

Workflows and holding pen

Jan Åge Lavik
CERN

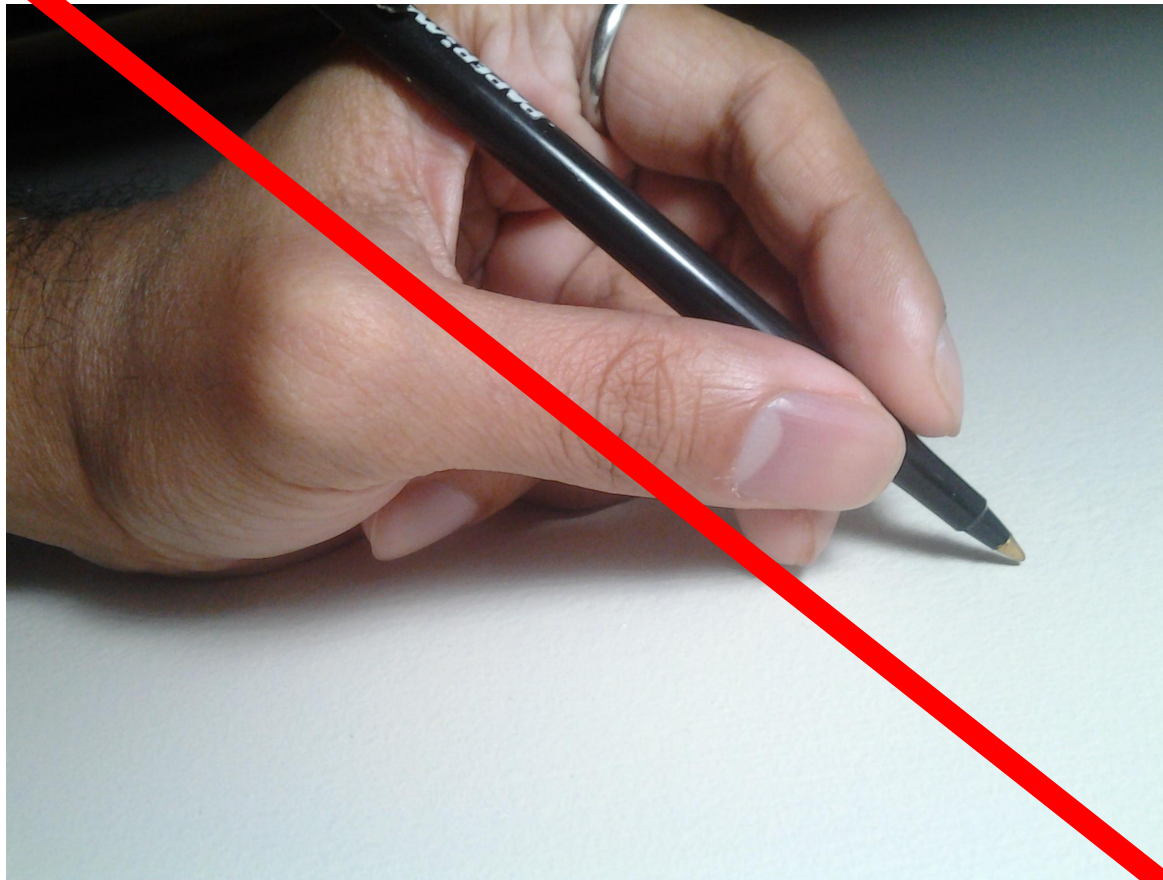
Invenio User Group Workshop
Jülich, Germany
17–20 November 2013

Holding pen?

First result in Google



First result in Google



holding pen

a pen where livestock is temporarily confined



holding pen

records 😊

a pen where ~~livestock~~ is temporarily confined



holding pen

records 😊

a pen where ~~livestock~~ is temporarily confined



**So, what about the
current Holding Pen?**

Currently too hidden

(Under: Home / Admin / OAI Harvest)

OAI Harvest Admin Interface

Overview of sources [\[?\]](#) [\[add new OAI source\]](#)

6 OAI sources currently present in the database

name	base URL	metadata prefix	frequency	last run	post-processes	actions
arxiv	http://export.arxiv.org/oai2	arXiv	Not Set	2013-11-16 04:00:12	c-p-a-f-u	edit / delete / test / history / harvest

Harvesting status [\[?\]](#)

Next oaiharvest task

- scheduled time: 2013-11-17 04:00:00
- current status: WAITING

[View Holding Pen](#)

Lack of features

Holding Pen

Show entries containing :

[Year 2013 \(23239 entries\)](#)

[2013-2 \(2461 entries\)](#)

