



Material Assay Database

Structure & Organization

James Loach, LBNL

This talk

- Design principles
- Database engine
- Data format
- User interface
- Implementation

Design principles

Data format

- Concise and structured
- Comprehensive and flexible

Database engine

- Open source
- Lightweight

User interface

- High quality
- Powerful search and display
- User/admin model

The screenshot shows a web application interface with a navigation bar at the top containing links: Search, Submit, Feedback, Instructions, and Preferences. Below the navigation bar is a search input field containing the text 'rock'. To the right of the input field are three buttons: a magnifying glass icon, 'Detail', and 'Expand'. Below the search bar, a list of search results is displayed, each preceded by a small square icon with a plus sign. The results are:

- Homestake mine, country rock with quartz, 4850L
- Homestake mine, country rock, 4850L
- Homestake mine, rhyolite dike rock, 4850L
- Homestake mine, rhyolite dike rock, 4100L
- Homestake mine, rhyolite dike rock, 1250L

MAJORANA / LBNL LBF have developed a
system on these principles

CouchDB



Open source non-relational database

- Stores a flat collection of JSON documents

```
"sample": {  
  "name":      "Fused silica",  
  "description": "Corning 7940, lot 56667",  
  "source":    "Mark Optics Ltd.",  
  "owner":     "LBNL LBF",  
}
```

- **Scheme free** so document structure can be vary
- Data aggregated and displayed with **views**
- Web applications can be stored as documents
- Distributed
- Interact with database via HTTP with Python, PERL, C++ etc.

Search



Cloud hosting



See the wikipedia article...

Management interface (native)

Viewing documents

The screenshot displays the CouchDBX management interface, titled "CouchDBX — Time to Relax". The main content area shows an "Overview" section with a table of databases:

Name	Size	Number of Documents	Update Seq
_users	20.1 KB	1	5
aarm	4.2 MB	13	36
assays_backup	9.5 MB	46	47
mj_assays	11.1 MB	46	60
test	8.1 KB	0	2

Below the table, it says "Showing 1-5 of 5 databases" and provides navigation links: "← Previous Page", "Rows per page: 10", and "Next Page →".

On the right side, there is a sidebar with the CouchDB logo and the text "relax". Below this, there is a "Tools" section with links to "Overview", "Configuration", "Replicator", "Status", "Test Suite", and "Hosting via Couchio". There is also a "Recent Databases" section listing "aarm". At the bottom of the sidebar, it says "Welcome to Admin Party! Everyone is admin. [Fix this](#)" and "Futon on Apache CouchDB 1.0.1".

The bottom part of the screenshot shows a detailed view of a document. It has a header with "Fields" and "Source" tabs. The "Fields" tab is active, showing a table with "Field" and "Value" columns. The document is a JSON object with the following fields:

- _id**: "cbf5d92e933784e6f89ede2c31001959"
- _rev**: "2-a608ff79125b395d011c22c44549636a"
- data_source**: reference "HOMESTAKE4-Concrete.doc", data_entry_name "James Loach", data_entry_contact "jcloach@lbl.gov"
- measurement**: institution "LBNL", technique "Gamma", date, requestor "n/a", requestor_contact "n/a", practitioner "Al Smith", practitioner_contact "ARSmith@lbl.gov", description "This document summarizes measurements made at the Lawrence Berkeley National Laboratory (LBNL) Low Background Facility (LBF) in preparation ...", results: count_length "58201 sec", data_file "23838", detector "MERLIN(BKY)"
- sample**: name "Homestake mine, rhyolite dike rock, 4850L", description "HST-18, rhyolite dike at 4850L, 5/09"

Futon main window

Management interface (native)

The screenshot shows the CouchDBX management interface titled "CouchDBX — Time to Relax". The main area displays an "Overview" of databases with the following table:

Name	Size	Number of Documents	Update Seq
_users	20.1 KB	1	5
aarm	4.2 MB	13	36
assays_backup	9.5 MB	46	47
mj_assays	11.1 MB	46	60
test	8.1 KB	0	2

Below the table, it says "Showing 1-5 of 5 databases" and provides navigation links: "← Previous Page", "Rows per page: 10", and "Next Page →".

