DSPACE

Introduction, Features & Technology

Mukesh Pund

Principal Scientist

CSIR-NISCAIR

Outline

- **What is DSpace and what does it do?**
- **The DSpace information model**
- **Components & features of DSpace**

What is DSpace?

- **A digital repository system**
- **Captures, stores, indexes, preserves and redistributes an organization's research material in digital formats**
- **Research institutions worldwide use DSpace for a variety of digital archiving needs –**
  - from institutional repositories (IRs) to
  - learning object repositories or electronic records management, and more.

DSpace is…

- **DSpace is a Digital Repository System**
  - Institutional Repositories
  - Learning Object Repositories

- **Open source development model**
  - BSD License

More than 1700 known instances of DSpace around the world

(source: http://registry.duraspace.org/registry/dspace, as on January 2015)

Who built DSpace?

- **The MIT Libraries and Hewlett-Packard (HP) jointly developed DSpace. The system is now freely available to research institutions world-wide as an open source system that can be customized and extended.**

- **Formed not-profit organization in July 2007: Dspace Foundation**

- **In July 2009, the DSpace Foundation ceased operation. DuraSpace took over supporting the DSpace project.**

- **Who manages DSpace?**
  - DSpace is freely available as open source software. The DSpace Community manages the code base and releases new versions of the software. An active community of developers, researchers and users worldwide contribute their expertise to the DSpace Community

Who built DSpace?

- Who can download the software?

Open-source systems like DSpace are available for anyone to download and run at any type of institution, organization, or company (or even just an individual).

Users are also allowed to modify DSpace to meet an organization's specific needs.

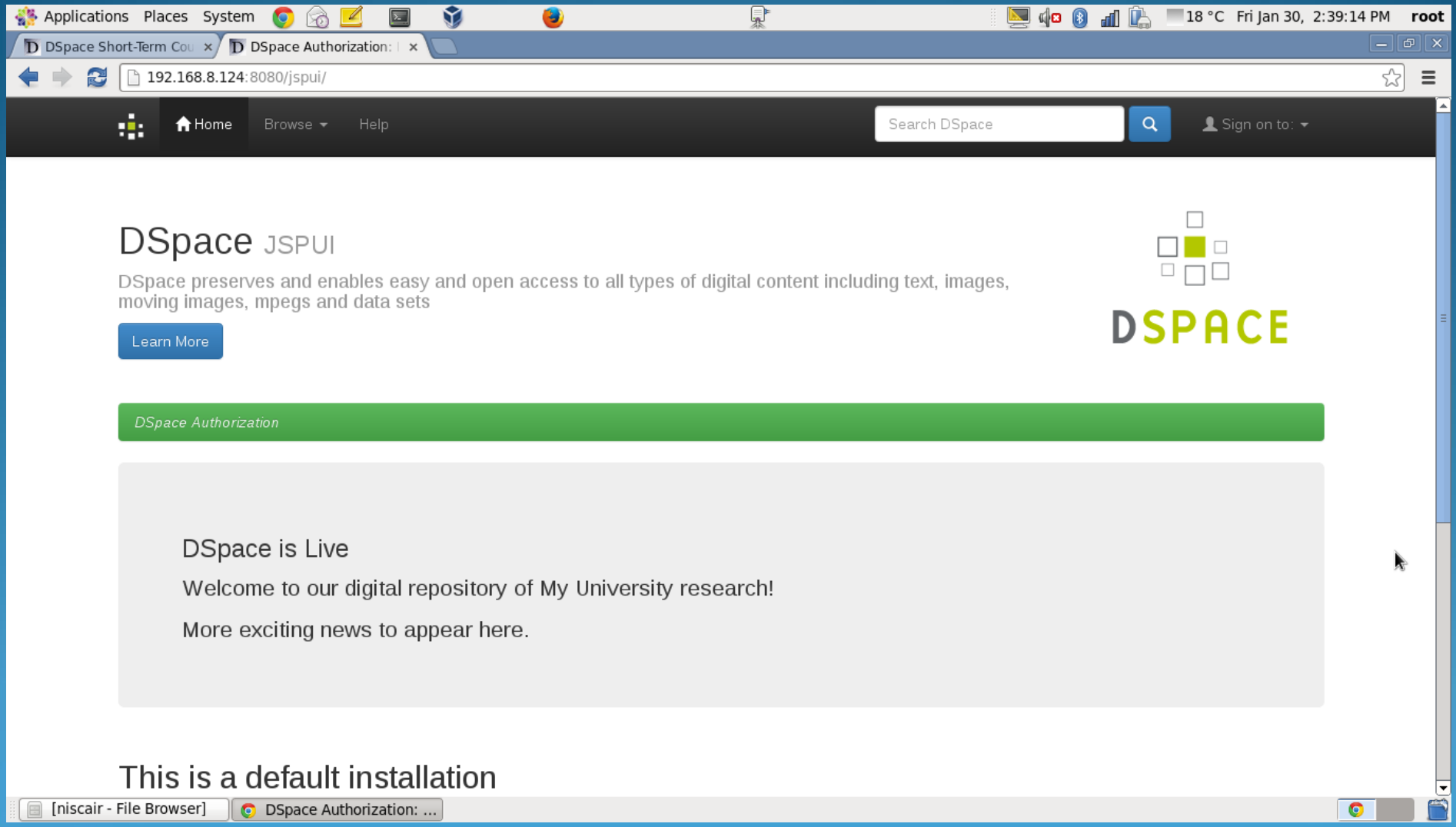
The BSD distribution license describes its specific terms of use.

DSpace is freely available as open-source software from SourceForge.