[2013-2-7 \(298 ENTRIES\)](#)

OAI Record ID	Insertion Date			
oai:arXiv.org:1006.3068	2013-02-07 06:03:58	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1007.4820	2013-02-07 06:03:58	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1101.0520	2013-02-07 06:03:58	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1106.2197	2013-02-07 06:03:58	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1109.3979	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1110.3816	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1110.6535	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1112.5435	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1201.3609	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1201.6658	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1202.1946	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1203.0151	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1203.6069	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1204.1053	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen

Limited search/filter

Holding Pen

Show entries containing :



[Year 2013 \(23239 entries\)](#)

[2013-2 \(2461 entries\)](#)

[2013-2-7 \(298 ENTRIES\)](#)

OAI Record ID	Insertion Date			
oai:arXiv.org:1006.3068	2013-02-07 06:03:58	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1007.4820	2013-02-07 06:03:58	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1101.0520	2013-02-07 06:03:58	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1106.2197	2013-02-07 06:03:58	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1109.3979	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1110.3816	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1110.6535	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1112.5435	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1201.3609	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1201.6658	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1202.1946	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1203.0151	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1203.6069	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen
oai:arXiv.org:1204.1053	2013-02-07 06:03:59	Compare with original	Apply changes	Delete from holding pen

Good BibEdit integration

	<input type="checkbox"/> a	arXiv:1006.3068 Same content	
	<input type="checkbox"/> c	astro-ph.CO Same content	
<input type="checkbox"/> 041__	<input type="checkbox"/> a	language (if not English)	
<input type="checkbox"/> 084__	<input type="checkbox"/> a	pacs	
	<input type="checkbox"/> 2	PACS	
<input type="checkbox"/> 100__	<input type="checkbox"/> a	Andersson, K. Same content	
	<input type="checkbox"/> u	Affiliation	
<input type="checkbox"/> 242__	<input type="checkbox"/> a	Translated title (English)	
<input type="checkbox"/> 245__	<input type="checkbox"/> a	X-ray Properties of the First SZE-selected Galaxy Cluster Sample from the South Pole Telescope	
<div style="border: 1px solid #ccc; padding: 10px; background-color: #fff9c4;"> <p>Field structure has changed. New value:</p> <pre> 245__ \$\$9 arXiv \$\$a X-ray Properties of the First SZE-selected Galaxy Cluster Sample from the South Pole Telescope </pre> <p><input type="button" value="✓"/> <input type="button" value="✗"/> <input type="button" value="+"/> <input type="button" value="Apply"/></p> </div>			
<input type="checkbox"/> 246__	<input type="checkbox"/> g	arXiv	
	<input type="checkbox"/> a	X-ray Properties of the First SZE-selected Galaxy Cluster Sample from the South Pole Telescope	
<input type="checkbox"/> 260__	<input type="checkbox"/> c	publication date YYYY-MM-DD	
<input type="checkbox"/> 269__	<input type="checkbox"/> c	2010-06	
		2010-06-15 <input type="button" value="✓"/> <input type="button" value="✗"/> <input type="button" value="+"/> <input type="button" value="Apply"/>	
<input type="checkbox"/> 300__	<input type="checkbox"/> a	28 Same content	
<input type="checkbox"/> 500__	<input type="checkbox"/> a	* Brief entry *	
<input type="checkbox"/> 500__	<input type="checkbox"/> g	arXiv Same content	
	<input type="checkbox"/> a	28 pages, 19 figures, submitted to ApJ Same content	
<input type="checkbox"/> 500__	<input type="checkbox"/> a	*Brief entry* Same content	

Focus points

- More accessible (web/api/cli)
- Highly customizable (workflows, widgets)
- User integration (“Your holding pen”)

Work in progress

Records in Holding Pen

[Home](#)
[Assigned to me](#)
[Pending Tasks ▾](#)
[Refresh!](#)
[Resolve selected](#)

10 records per page
 Select Columns
Search:

