

DocDB for CMS

Communications Meeting
Eric Vaandering
FNAL
29 April 2010

Goal

- Collect the O(100K) CMS documents not covered by
 - EDMS
 - Indico
 - Papers and conference reports
- 100K documents? Really?!
 - Document has a very loose definition: something you create you want to share with others: memos, video documenting a procedure, etc.
 - Yes, really: Experience shows ~5-10 documents/year/collaborator
- Harvest documents now, transfer with metadata to CDS later
 - Depends on CDS interface development
 - DocDB takes a user ~60 seconds to create a new document

Overview

- Long history (since 2001) and used by most experiments at FNAL (and several outside of FNAL)
- It is a “collaborative document catalog”
 - Maintains documents (collections of files) and meta-data
 - Topics, authors, publication references, events, document type
 - Documents are versioned and agnostic about file content
 - Metadata hyperlinked to searches
 - Collaborative means permissions are group based
 - Users are members of one or more groups (HCAL, Higgs, Communications)
 - Some groups can be read-only (e.g. reviewers)
 - Each document has a list of groups who can read, who can modify
- Some experience within CMS already
 - Pixels since 2005
 - BPM, PLT, DCS more recently

Current Status

- DocDB is fully deployed at <https://cms-docdb.cern.ch/cgi-bin/DocDB/DocumentDatabase>
 - Thanks to FacOps for several rounds of support
- Migrated existing Pixel, DCS, etc. content into this server
 - Some work reorganizing topic hierarchy, proper metadata on docs
- Integrated with CERN Single Sign On
 - Depends on e-groups, all “cms-users” have access, smaller e-groups mapped onto DocDB's groups
- Topic structure defined for ECAL as next “user”
 - Use experience for other groups going forward

Future Developments

- DocDB has its own integrated agenda creator
 - Disable this, keeping event metadata field
 - Some integration with Indico if possible
- Meta-data entry scales well up to ~few hundred collaborators
 - May be too cumbersome for ~few thousand working on diverse topics
 - Some planning for this by associating topics, etc, to groups.
- A few other refinements for collaborations the size of CMS
 - Code is well tested and mature, so should not need much

Demo

- Take a look at some of the documents not covered by existing systems
- Click around the interface
- Demonstrate updating an existing document
- <https://cms-docdb.cern.ch/cgi-bin/DocDB/DocumentDatabase>
- Search on “installation video” and “specification”