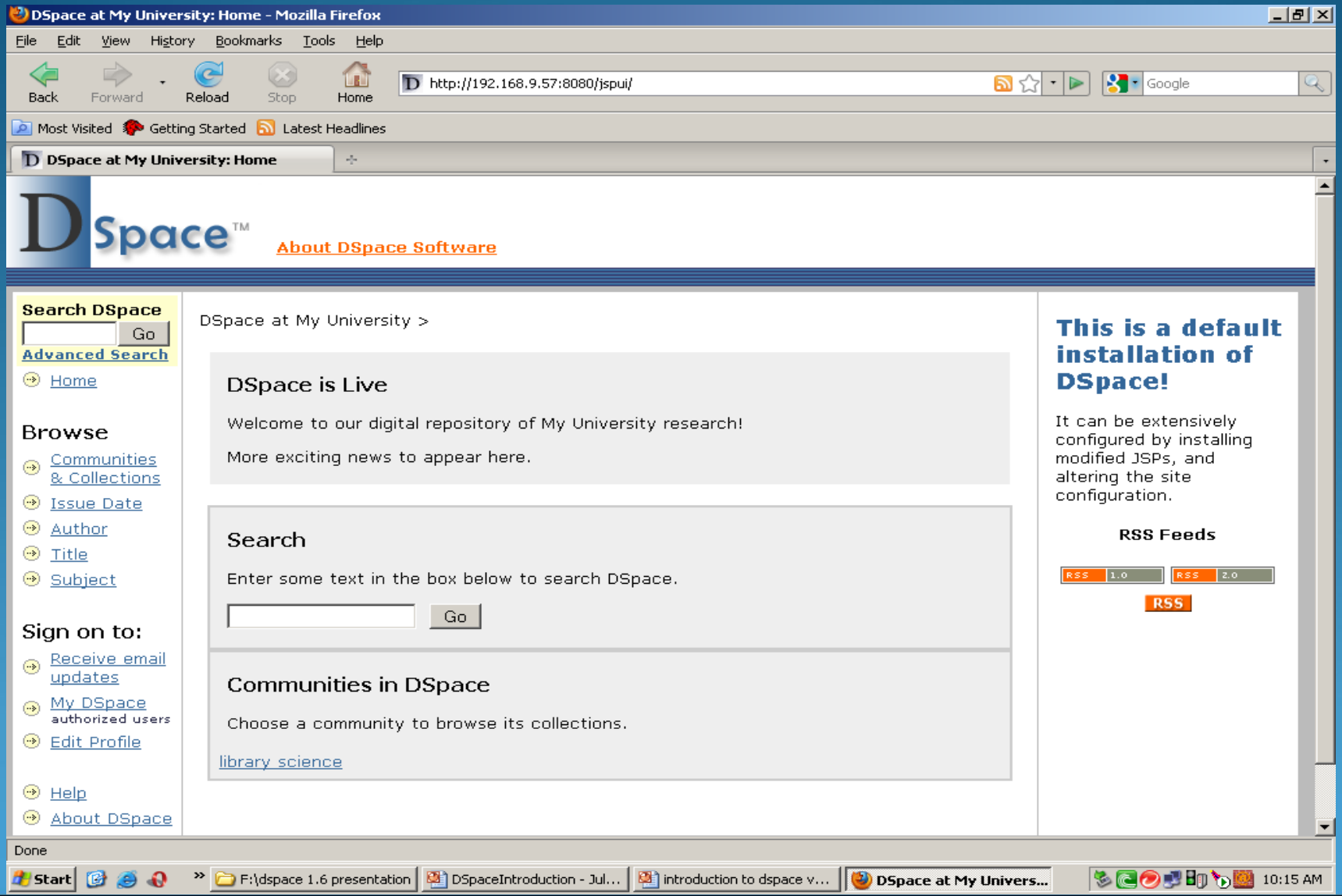
DSpace…

- **Captures**
  - Digital research material directly from the creators
- **Describes**
  - Allows descriptive, technical, and rights metadata
  - Assigns persistent identifiers
- **Distributes**
  - Searches metadata & full text
  - Delivers content over the web
- **Preserves**
  - Content in supported formats for long term preservation

