-out chips and two different silicon sensor technologies were developed. For IBL was reduced to 30 by 128 micrometers and a CMOS-based readout system was introduced as opposed to the CDSR. New coating bump structures were invented to support the modules that make up the IBL. These were built to be just fine enough to serve an mechanical support but flexible enough to be inserted.

As remarkable as the developments were, even more remarkable is the collaborative nature of the project. Forty-seven institutes from 14 countries were involved in the IBL team. In reality, as does everything else in ATLAS, depended on the members. "The achievement in the IBL during the insertion was pivotal," says Silvestre Wittels, who together with Raphael Valleron, coordinated the engineering and construction. "It was a lot of confidence there because of the many practice sessions, but more importantly, there was a lot of trust."

Posted by Dan 0'Loughlin on 20 May 2014, last updated 21 May 2014, 16:28.
**New API for exporting/embedding image slideshows**

- `CFG_SITE_URL/images/{recid}/export?format={format_type}`
- `CFG_SITE_URL/images/keywords/{keyword}?format={format_type}`
- `format_type = {slideshow, sspp, espp}` # different transition types

<iframe width="480" height="360" scrolling="no" frameborder="0" src="https://cds.cern.ch/images/CERN-GE-1311274/export?format=sspp" allowfullscreen></iframe>

<iframe width="200" height="200" scrolling="no" frameborder="0" src="https://cds.cern.ch/images/keywords/CERN%20PHOTOWALK?format=slideshow&w=200&h=200" allowfullscreen></iframe>
New API for embedding a video

- CFG_SITE_URL/video/{video-id} # video-id = report number (persistent identifier)

```html
<iframe src="http://cdsweb.cern.ch/video/CERN-MOVIE-1967-001" allowfullscreen="" frameborder="0" height="360" width="480"></iframe>
```

- Improves:
  - **Integration and accessibility** (device-specific embedding code)
  - **Persistence** (the URL in the embedding code will never change even if record id will change, streaming get’s updated, slaves get regenerated)
  - **Sharing and connectivity** (inside the embedding video frame: sharing tools and related videos)
Embedding of Videos

IT Department @ CERN OpenDays 2013

Welcome to the Information Technology Department!

As part of the CERN OpenDays, the IT Department literally opened the doors to its Data Centre so you could discover more about our role in discoveries and research at CERN and beyond.

We have compiled here all the material on display during the OpenDays, as well as some more information to help you further explore some of the things you may have seen.

The Itinerary

After the video, there were separate guided tours of the 2 main computing rooms - one in our immense hall of racks and servers, and the other down into the basement to walk through our data storage vault.

In addition to the main visit of the Data Centre, some demonstrations of our computing took place next to the Data Centre. There was an opportunity to find out more about Grid computing and talk to some of its experts. Visitors could immerse themselves in our special Liquid Galaxy ‘surround screen’ and go on a Google Virtual Tour of the Data Centre.

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Data Centre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projects</td>
<td></td>
</tr>
<tr>
<td>Meeting catalog</td>
<td></td>
</tr>
<tr>
<td>OpenDays 2013</td>
<td></td>
</tr>
<tr>
<td>What We Do</td>
<td></td>
</tr>
<tr>
<td>The Data Centre</td>
<td></td>
</tr>
<tr>
<td>Birth of the Web</td>
<td></td>
</tr>
<tr>
<td>Grid computing</td>
<td></td>
</tr>
<tr>
<td>Media</td>
<td></td>
</tr>
</tbody>
</table>
Push to YouTube

Demo
Push to YouTube

CMS animation of the high-energy collisions at 7 TeV on 30th March 2010
by CMS team © 2010 CERN
Run 132440 - Event 2732271
Upload video to youtube
Push to YouTube

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

**Description**
Run 132440 - Event 2732271

**Keywords**
LHC, CMS, LHCfirstphysics

**Privacy**
Private

**Category (YouTube)**
Education

Please note that the upload process sometimes it can take up to several minutes.
The video successfully uploaded with id: k4zn2z8v10k

Please note that the upload process sometimes it can take up to several minutes.
Way to go, Ludmila Marian!
Looks like you uploaded your first ever YouTube video.

CMS animation of the high-energy collisions at 7 TeV on 30th March ...
Hope you enjoy your video, and welcome to the community!
Your latest fan, The YouTube team
Copyright and License

- CERN is progressively moving from CERN license to CC-BY-SA license for photos (and videos);
- Pilot project (89 images so far);
- Current work on consolidating the copyright/license usage across Invenio:
  - better support at the level of BibDocs (each BibDoc can have its own copyright&license); can be updated via the interface.
  - better support in the MARC (several copyright&license statements supported: for the record, for each of the files attached);
  - updating EXIF (for photos) with the copyright&licence from MARC;
Photos

- Current state
  - one record - several images (one photo shoot) *(even 300 images per record)*
  - Photos stored either locally (BibDocFile) or externally (MARC) or both

- Shortcomings of the current model
  - Needs workaround for referencing one image *(http://cds.cern.ch/images/CERN-EX-0705021-01)*
  - Not possible to add per image metadata (tags, keywords, description, comments, ..), to add to baskets (possible only as external links)
  - .. (not possible to reference each image as an object)
New document model: Albums

- already (slightly) in use for Videos
- ex: http://cds.cern.ch/record/1606846 (MARC for TOC, MARC for Asset)
1 photo per record (main photo + slaves)

1 record representing the Album, containing references of the photos + possible order in the Album

the photo record will be a 1st class citizen: baskets, comments, sharing, exporting = out of the box with Invenio

also many improvements:

- more flexible and customizable slideshow embedding for albums
- photo overlays
- widget framework: for easy addition of new features
- easy creation of new Albums by selecting photos from existing Albums
- read/write EXIF (at submission time, for pre-filling the form and exporting metadata with the photo)
Photo Albums

Servicing the first web server - Tim Berners-Lee's NeXT

In August 2009 a team from the Association aBCM in Lausanne came to CERN to give the world's first web server a health check under the watchful eye of web pioneer Robert Caillau. They took an image of the hard drive at this time, copies of which were given to Robert Caillau and Tim Berners-Lee.

21 photos

Date: 2009
Photographer: A
Keywords: Computing, World Wide Web, NeXT, Computers and Control Rooms
Photo Albums

**Settings**

- **Width**: 500
- **Height**: 320
- **Transition Speed**: 800
- **Color**: #000000
- **Autoplay Speed**: 12000

**Autoplay**: Yes

**Show Navigation**: Yes

**Transition Effect**: Slide

```
<iframe width="500" height="320" src="http://ubuntu13/media/gallery?format=gallery&album_id=100&options=width=500, height=320, transitionSpeed=800, color=%23000000, autoplaySpeed=12000, enableAutoplay=true, showNavigation=true, effect=slide, scrolling="no" frameborder="0"></iframe>
```
Ongoing R&D

- Image similarity
  - finding exact match (disregarding image processing like cropping, rotation, resizing) - for better interlinking between photos and records using the photos (Bulletin, Currier, etc.)
  - finding visual similar images (content based image retrieval) - for improving the image search

- How
  - LIRE (Lucene Image REtrieval) for feature computation - not enough
  - combine features, CLUE, SIFT-like features
Image Tagging

- offer the possibility of tagging people/objects on an image;
- face recognition;
- exporting the tags with the image;
Questions?

info@invenio-software.org

 cds.support@cern.ch

 ludmila.marian@cern.ch