

# Customized Galaxy with applications as Web Services and on the Grid for Open Source Drug Discovery (OSDD)



A CSIR led team India consortium with global partnership for affordable healthcare



**Anshu Bhardwaj**  
Council of Scientific & Industrial Research (CSIR),  
India

**Chintalapati Janaki,**  
Center for Development of Advanced Computing (C-DAC),  
India

# TB Kills

One person  
every 20 seconds

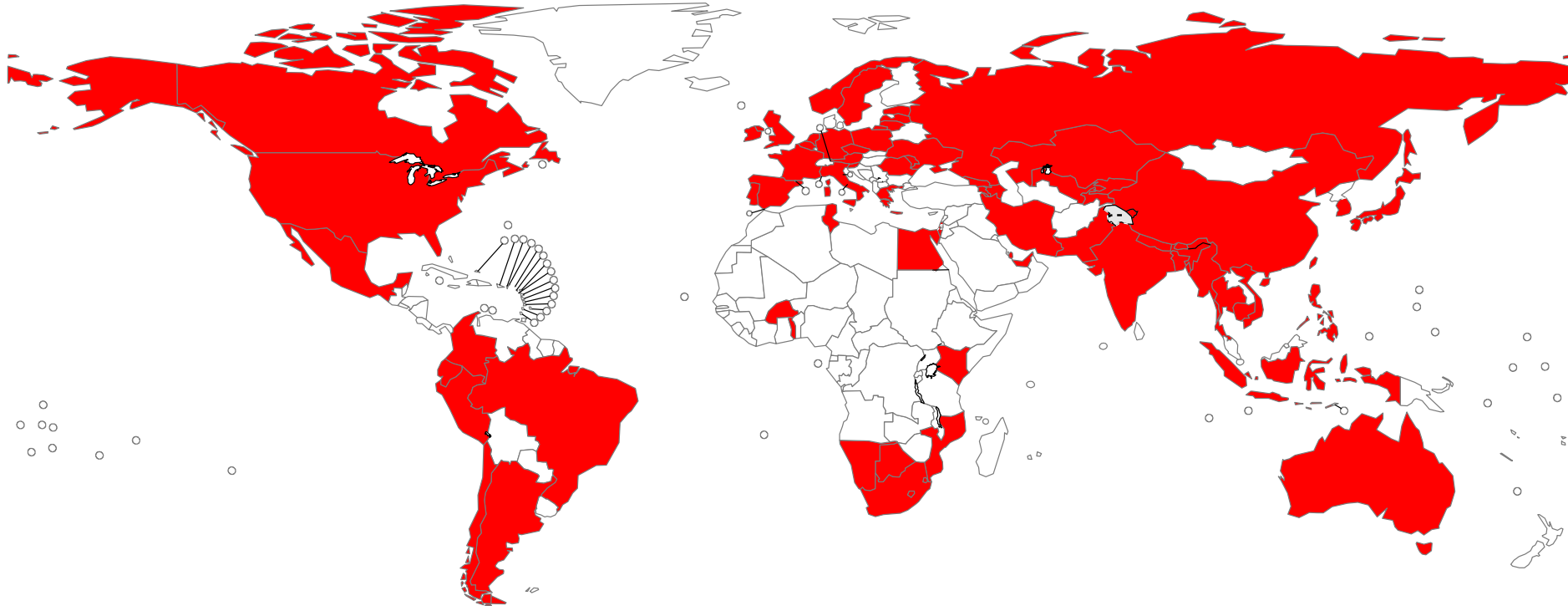
1.7 Million people  
every year

**In India**

Nearly 1000 people  
every day

Two people  
every 3 minutes

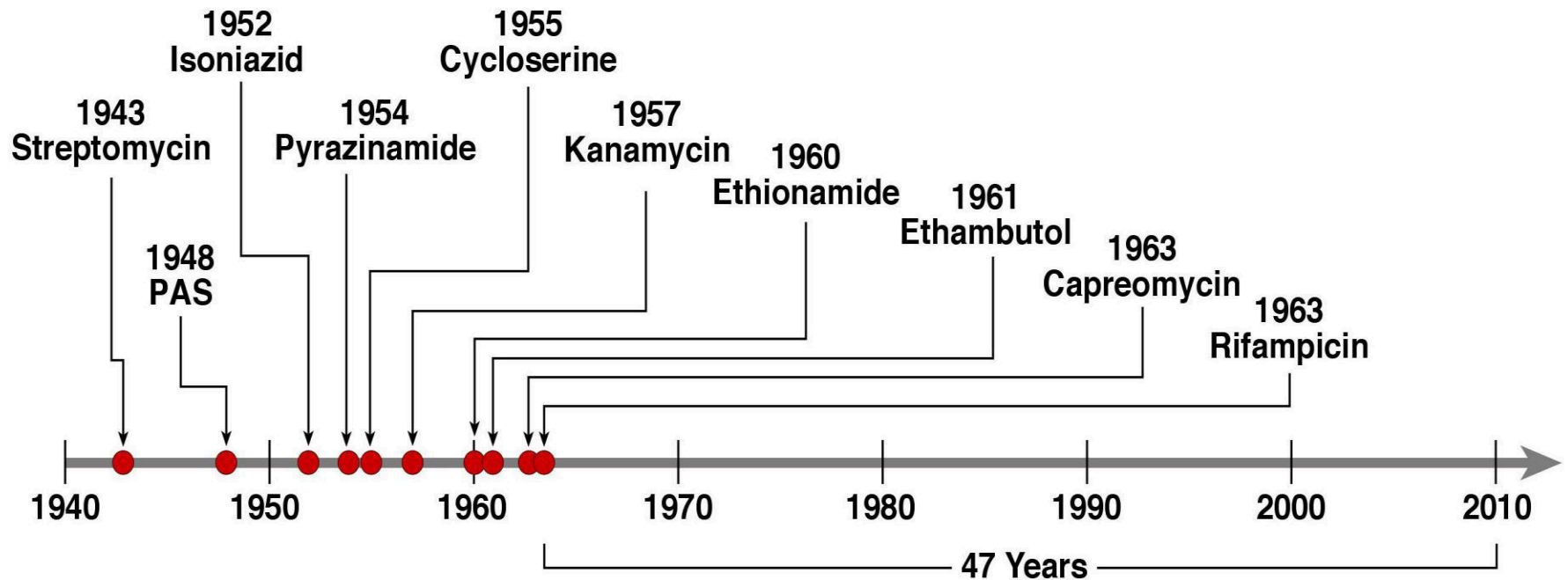
# Countries that had reported at least one XDR-TB case by end March 2011



Argentina	Bhutan	France	Japan	Namibia	Republic of Korea	Thailand
Armenia	Cambodia	Georgia	Kazakhstan	Nepal	Republic of Moldova	Togo
Australia	Canada	Germany	Kenya	Netherlands	Romania	Tunisia
Austria	Chile	Greece	Kyrgyzstan	Norway	Russian Federation	Ukraine
Azerbaijan	China	India	Latvia	Pakistan	Slovenia	United Arab Emirates
Bangladesh	Colombia	Indonesia	Lesotho	Peru	South Africa	United Kingdom
Belgium	Czech Republic	Iran (Islamic Rep. of)	Lithuania	Philippines	Spain	United States of America
Botswana	Ecuador	Ireland	Mexico	Poland	Swaziland	Uzbekistan
Brazil	Egypt	Israel	Mozambique	Portugal	Sweden	Viet Nam
Burkina Faso	Estonia	Italy	Myanmar	Qatar	Tajikistan	