DSpace: New CINECA Theme

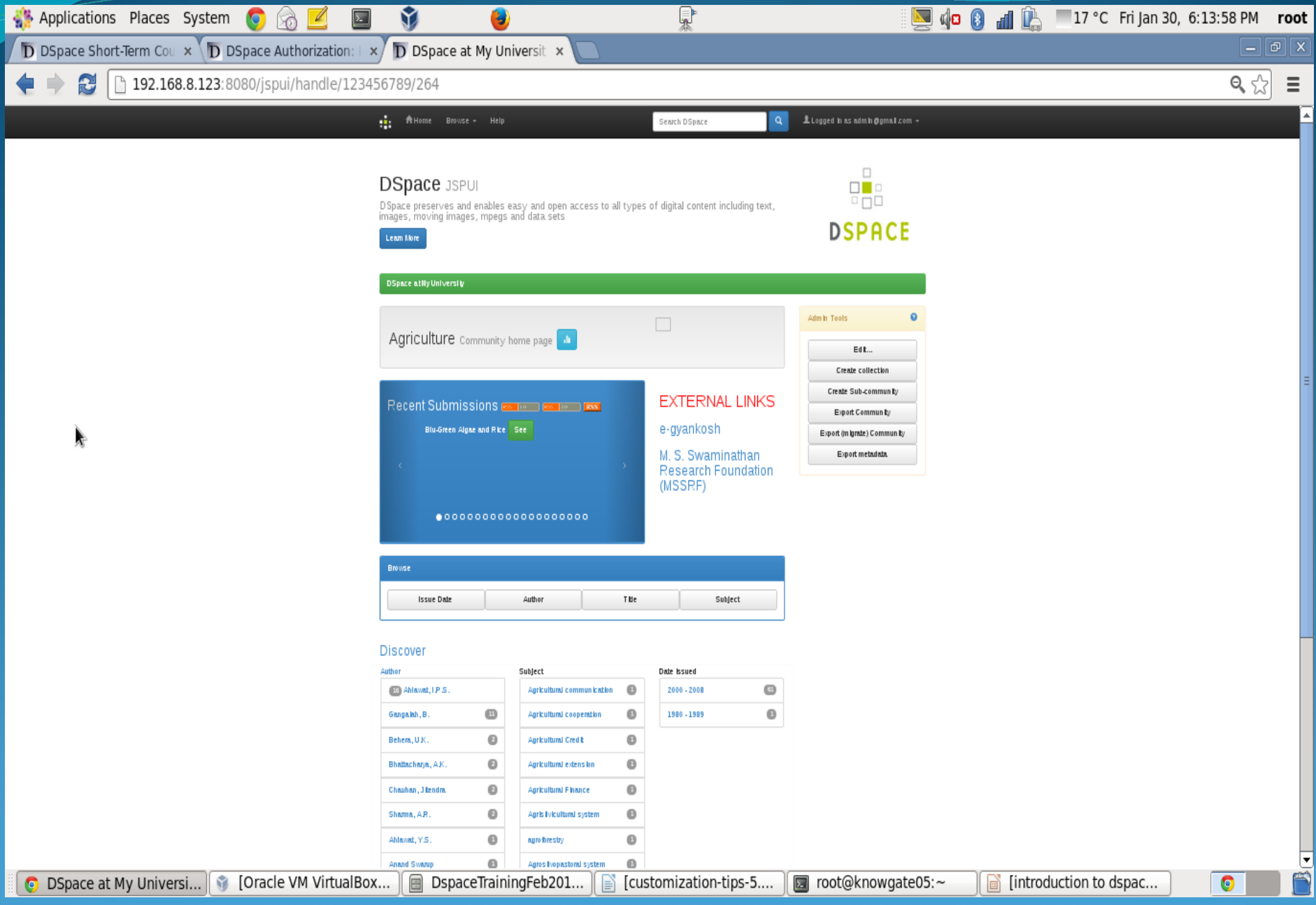


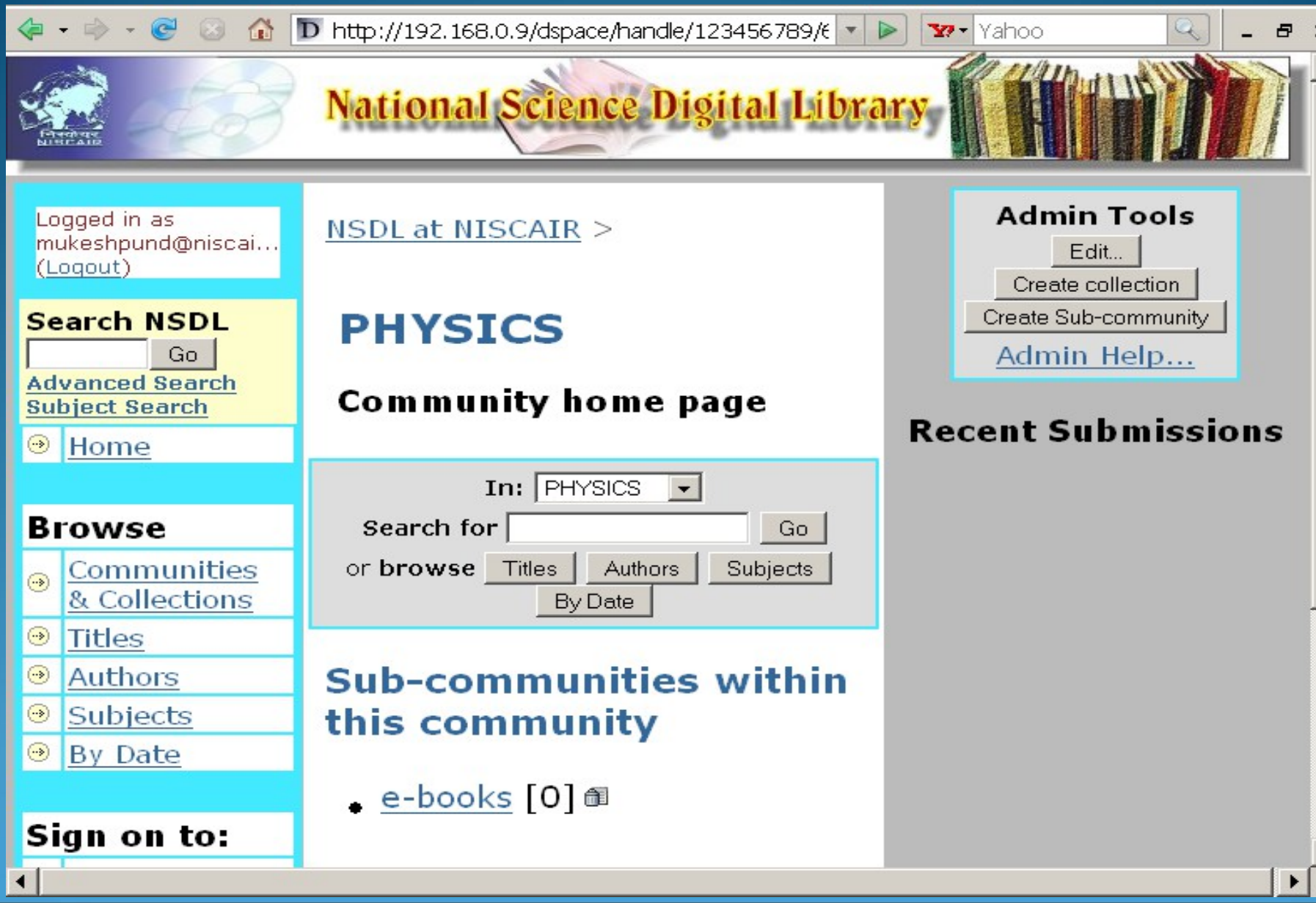
DSpace: Old Theme



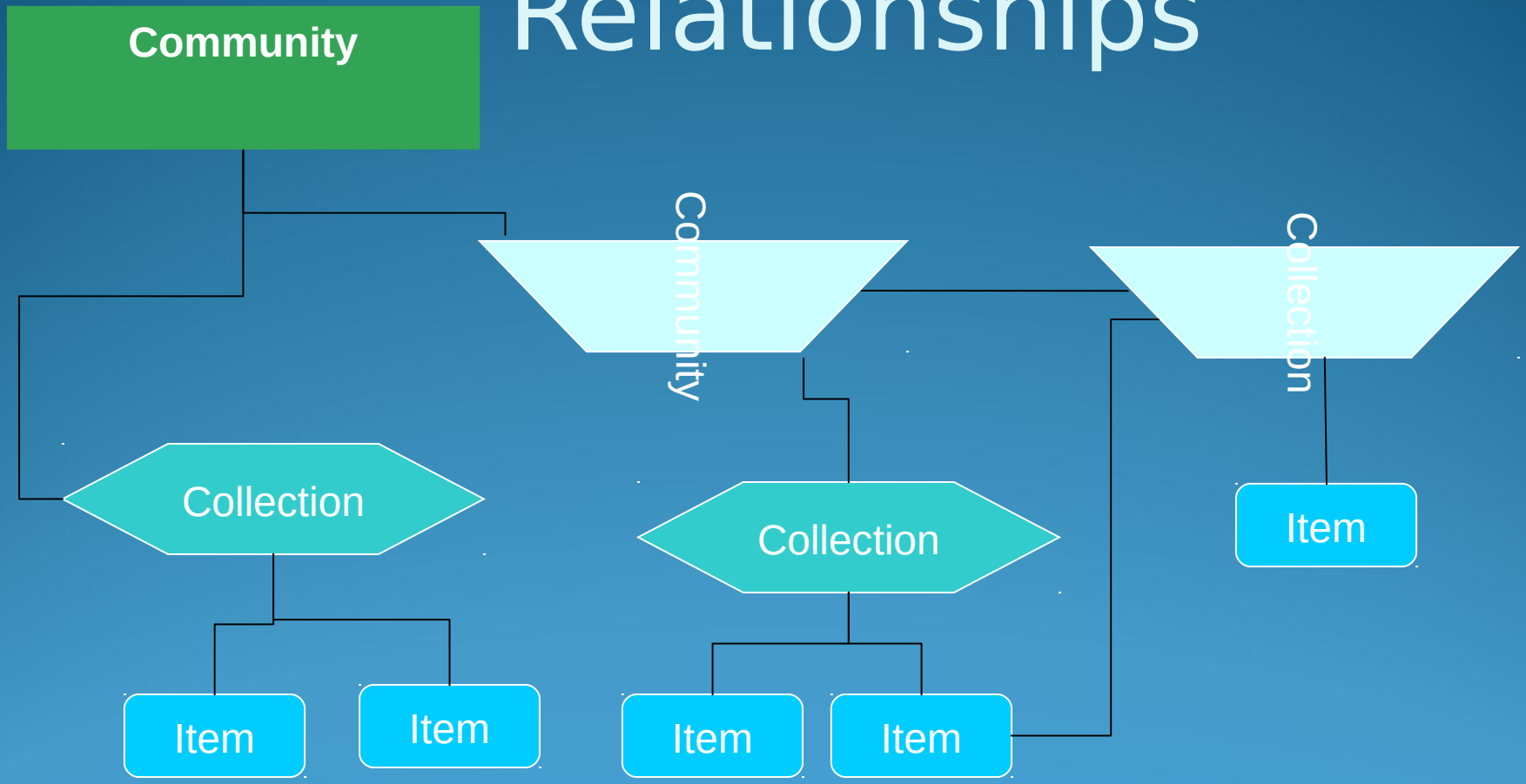
DSpace: Information Model

- **Communities & Collections**
  - Hierarchical organization of items in the repository
- **Items**
  - Logical units of content
  - Receive persistent identifiers
- **Bitstreams & Bundles**
  - Individual digital files


Applications Places System
17 °C Fri Jan 30, 6:13:58 PM root
DSpace Short-Term Cou
DSpace Authorization:
DSpace at My Universit
192.168.8.123:8080/jspui/handle/123456789/264
Home Browse Help
Search DSpace
Logged in as admin@gmail.com
DSpace JSPUI
DSpace preserves and enables easy and open access to all types of digital content including text, images, moving images, mpegs and data sets
Learn More
DSPACE
DSpace at My University
Agriculture Community home page
Admin Tools
Edit...
Create collection
Create Sub-community
Export Community
Export (migrate) Community
Export metadata
Recent Submissions
Blu-Green Algae and Rice
See
EXTERNAL LINKS
e-gyankosh