On the right side, there is a sidebar with the CouchDB logo and the text "CouchDB relax". Below this, there is a "Tools" section with links: "Overview", "Configuration", "Replicator", "Status", "Test Suite", and "Hosting via Couchio". Under "Recent Databases", "aarm" is listed. At the bottom of the sidebar, it says "Welcome to Admin Party! Everyone is admin. [Fix this](#)" and "Futon on Apache CouchDB 1.0.1".

In the foreground, a "Replication" window is open. It has two sections: "Replicate changes from:" and "to:". Both sections have a radio button for "Local database:" (selected) and a text input for "Remote database:" (containing "http://"). A double-headed arrow connects the two sections. At the bottom right of the window, there is a checkbox for "Continuous" and a "Replicate" button. Below the replication window, there is an "Event" section with the text "No replication".

Replication

Futon main window

Management interface (cloud)

www.cloudant.com

(also see www.couchbase.com)

The screenshot shows the Cloudant management interface for a database named "lbf / assays". At the top, the database URL is displayed as "https://lbf.cloudant.com/assays/" with a "Copy" button. Below this are several action buttons: "Data", "Stats", "Search" (highlighted with a yellow border and a "NEW" tag), "Permissions", "View in Futon", and a red "Delete database" button.

The main section is titled "All documents". It includes a dropdown menu for "All documents" and a search bar labeled "search by key". Below the search bar, a document is selected, showing its key "050f40be82ad8c96fb3dea03". To the right of the key are links for "View source", "Reload", "View raw", and "Delete document".

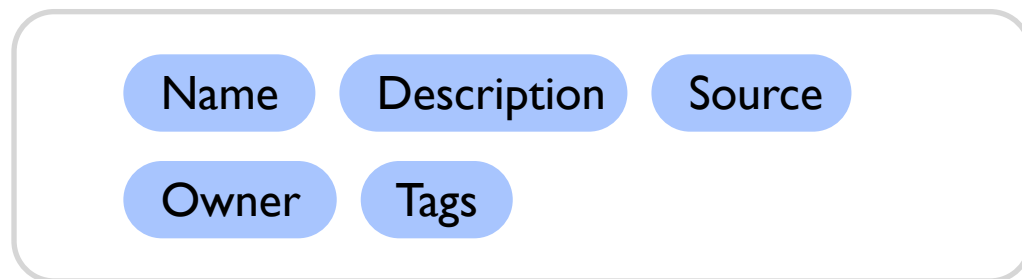
The document's metadata and fields are displayed in a list format:

- `_id` : 050f40be82ad8c96fb3dea03153fb49e
- `_rev` : 1-572983d5eb079b7aac402520666d7ddf
- `date` : 26/03/2008
- `facility` : LBNL
- `detector` : MERLIN (BKY)
- `responsible_person` : Al Smith
- `requester` : Kam-Biu Luk
- `requester_email` : K_Luk@lbl.gov

What is an assay?

