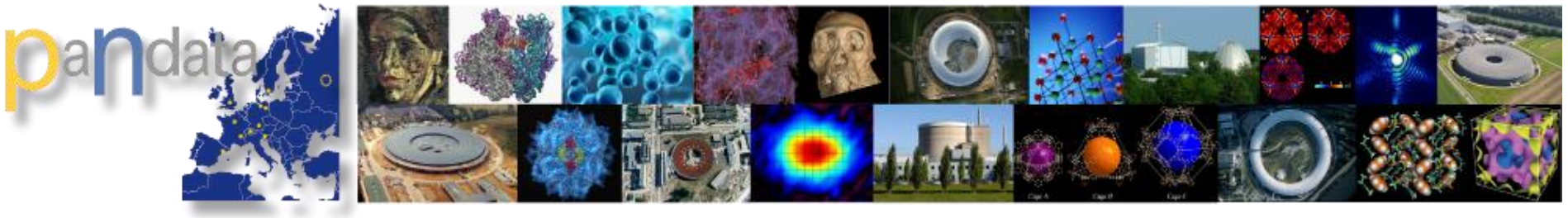
PANData Software Catalogue

Jamie Hall

Developer

IT Services, Institut Laue-Langevin

30th January 2013. CERN

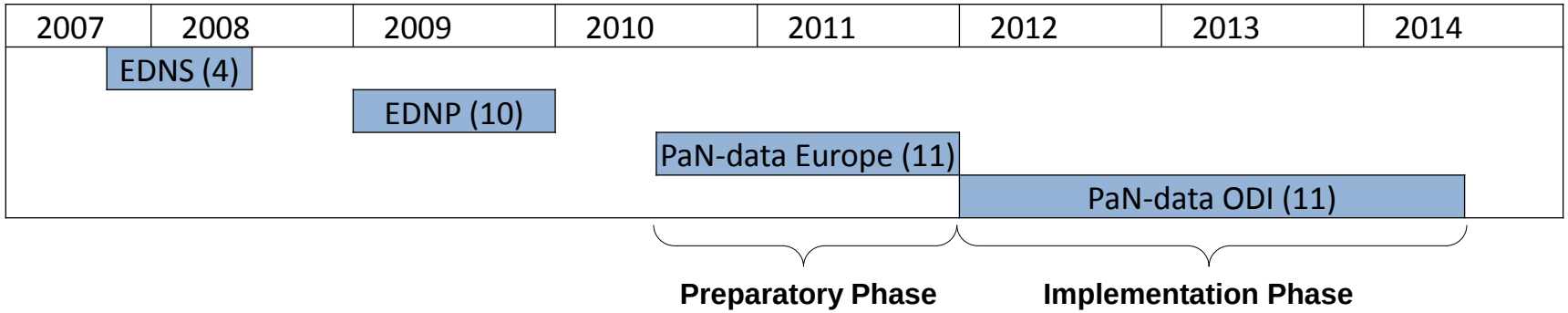


PHOTON AND NEUTRON DATA INFRASTRUCTURE

History

Established 2007 with 4 facilities

Expanded since to 13 facilities

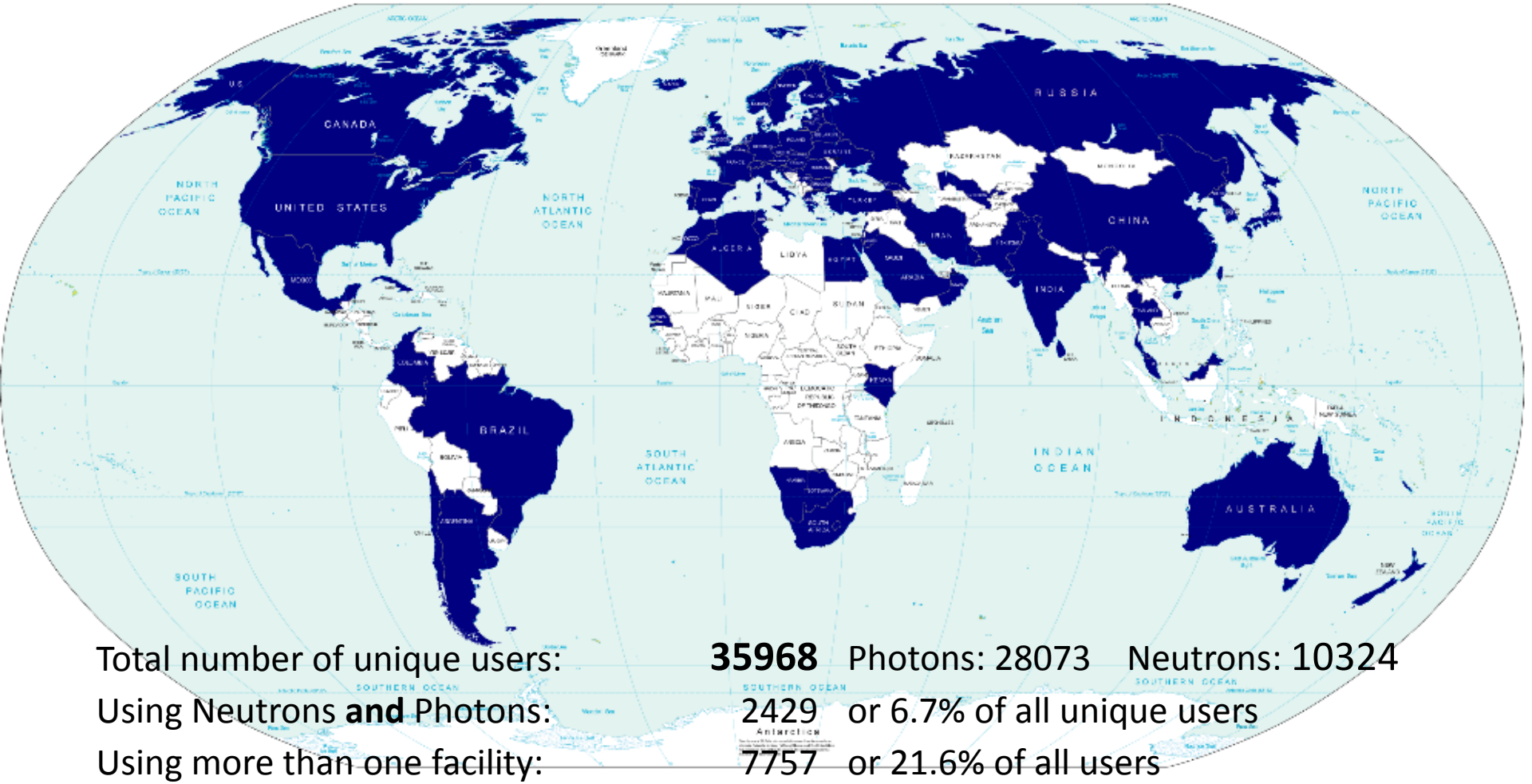


PaNdata ODI Partners

PaNdata is a joint effort of currently 13 major world class European x-ray and neutron research facilities providing experimental opportunities to a broad range of scientific communities world wide.

PaNdata aims to create a fully integrated pan-European, information infrastructure supporting the scientific process.