M. S. Swaminathan Research Foundation (MSSRF)
Browse
Issue Date
Author
Title
Subject
Discover
Author
Ahlawat, I.P.S.
Gangaiah, B.
Behera, U.K.
Bhattacharya, A.K.
Chauhan, Jitendra
Sharma, A.R.
Ahlawat, Y.S.
Anand Swarup
Subject
Agricultural communication
Agricultural cooperation
Agricultural Credit
Agricultural extension
Agricultural Finance
Agrisilvicultural system
agroforestry
Agrosilvopastoral system
Date Issued
2000 - 2008
1980 - 1989
DSpace at My Universi...
[Oracle VM VirtualBox...
DspaceTrainingFeb201...
[customization-tips-5....
root@knowgate05:~
[introduction to dspac...


http://192.168.0.9/dspace/handle/123456789/
Yahoo
National Science Digital Library
Logged in as mukeshpund@niscai... (Logout)
Search NSDL
Go
Advanced Search
Subject Search
Home
Browse
Communities & Collections
Titles
Authors
Subjects
By Date
Sign on to:
NSDL at NISCAIR >
PHYSICS
Community home page
In: PHYSICS
Search for
Go
or browse
Titles
Authors
Subjects
By Date
Sub-communities within this community
e-books [0]
Admin Tools
Edit...
Create collection
Create Sub-community
Admin Help...
Recent Submissions