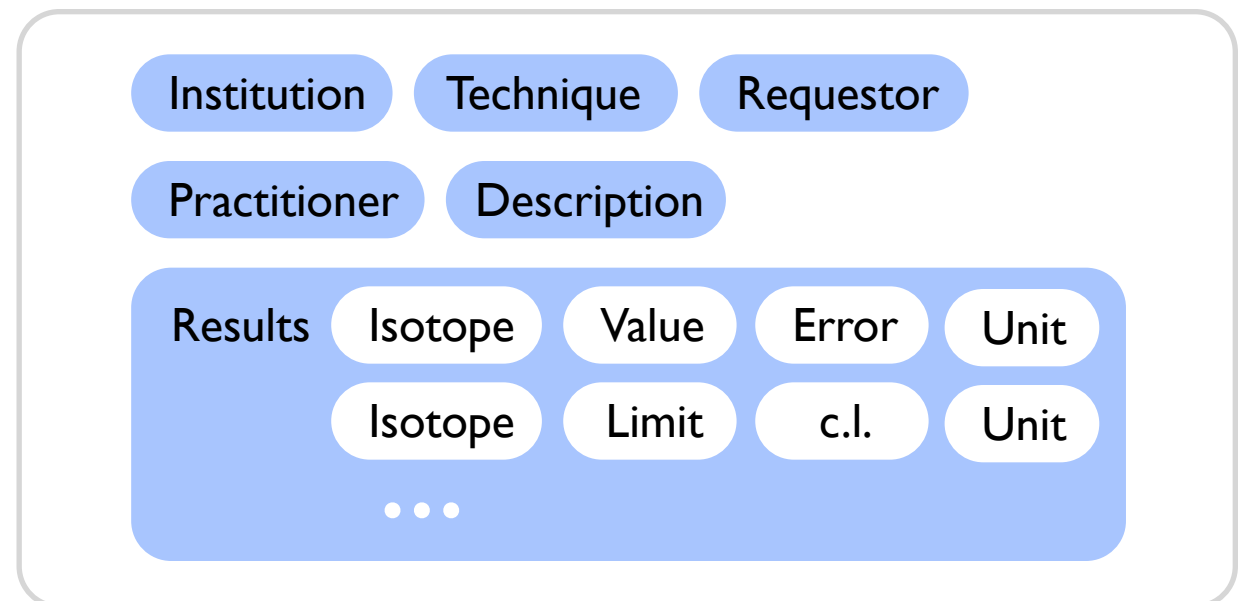
Sample

The thing that is being counted



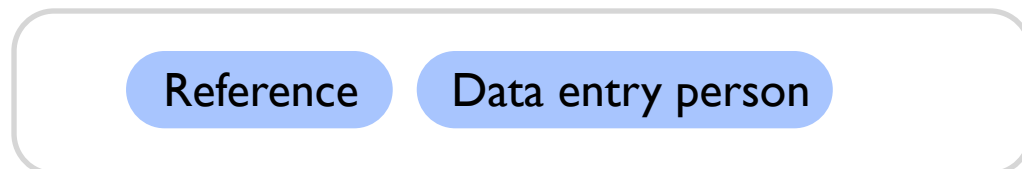
Measurement

The measurement and its results



Data source

Where the data came from and who entered it



This concept must be expressed in terms of fields and rules

Fields & rules

Fields

```
{  
  "type": "measurement",  
  "sample": { },  
  "measurement": { },  
  "data_source": { }  
}
```

```
"sample": {  
  "name": "",  
  "description": "",  
  "source": "",  
  "owner": "",  
  "tags": [ "", " ", " ", ... ]  
}
```

```
"measurement": {  
  "institution": "",  
  "technique": "",  
  "date": "",  
  "requestor": "",  
  "requestor_contact": "",  
  "practitioner": "",  
  "practitioner_contact": "",  
  "description": "",  
  "results": [  
    { "isotope": "", "value": "", "error": "", "unit": "" },  
    { "isotope": "", "limit": "", "c.l.": "", "unit": "" }  
  ]  
}
```

Rules

Date	YYYY-MM-DD
Isotope name	238U, U-238
Unit	ppm, ppt, mBq/kg, ...
Contact	name@email.com

```
"data_source": {  
  "reference": "",  
  "data_entry_name": "",  
  "data_entry_contact": ""  
}
```

JSON data format

Extendability

In non-relational databases
the field structure need not be fixed

```
"sample": {  
  "name": "",  
  "description": "",  
  "source": "",  
  "owner": "",  
  "tags": [ "", "", "", ... ]  
}
```

```
"sample": {  
  "name": "",  
  "description": "",  
  "source": "",  
  "owner": "",  
  "tags": [ "", "", "", ... ],  
  "geometry": "",  
  "mass": "",  
  "color": "",  
  "storage box": ""  
}
```

Core

Extension

But the best way to harness this power?