# TB Drug Discovery

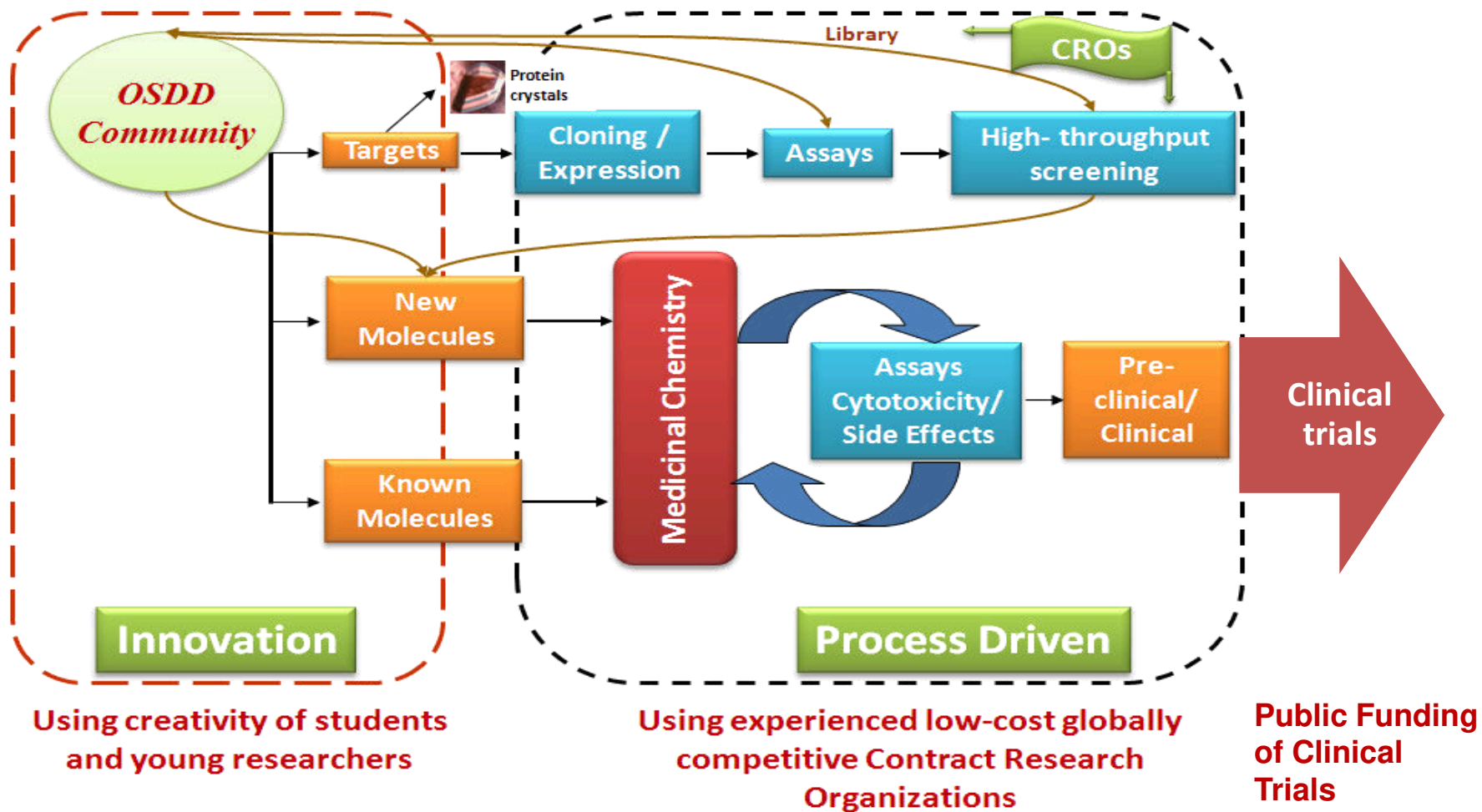
---



# Why Open Source Drug discovery ?

- Many eye balls make the bug shallow!
- Lack of market incentive for TB
- Successful Open Source Models
  - Human Genome Sequencing Initiative
  - Open Source Software Initiative (eg: Linux OS)
  - Android
  - The WWW