Community & Collection Relationships



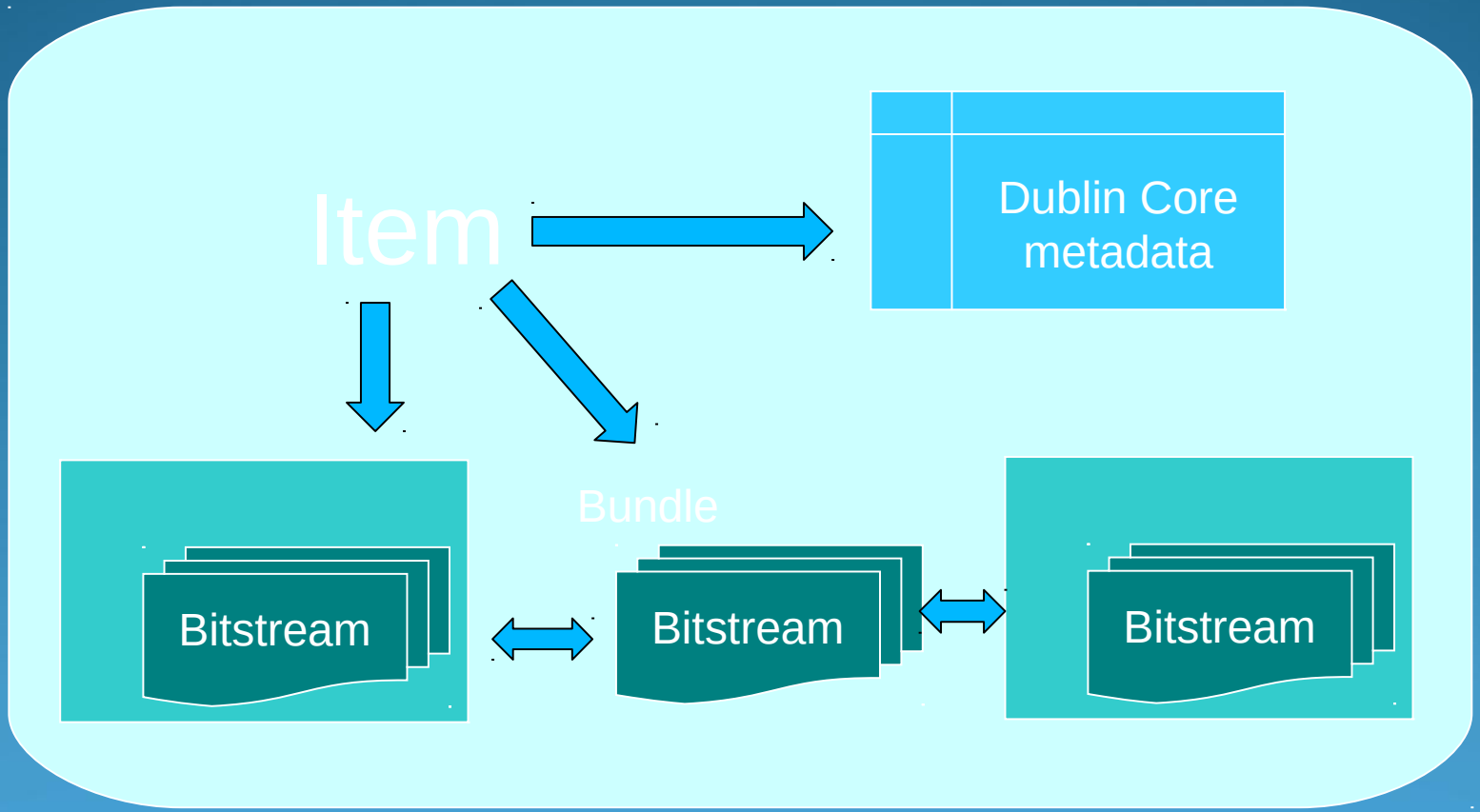
Communities & Collections

- **Collections and Communities organize items into a hierarchical form**
- **Metadata:**
  - Limited descriptive metadata available
    - Name, description, license, etc…
- **Example:**

Communities
College of Architecture
Office of Graduate Studies

Collections
A Department's Technical Reports
A faculty member's publications

Item Composition



Items

- **Items are logical units of content**
- **Metadata:**
  - All items have qualified Dublin Core metadata
  - May contain metadata in other formats encoded as a bitstream
- **Example:**
  - Book
  - ETD thesis
  - Web page (Images, CSS, HTML)
  - Photographs

Bitstreams

- **Bitstreams are Individual Digital files**

- **Metadata:**
  - Limited descriptive metadata available
    - name, file format, size, etc…
- **Example:**

- PDF file
- Word document
- JPEG picture

- Executable program
- HTML file
- CSS file

Bundles

- **Bundles group related bitstreams together**
- **Metadata:**
  - No metadata
- **Example:**
  - HTML files and images that compose a single HTML document may be organized into a bundle