Option A : Extendable format

Option B : Extendable specification

Option C : Filtered replication

User interface

Viewer

- Search form
- Submission form
- Flexible data display
- Data export

Management

- Approval
- Editing / deleting

Write in HTML/JavaScript

Store as a CouchDB document

Multiple interfaces are allowed,
provided they respect the data specification



Searching

Search

Submit

Feedback

Instructions

Preferences

Detail

Expand

Persephone v0.2 (demo for AARM)

[Search](#) [Submit](#) [Feedback](#) [Instructions](#) [Preferences](#)

[🔍](#) [Detail](#) [Expand](#)

- ☐ Homestake mine, country rock with quartz, 4850L
- ☐ Homestake mine, country rock, 4850L
- ☐ Homestake mine, rhyolite dike rock, 4850L
- ☐ Homestake mine, rhyolite dike rock, 4100L
- ☐ Homestake mine, rhyolite dike rock, 1250L

Persephone v0.2 (demo for AARM)

Search

Submit

Feedback

Instructions

Preferences

rock

Detail

Expand

Homestake mine, country rock with quartz, 4850L

Sample	Description	HST-19-7/09 (sample 2), country rock, 1/3 to 1/2 quartz, 4850 vent drift		
	Results	U	0.77 (0.01)	ppm
		Th	1.59 (0.03)	ppm
		K	0.92 (0.01)	pct

Homestake mine, country rock, 4850L

Sample	Description	HST-19-7/09 (sample 1), country rock, 4850 vent drift		
	Results	U	4.42 (0.02)	ppm
		Th	8.76 (0.06)	ppm
		K	2.49 (0.01)	pct

Homestake mine, rhyolite dike rock, 4850L

Sample	Description	HST-18, rhyolite dike at 4850L, 5/09		
	Results	U(early)	8.58 (0.01)	ppm
		U(late)	8.16 (0.04)	ppm
		Th	10.59 (0.01)	ppm
		K	3.97 (0.02)	pct
		Eman	5	pct

Homestake mine, rhyolite dike rock, 4100L

Sample	Description	HST-17, rhyolite dike at 4100L (near Yates Shaft), 4/09		
	Results	U(early)	9.53 (0.12)	ppm
		U(late)	8.57 (0.05)	ppm
		Th	11.4 (0.1)	ppm
		K	7.60 (0.02)	pct

Search

Submit

Feedback

Instructions

Preferences

rock

Detail

Expand

☐ Homestake mine, country rock with quartz, 4850L

Sample	Description	HST-19-7/09 (sample 2), country rock, 1/3 to 1/2 quartz, 4850 vent drift
	Source	Homestake mine
	Owner	LBNL
	Tags	Homestake samples
	Geometry	S6MB (full)
Measurement	Technique	Gamma
	Institution	LBNL
	Date	7 / 2009
	Requestor	n/a (n/a)
	Practitioner	Al Smith (ARSmith@lbl.gov)
	Description	This document summarizes measurements made at the Lawrence Berkeley National Laboratory (LBNL) Low Background Facility (LBF) in preparation for converting the Homestake Mine facilities into a deep underground experimental facility for the Nuclear Sciences. All analyses have been performed using a high-resolution HPGe detector gamma-ray spectrometer, to identify and quantify all gamma-emitters in sample materials. Except for surface samples which may contain radionuclides from mid 20th century atmospheric nuclear weapons testing, the gamma-emitters of significance are the natural terrestrial radionuclides (U,Th,K): the uranium series, the thorium series, and potassium. The following list summarizes results obtained from bulk samples collected from the underground workings of the mine as these areas have become accessible, starting in September 2007.
	Count length	82801 sec
	Data file	24113
	Detector	MERLIN(BKY)
	Results	U 0.77 (0.01) ppm Th 1.59 (0.03) ppm K 0.92 (0.01) pct
Data	Reference	HOMESTAKE4-Concrete.doc
	Entry by	James Loach (jcloach@lbl.gov)

☐ Homestake mine, country rock, 4850L

Sample	Description	HST-19-7/09 (sample 1), country rock, 4850 vent drift
--------	-------------	---