User Community



Total number of unique users: **35968** Photons: 28073 Neutrons: 10324

Using Neutrons **and** Photons: 2429 or 6.7% of all unique users

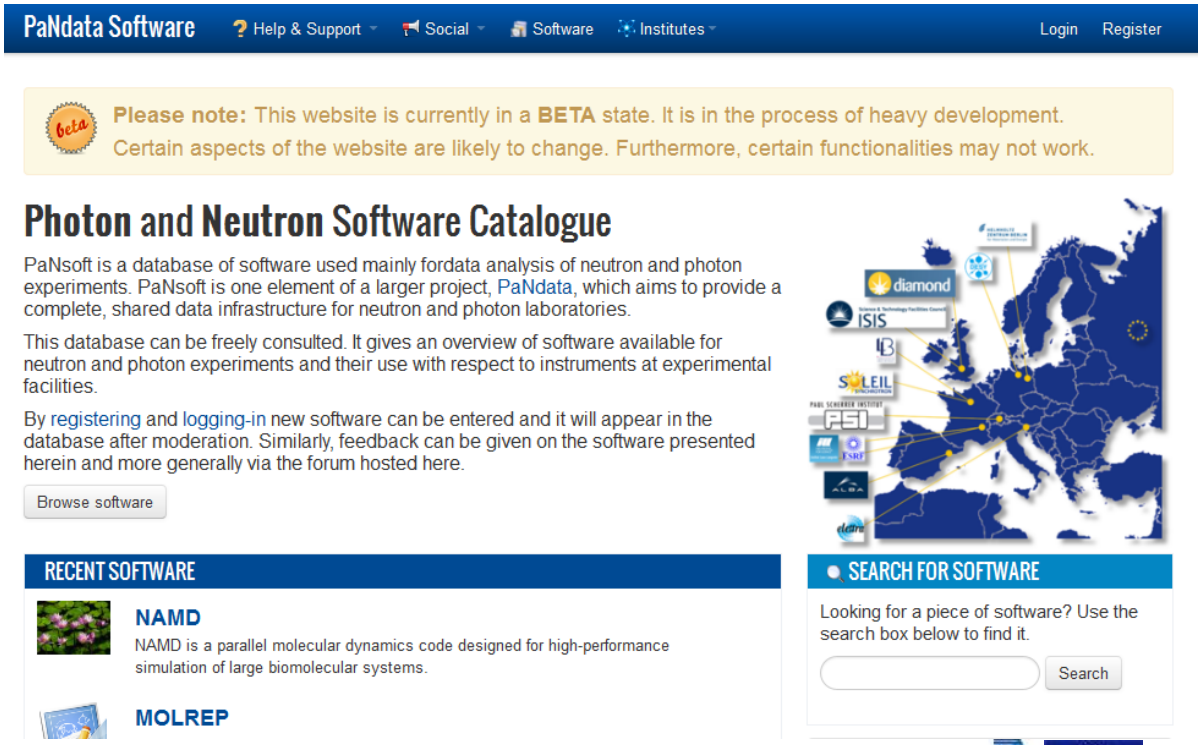
Using more than one facility: 7757 or 21.6% of all users

Objectives

- Harmonise authentication and authorisation
- Standardise data formats and annotation of data
- Allow transparent and secure remote access to data
- Establish sustainable and compatible distributed data catalogues (cross search engine)
- Provide compatible data analysis software
- Promote data policies in laboratories
- The Infra should cover the whole continuum (from proposal to publication)
- Allow long term preservation of data