Title	Author	Source	Category	Created	Status	Details	Actions
Two-colour QCD with heavy quarks at finite densities	Langfeld, Kurt	arXiv	hep-lat	13 minutes ago	Halted	Details	Approve Record
Zeros of QCD partition function from finite density lattices	Jin, Xiao-Yong	arXiv	hep-lat	13 minutes ago	Halted	Details	Approve Record
Bottomonium spectrum at finite temperature	Harris, Tim	arXiv	hep-lat	13 minutes ago	Halted	Details	Approve Record
Non-perturbative treatment of strong interactions in field theories	Trachenko, K	arXiv	hep-th	13 minutes ago	Halted	Details	Approve Record
Eigenvalue spectrum of lattice $\mathcal{N}=4$ super Yang-Mills	Weir, David J	arXiv	hep-lat	13 minutes ago	Halted	Details	Approve Record
Scaling, topological tunneling and actions for weak coupling DWF calculations	McGlynn, Greg	arXiv	hep-lat	13 minutes ago	Halted	Details	Approve Record
The Gradient Flow Coupling in Minimal Walking Technicolor	Rantaharju, Jarno	arXiv	hep-lat	13 minutes ago	Halted	Details	Approve Record
Towards the continuum limit in transport coefficient computations	Francis, A	arXiv	hep-lat	13 minutes ago	Halted	Details	Approve Record

How did we get there?

Early ideas

- Re-using core features (search, indexing)
- Sense of urgency
- Old methodologies (2.0 not yet born)

Idea #1: Special Holding Pen collection

- Taking up/fragmenting “record space”
- How to store technical information?
MARC?
- Prototyped with special Holding Pen indexes
- Didn't feel right..

Idea #2: Re-using baskets

- Interesting, outside of record store
- Felt like “trying to fit a square in a circle”
- Needed something specific

Back to the origins: What is ingestion?

Ingestion workflows

Needed a workflow module

BibWorkflow aka 'workflows' (2.0)

- New module (off “workflow 1.01” lib)
- Tasks and workflows in pure Python
- Interruptible/resumable workflows
- Backbone of new modules
(deposition/holding pen)

Example task:

```
def add_data(a):  
    """ Task with parameters """  
    def _add_data(obj, eng):  
        obj.data = obj.data + a  
    return _add_data
```

Example workflow:

```
class my_workflow(object):  
    workflow = [  
        add_data(1),  
        add_data(2),  
    ]  
    title = "Sample workflow"
```


API

```
In [1]: from invenio.modules.workflows.api import start
```

```
In [2]: start("sample_workflow", data=[1])
```

```
INFO 2013-11-17 16:46:21,685 workflow.67d8e272-4f9f-11e3-b558-0024d7bf88e0 Workflow saved to db.  
INFO 2013-11-17 16:46:21,700 workflow.67d8e272-4f9f-11e3-b558-0024d7bf88e0 Workflow has been started  
INFO 2013-11-17 16:46:21,708 object.67d8e272-4f9f-11e3-b558-0024d7bf88e0_2 Object is selected for  
processing  
INFO 2013-11-17 16:46:21,718 object.67d8e272-4f9f-11e3-b558-0024d7bf88e0_2 Object processing is finished  
INFO 2013-11-17 16:46:21,727 workflow.67d8e272-4f9f-11e3-b558-0024d7bf88e0 Done saving object: 2  
INFO 2013-11-17 16:46:21,739 workflow.67d8e272-4f9f-11e3-b558-0024d7bf88e0 Workflow saved to db.
```

```
Out[2]: <BibWorkflow_engine(sample_workflow)>
```

```
In [3]: [obj.data for idx, obj in _.getObjects()]
```

```
Out[3]: [4]
```

```
start_delayed() # asynchronous execution
```

Workers

- Distribute workflows
- Asynchronous execution
- Plugin based wrappers
- Celery & RedisQueue implemented



```
start_delayed() # asynchronous start()
```

Holding Pen as “view”

- Overview of objects that require attention
- Marked for a certain “action” or contains errors

Category	Created	Status	Details	Actions
hep-lat	13 minutes ago	Halted	Details	Approve Record

```
eng.halt("Needs human intervention")
```

Special “HP Widgets”

Predictions for LHC from SO(4) MWT

Hietanen, Ari ; Pica, Claudio ; Sannino, Francesco ; Søndergaard, Ulrik

2013

Abstract: We investigate the vector, axial and pseudo scalar mass spectrum of an SO(4) - MWT gauge theory with fermions in the vector representation of SO(4). Here we present the preliminary lattice results for the masses of vector and axial vector meson using Wilson fermions. These spectra are crucial for the discovery and to guide the searches of composite dynamics at the LHC.

Note: Comments: 7 pages, 5 figures, presented at the 31st International Symposium on Lattice Field Theory (Lattice 2013), 29 July - 3 August 2013, Mainz, Germany

Reject

Accept

Reject

Accept

- Resolve halted objects via customized UI
- **Main** and **quick** versions

```
eng.halt(msg="Needs approval",  
         widget="approval_widget")
```



Small demo

Next steps..

- UI improvements (requirements, users)
- Set of building blocks (widgets, tasks)
- Full ingestion workflows