Search
Submit
Feedback
Instructions
Preferences

Submit for approval
Check
Clear warnings
Clear form

Sample

Name	<input type="text" value="Brief description"/>	
Description	<input type="text" value="Detailed description"/>	
Source	<input type="text"/>	
Owner	<input type="text"/>	
Tags	<input type="text" value="Tags separated by spaces"/>	
Mass	<input type="text"/>	
Geometry	<input type="text"/>	

Measurement

Technique	<input type="text"/>	
Institution	<input type="text" value="Where it was counted"/>	
Date	<input type="text" value="mm/dd/yyyy"/>	
Requester	<input type="text" value="Name"/>	<input type="text" value="Email or institution"/>
Practitioner	<input type="text" value="Name"/>	<input type="text" value="Email or institution"/>
Description	<input type="text" value="Detailed description"/>	

[Search](#) [Submit](#) [Feedback](#) [Instructions](#) [Preferences](#)

[Send feedback](#)

Feedback

Name

Full name

Email

Email address

Comment

Feedback

[Search](#)[Submit](#)[Feedback](#)[Instructions](#)[Preferences](#)

Searching

Search returns documents containing one or more of the search terms.

You can alter this default behavior using wildcards and operators such as:

"rhyolite dike rock"
rhyolite AND rock
rhyolite OR rock
4??0L

By default results are presented in a concise form. Click 'Detail' to show the full detail.

Enter 'all' to show all documents.


Submitting

Data should be entered into the search form as indicated.

Grayed out field names are optional.

Submitted data is not immediately searchable. The moderator must sign off on each document.

Repository

 nepahwin / persephone

[Watch](#) 2 [Fork](#) 2

[Code](#) [Network](#) [Pull Requests](#) 1 [Issues](#) 1 [Wiki](#) [Graphs](#)



A material assay database for the low-background physics community — [Read more](#)

[Clone in Mac](#) [ZIP](#) [HTTP](#) [Git Read-Only](#) [Read-Only access](#)






[branch: master](#) [Files](#) [Commits](#) [Branches](#) 1 [Tags](#) [Downloads](#)

🕒 Latest commit to the **master** branch

Change format of the README file

 nepahwin authored 3 months ago  commit 2fbc7f1424

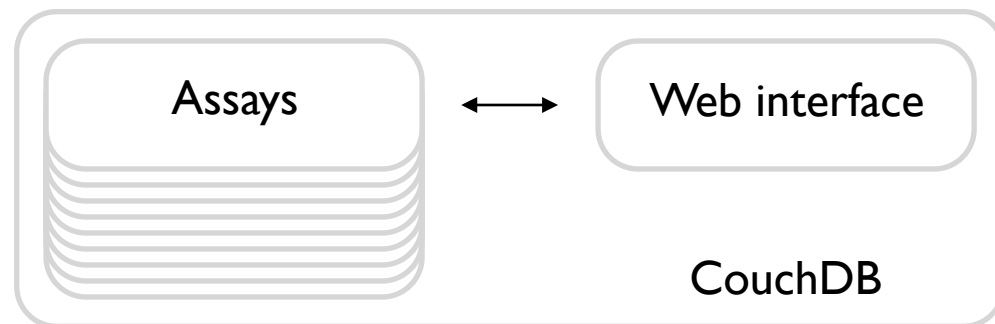
persephone /

name	age	message	history
 doc	3 months ago	Structure reorganization [nepahwin]	
 html	3 months ago	General code reorganization [nepahwin]	
 .gitignore	3 months ago	Add .gitignore file [nepahwin]	
 LICENSE	3 months ago	Add a brief license & terms of use [nepahwin]	
 README.md	3 months ago	Change format of the README file [nepahwin]	

github.com/nepahwin/persephone

Implementation

Database



The two things we give to the community :

- Data specification
 - a document
- Web interface
 - a piece of code that knows the contents of the document

Usage

Database can exist in many instances :

- Central institution
 - big collection of assays for public query
 - mirrored worldwide
 - mirrored to laptops, cell phones
- Collaboration
 - restricted collection of assays for private query
- Counting institution
 - restricted collection of assays for private query

The way forward

- Write the data specification
- Finish the v1.0 coding
- Port some existing datasets
- Release!

This is not a huge amount of work