Components & Features of DSpace

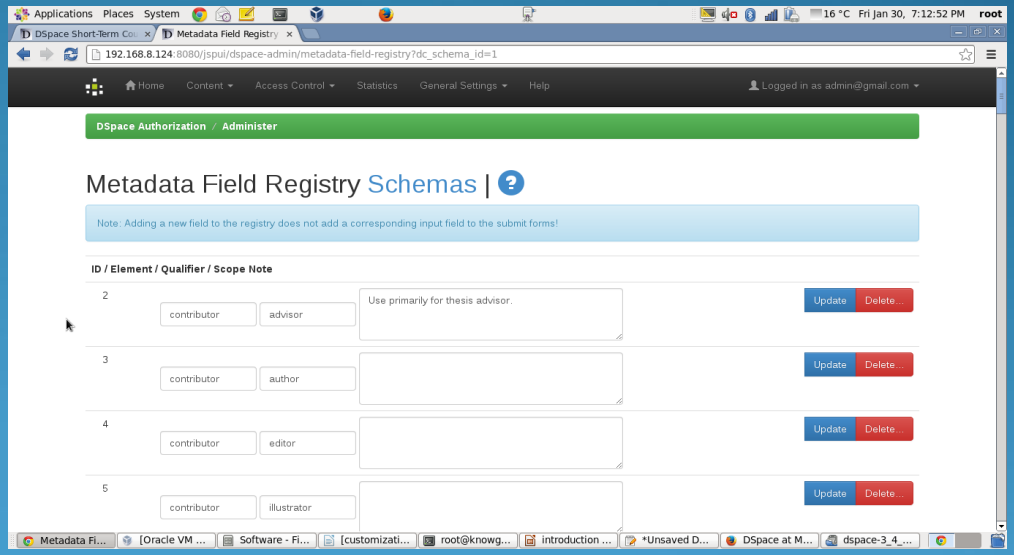
Item Metadata

- **Descriptive**
  - Qualified Dublin Core
  - Non Dublin Core supported also (as long as it's still flat)
  - Any other format may be added as a bitstream
    - However, it will not be searchable
- **Administrative**
  - Who can access, remove, or modify an item
  - Stored in the database, no standard format used
- **Structural**
  - Very basic
  - What bitstreams are contained in an item
  - What collections and communities does an item belong to

Dublin Core registry

Maintain what metadata fields may exist for an item in DSpace.

- **Three components**
  - Schema (new)
  - Element
  - Qualifier
  - Scope Note

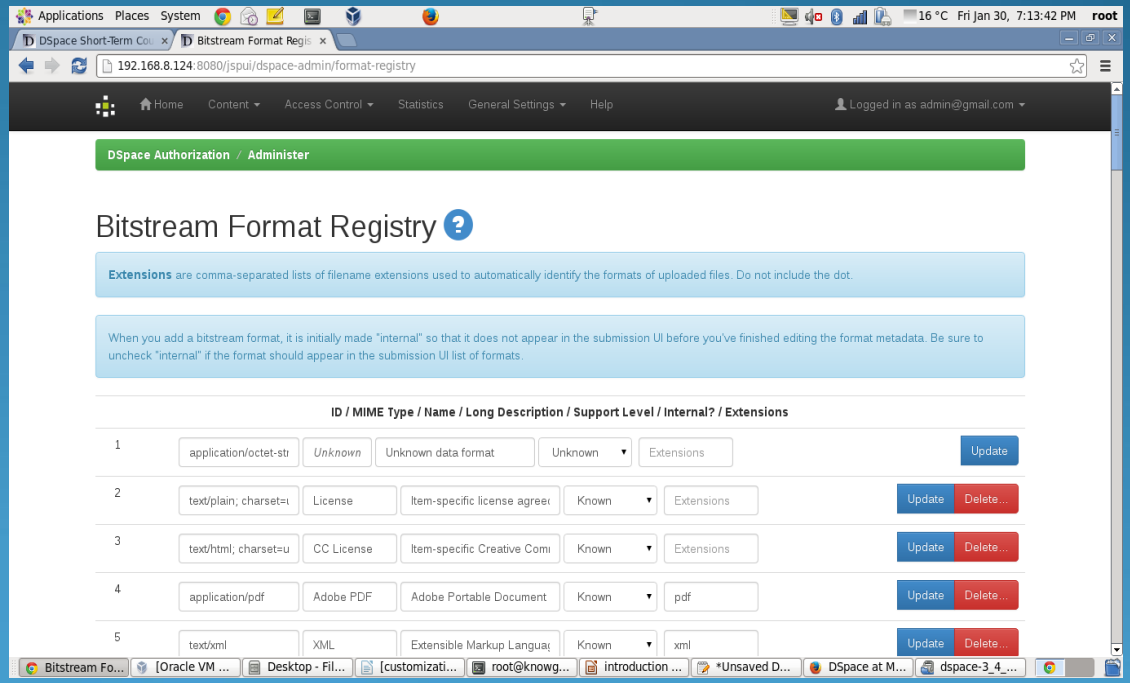


Format Registry

- Maintain a registry of file formats

Three levels:

- Supported
- Known
- Unknown



Handle System

- **CNRI- Corporation for National Research Initiatives**
- **Provides a persistent identifier**

- **Standard URL's change**
  - Hardware or software changes
  - Political changes
  - Network changes

- **Handles attempt to address these problems by creating a permanent URL independent of the repository.**

- **Example:**
  - http://hdl.handle.net/1969.1/3356

E-People

- **DSpace user accounts are called E-people**
- **If permitted, an e-person may:**
  - Login to the site
  - Sign up to receive notifications about changes to a collection
  - Submit new items to collections
  - Administer collections/communities
  - Administer the DSpace site.