THE SOFTWARE CATALOGUE

PANData Software is a web application which provides users with the necessary tools to find, collaborate and get involved with Photon and Neutron software



Why do we need a software catalogue?

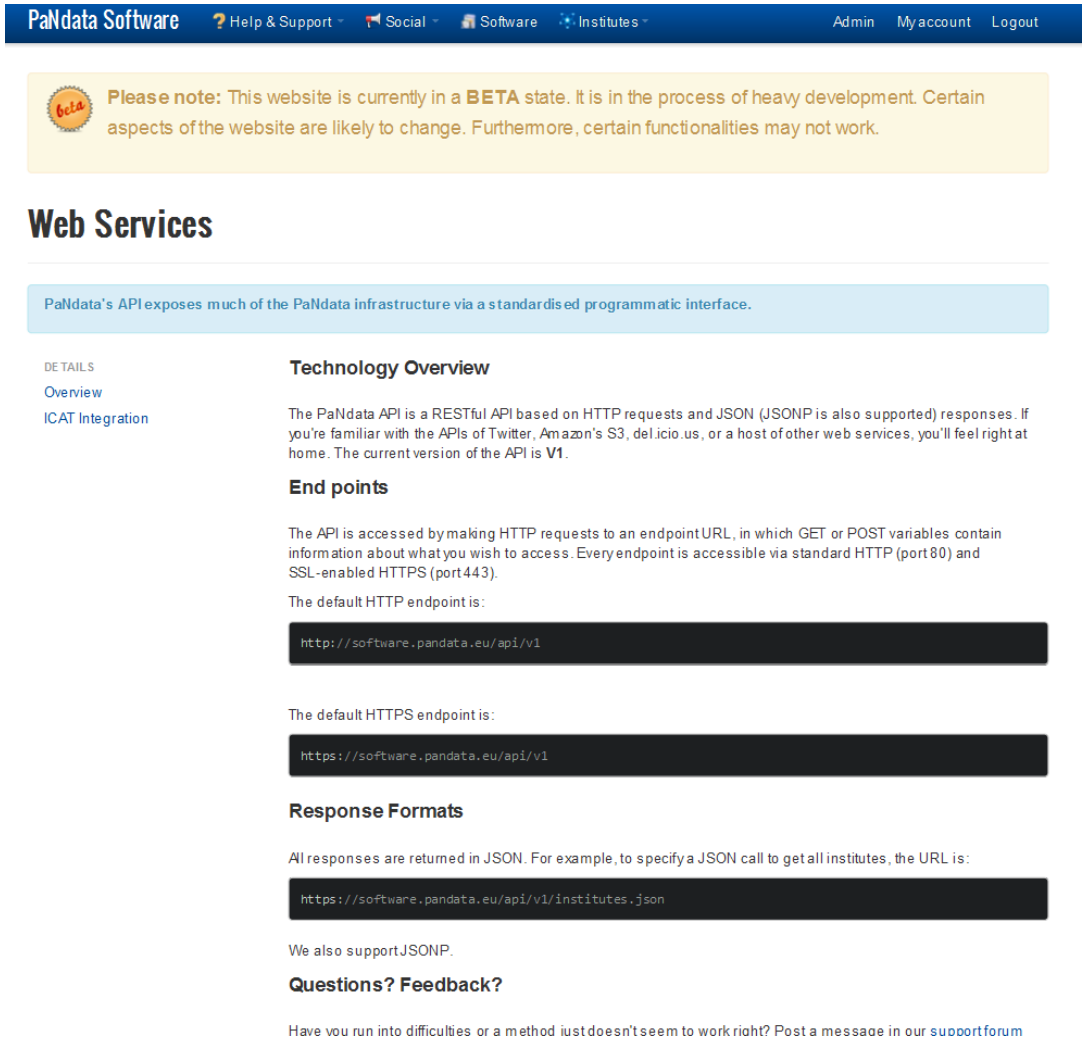
- Provide scientists with a centralised catalogue of software they can use to solve their problem
- Help users decide on the best software to use to fulfill their needs
- Help developers through the use of versioning, best practices and hierarchical data architecture to publicise their software
- Unique opportunity – de-facto place to go for Photon and Neutron Software
- Aggregate and share knowledge