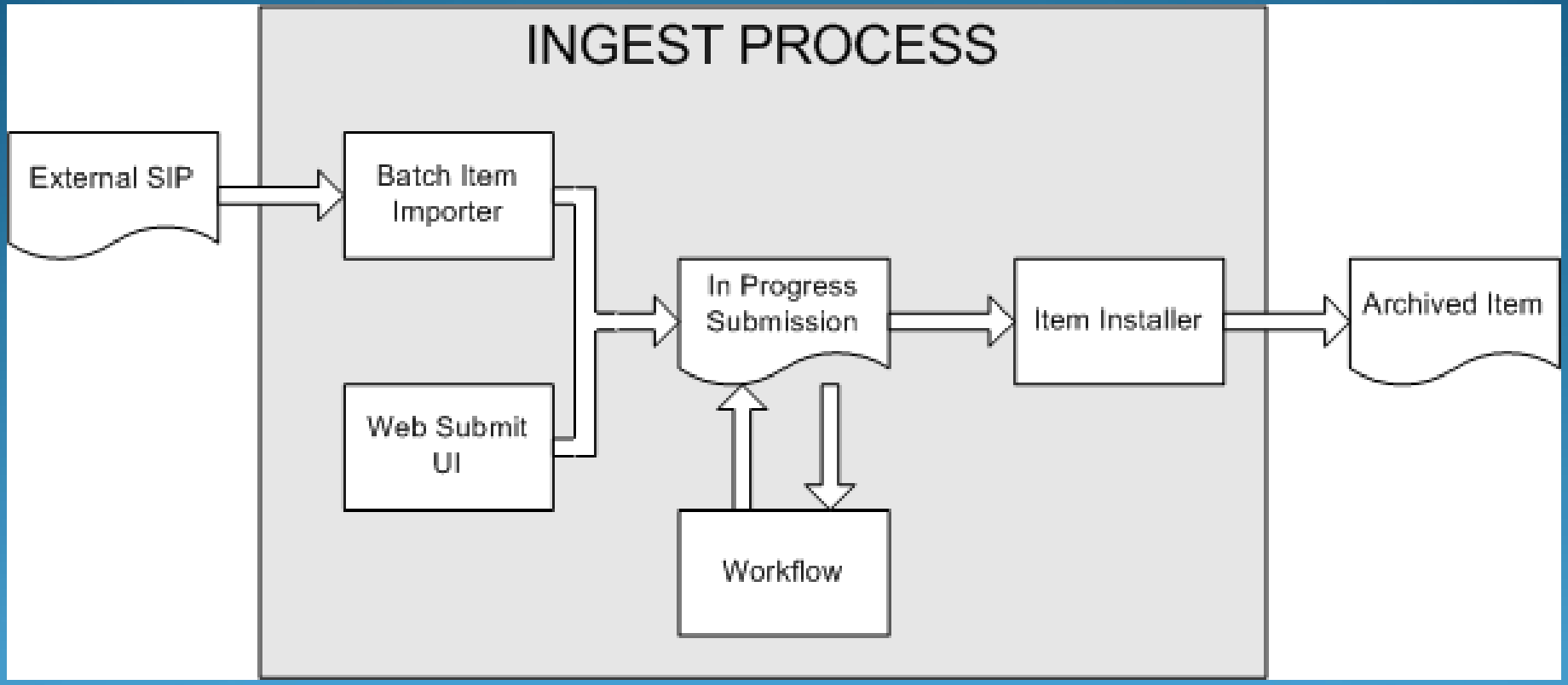
Authorization

- **The DSpace authorization system enables administrators to give e-people the ability to perform the following operations on an object.**
  - Add / Remove
    - Enable an e-person to add or remove any object (community, collection, item)
  - Collection Administrator
    - Enable an e-person to edit an item’s metadata, withdraw items, or map items into the collection.
  - Write
    - Enable an e-person to add or remove bitstreams
  - Read
    - Enable an e-person to read bitstreams

Ingestion

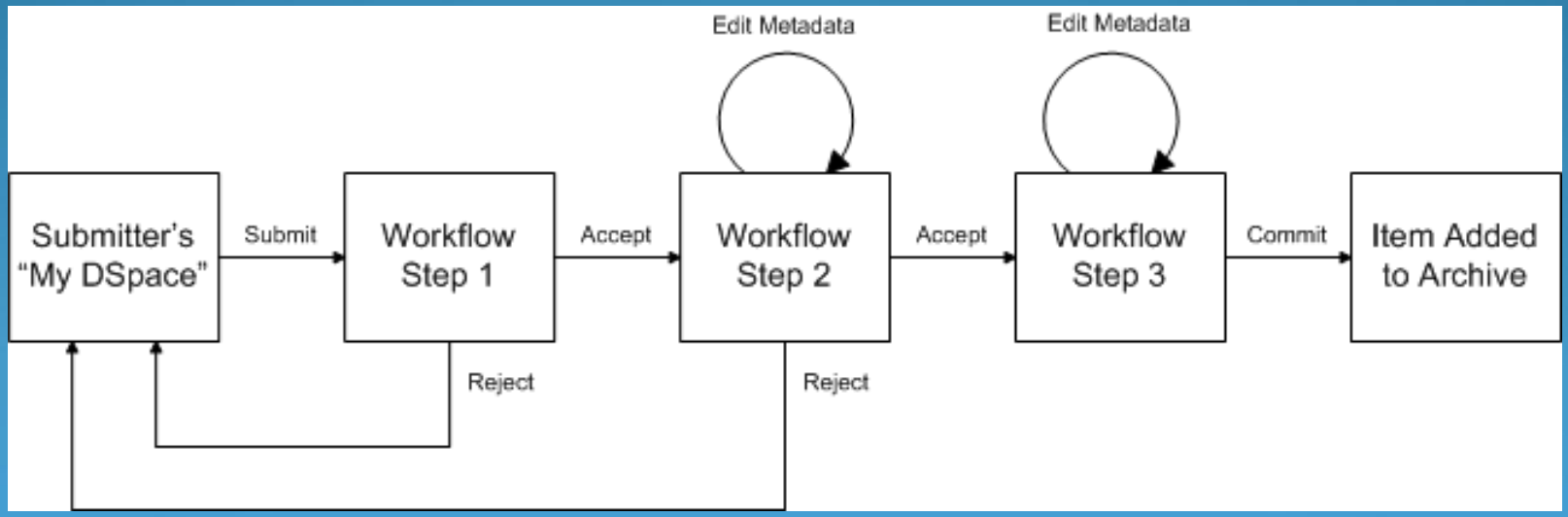
- **Ingestion = getting stuff into DSpace**
- **Batch import**
  - Many at a time
  - Needs to be in a specific format
    - XML encoded metadata
    - Bitstreams
- **Web based submission**
  - One at a time
  - Workflow processes

Ingestion Processes



Workflow

- **Step 1: May reject the submission**
- **Step 2: Edit metadata or reject**
- **Step 3: Edit Metadata**



Search & Browse

- **Users may browse any item in DSpace**
  - Title
  - Author
  - Date
  - Community / Collection
  - Subject (new)