Unknowns

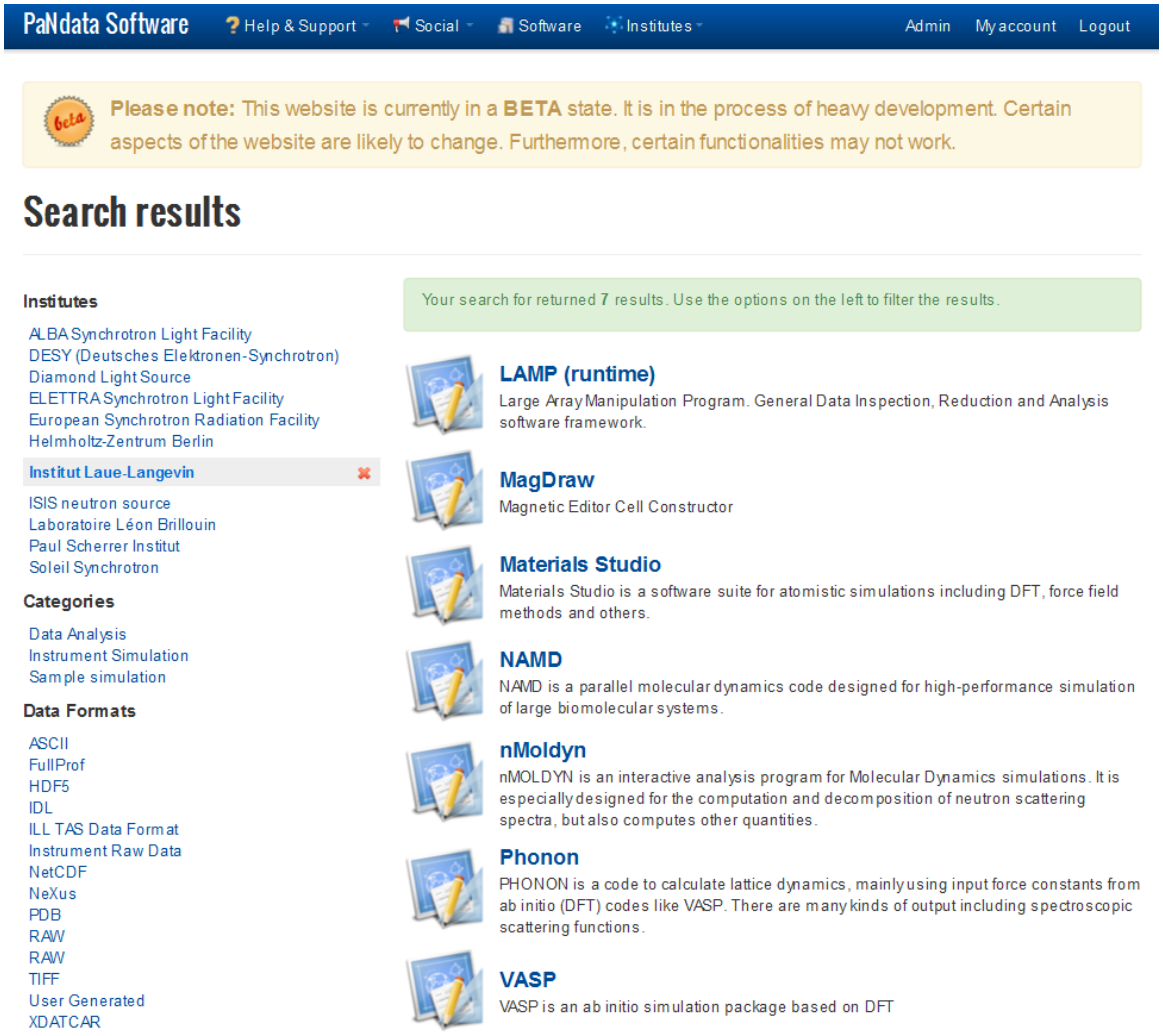
Lots of software and we don't know their state:

- Version
- Authors and contributors
- License
- Data formats
- Hardware requirements
- Software requirements
- Attributed to an institute
- Website
- Languages
- And much more…

What do we provide ?



What do we provide ?



What do we provide ?

PaNdata Software | Help & Support | Social | Software | Institutes | Admin | My account | Logout

Please note: This website is currently in a **BETA** state. It is in the process of heavy development. Certain aspects of the website are likely to change. Furthermore, certain functionalities may not work.

NAMD

NAMD is a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems.

Details | Screenshots

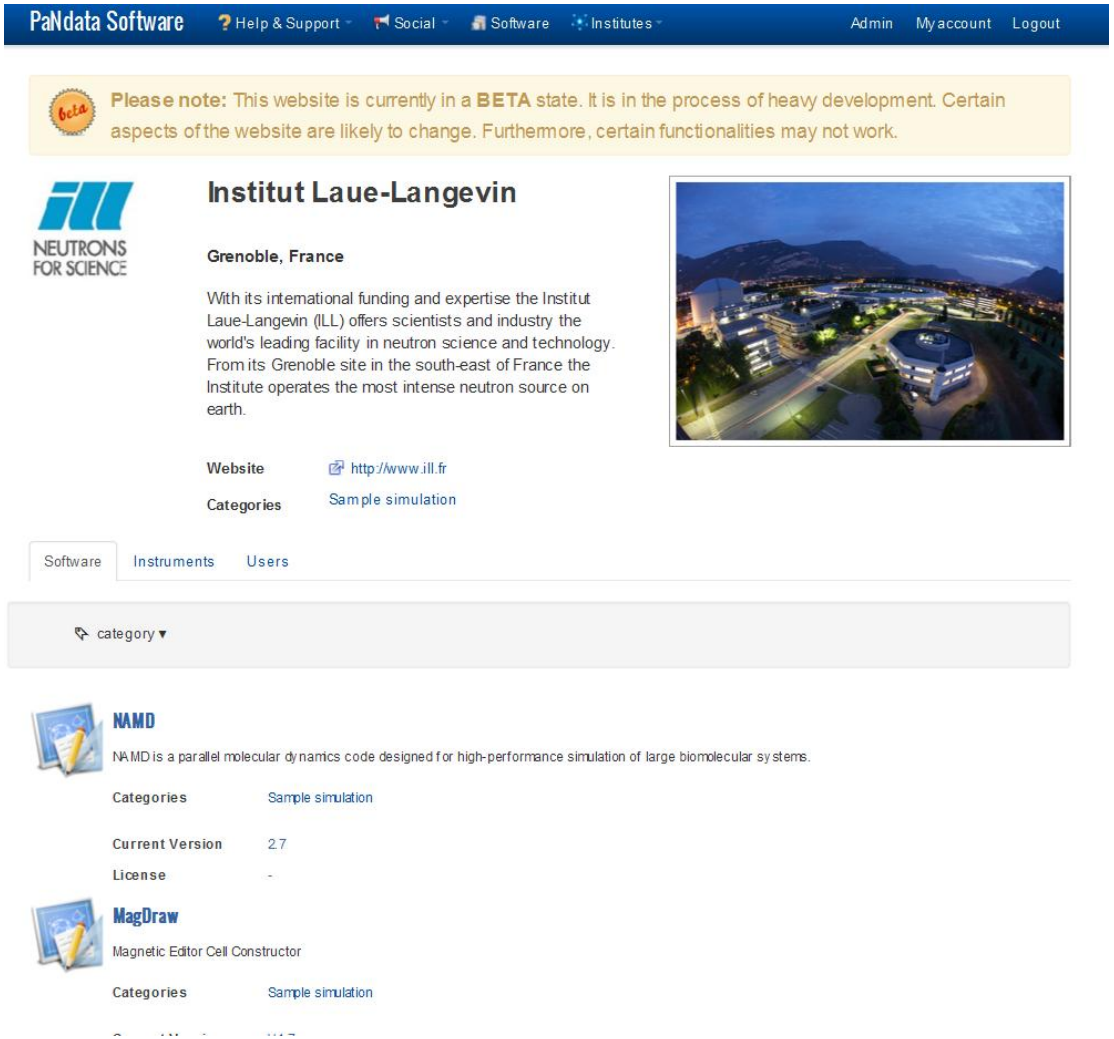
Overview

Website	http://www.ks.uiuc.edu/Research/namd//
Current Version	2.7
License	Unknown
Categories	Sample simulation
Beam types	
Institute	Institut Laue-Langevin