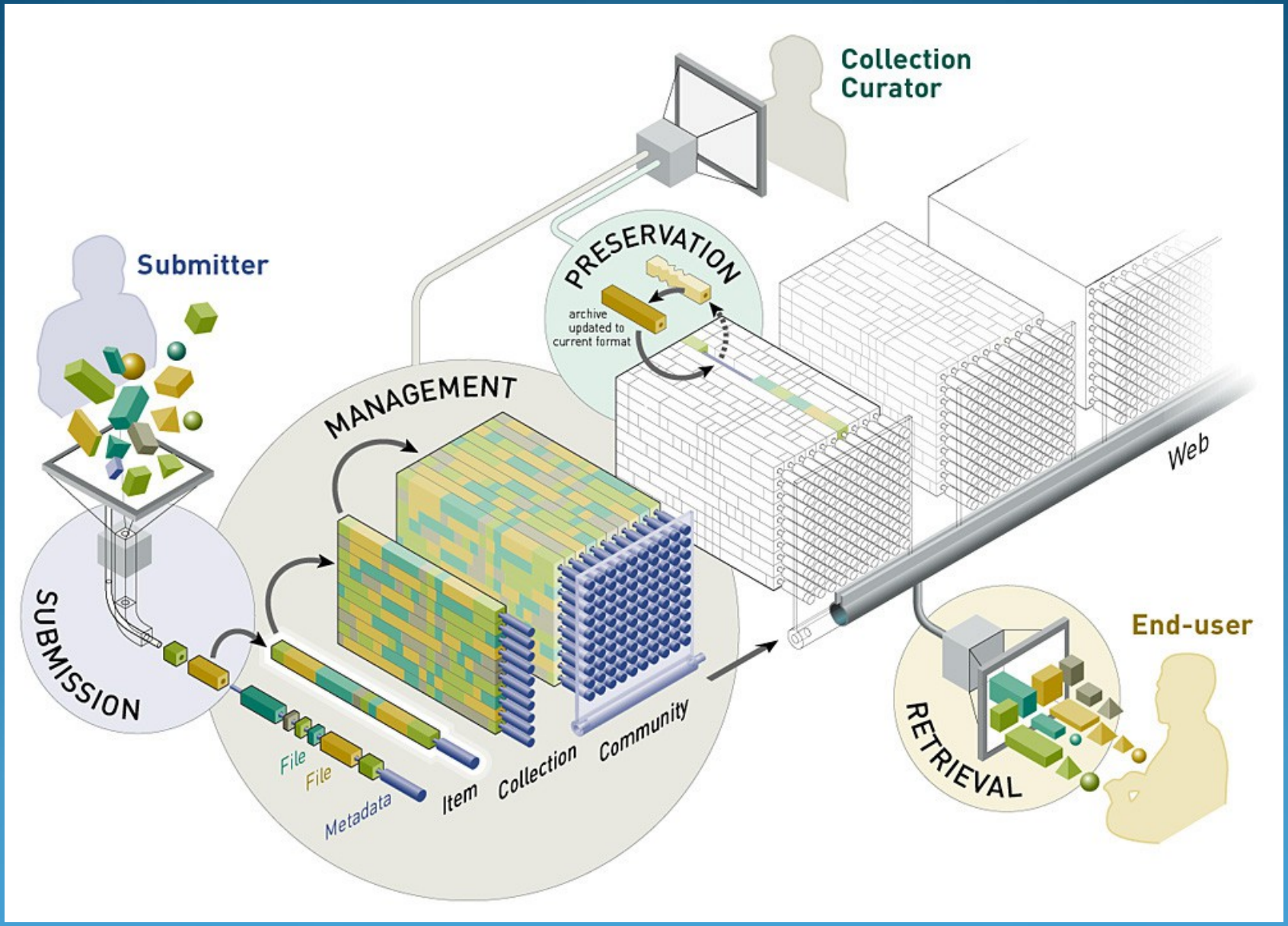
- **Users may search for any item in DSpace based upon any Dublin Core value or a full text search.**

OAI-PMH

- **Enables other sites to harvest metadata from a DSpace repository**
- **Collections are exposed as OAI sets**
- **Only Dublin Core metadata is available**

Statistics

- **Analysis the DSpace logs to generate a set of statistics on how DSpace is being used.**
- **Metrics collected:**
  - Number of item visits
  - Number of collection visits
  - Number of community visits
  - Number of OAI requests
  - Number of logins
  - Most popular searches
- **Presented in a by-month form or in-total form.**
- **SOLR Statistics**
- **Google Analytic**


Collection
Curator
Submitter
PRESERVATION
archive
updated to
current format
MANAGEMENT
Web
SUBMISSION
End-user
RETRIEVAL
File
File
Metadata
Item
Collection
Community

DSpace: Software Components

- **UNIX-like OS or Microsoft Windows**
- **Sun Java JDK 7 or later (standard SDK is fine, you don't need J2EE)**
- **Apache Maven -3.0.4 or later (Java build tool)**
- **Apache Ant 1.8.4 or later (Java build tool)**
- **Relational Database: (PostgreSQL or Oracle).**
- **Servlet Engine: (Jakarta Tomcat 7.x, Jetty, Caucho Resin or equivalent).**
- **Perl (required for [dspace]/bin/dspace-info.pl)**
- **Dspace Software Source code**

New Features in Dspace 4.x and 5.x

- **New Theme of JSPUI (by CINECA)**
- **Curation tasks administrative UI**
- **Advanced Embargo feature**
- **Item level versioning feature**
- **AJAX progress bar for file upload the submission upload step**
- **Sherpa/Romeo integration in the submission upload step**

New Features in Dspace 4.x and 5.x

- **Discovery: Search & Browse is now enabled by default in both XMLUI and JSPUI**
- **Facet Search (in Advance Search)**
- **More:**
  - Dspace-4: https://wiki.duraspace.org/display/DSDOC4x/Release+Notes
  - Dspace-5: https://wiki.duraspace.org/display/DSDOC5x/Release+Notes

Thank You