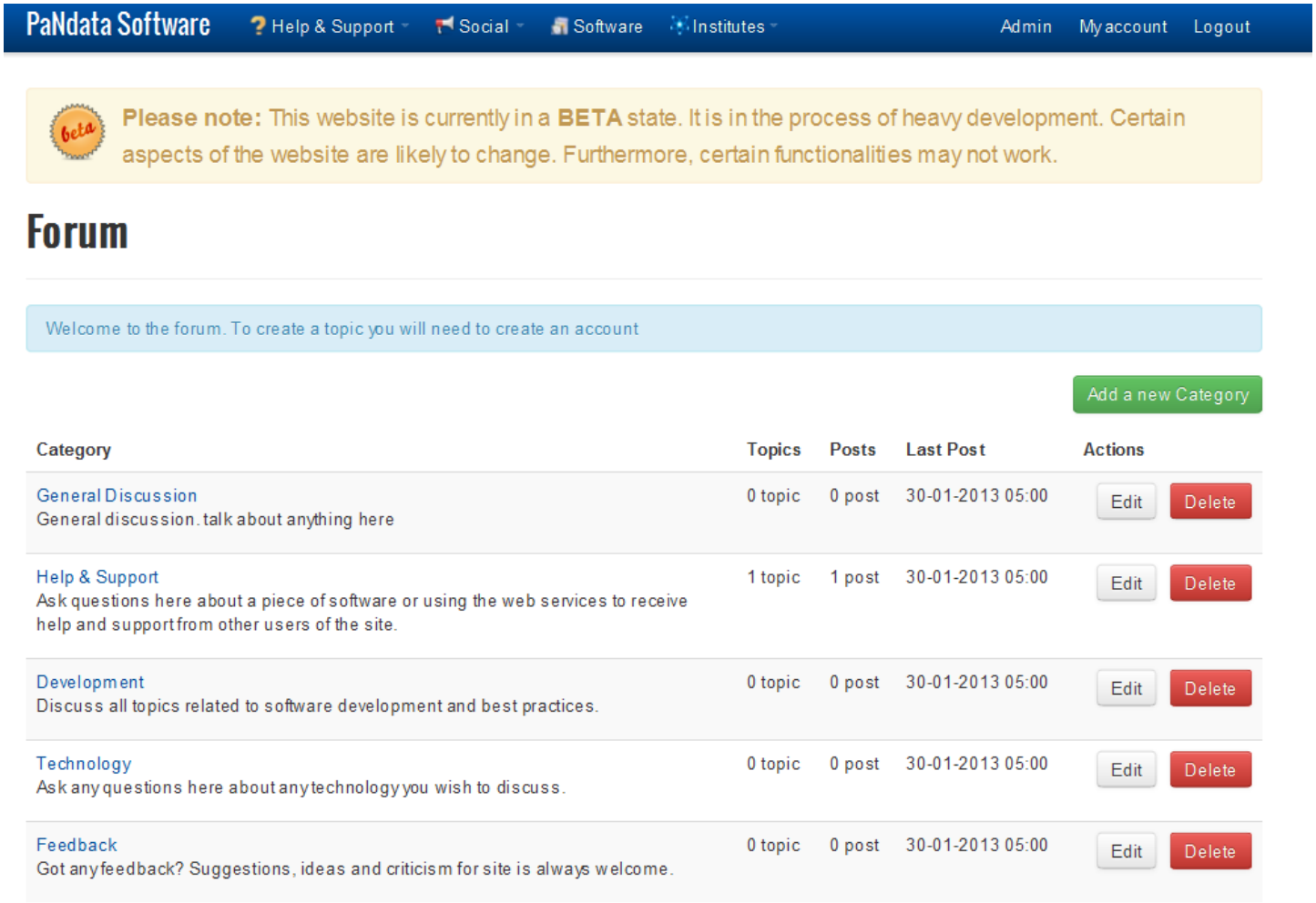
Users

hall Maintainer Contributor

What do we provide ?



What do we provide ?



Web services

- Defined as "a software system designed to support interoperable machine-to-machine interaction over a network”
- Communication over the HTTP or the HTTPS protocol
- Return hierarchical data that can be easily read by any programming language
- Expose the data infrastructure via a standardised programmatic interface through a single end point
- Primarily used by developers
- All web services are public
- A user can access data stored in the catalogue through any programming language of their choice
- Return responses in JSON

Search Engine

- PaNdata-ODI Software provides a flexible and scalable search engine which produces realtime and instantaneous results
- Flexible filtering
- Statistics to provide a real insight into:
  - What users are searching for?
  - Which terms are most popular?
  - Which filters are the most interesting?
  - And much more.

User authentication

- Two options of user authentication
  - Basic authentication
    - Local account with username and password
  - Federated authentication
    - SAML Single Sign On
    - Test implementation with Umbrella
    - Open to other protocols

What we store about software

- Name
- Description
- License
- Contact details
- Categories
- Instruments
- Beam type(s)
- Institute
- Website
- Users
- Screenshots
- Logo
- Versions
  - Version number
  - Software requirements
  - Hardware requirements
  - Platforms
  - Languages
  - Input formats
  - Output formats
- Comments
- Ratings

Social Web meets Software Catalogue

- Enable users to effortlessly discuss, collaborate and get involved via:
  - Voting and commenting on software
  - Participating in the forums
  - Personal messaging to other users
  - Following PaNdata-ODI Software on twitter
  - Newsletters
  - Individual profiles
  - Taking control of their own software and getting involved in others software

Future roadmap

- Integration with ICAT
- Publicity
- Software downloads (mirror)
- Continuous integration
- Apps developed with the PaNdata-ODI software web services (Android, iOS)
- OAuth web services integration
- Statistics
- Continue pushing the NeXus format
- Intelligent software relations
- Long term sustainability
- Linking publications to software
- Persistent Identifiers for software

Persistent Identifiers

- How to identify and refer to software persistently ?
  - Sustainability of software's repository ?
  - Sustainability of infrastructure?
  - Ownership ?
  - Contract & licensing ?
  - What to identify? Software, releases, releases per OS, libraries
  - DOI ? Other mechanisms (PURL(Z), ARK, URN, …)

- Absolute necessity for data preservation and provenance activities.

Questions ?

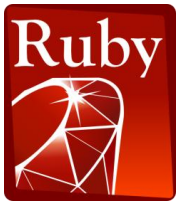
mongoDB

debian

REDMINE
flexible project management

doctrine

Symfony


Ruby

git

php

bit.ly

TWIG

VARNISH
CACHE

GRAVATAR

Bootstrap, from Twitter
Bootstrap is a toolkit from Twitter designed to kickstart development of webapps and sites.
It includes base CSS and HTML for typography, forms, buttons, tables, grids, navigation, and more.
Nerd alert: Bootstrap is built with Less and was designed to work out of the gate with modern browsers in mind.

jQuery

twitter

MySQL®

capifony

ELASTIC SEARCH