# OSDD Process Flow



# Status: OSDD Projects

Drug Target  
Identification

Virtual  
Screening

Chemical  
Synthesis/  
library

Screening/ Hit  
identification

Hit to Lead

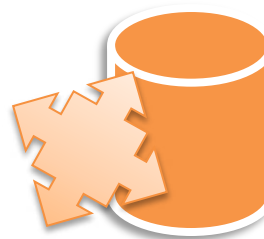
18

19

9

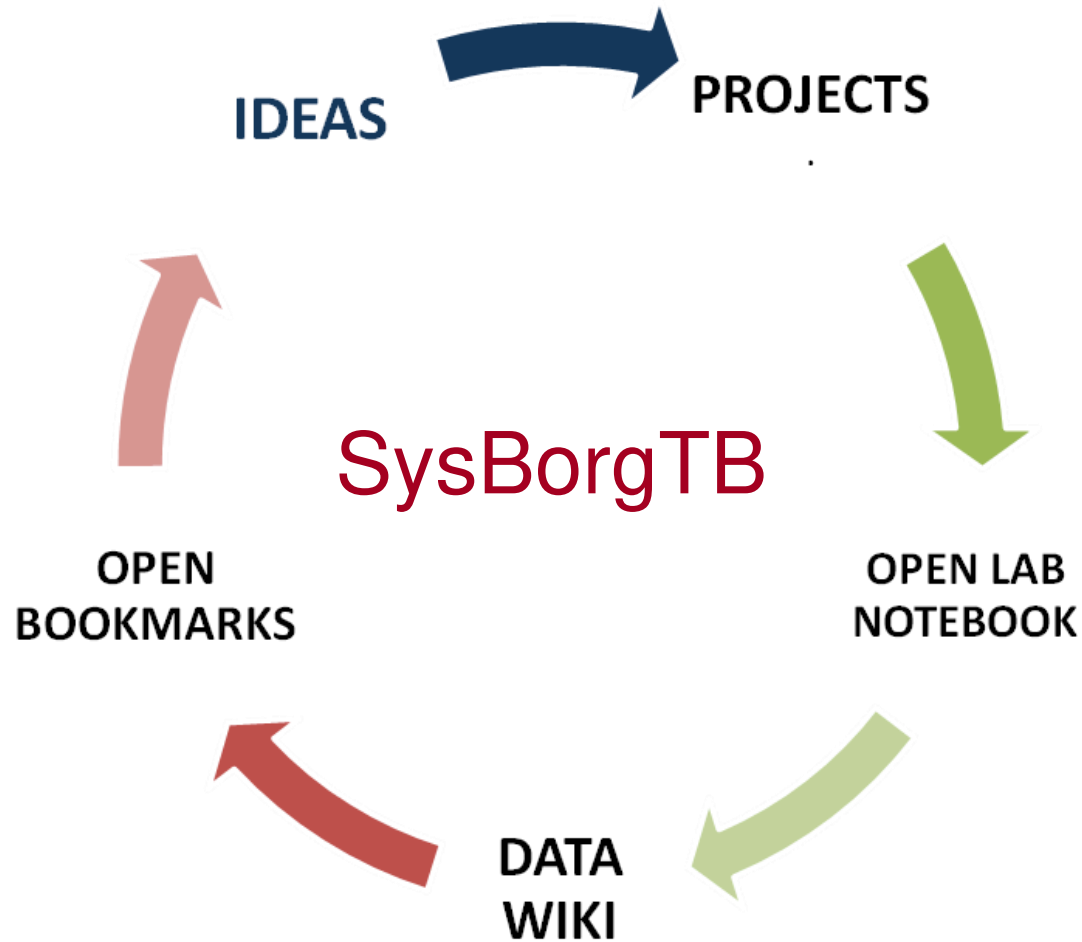
6

2



Other projects aim to  
develop tools,  
databases and  
repositories for the  
OSDD community

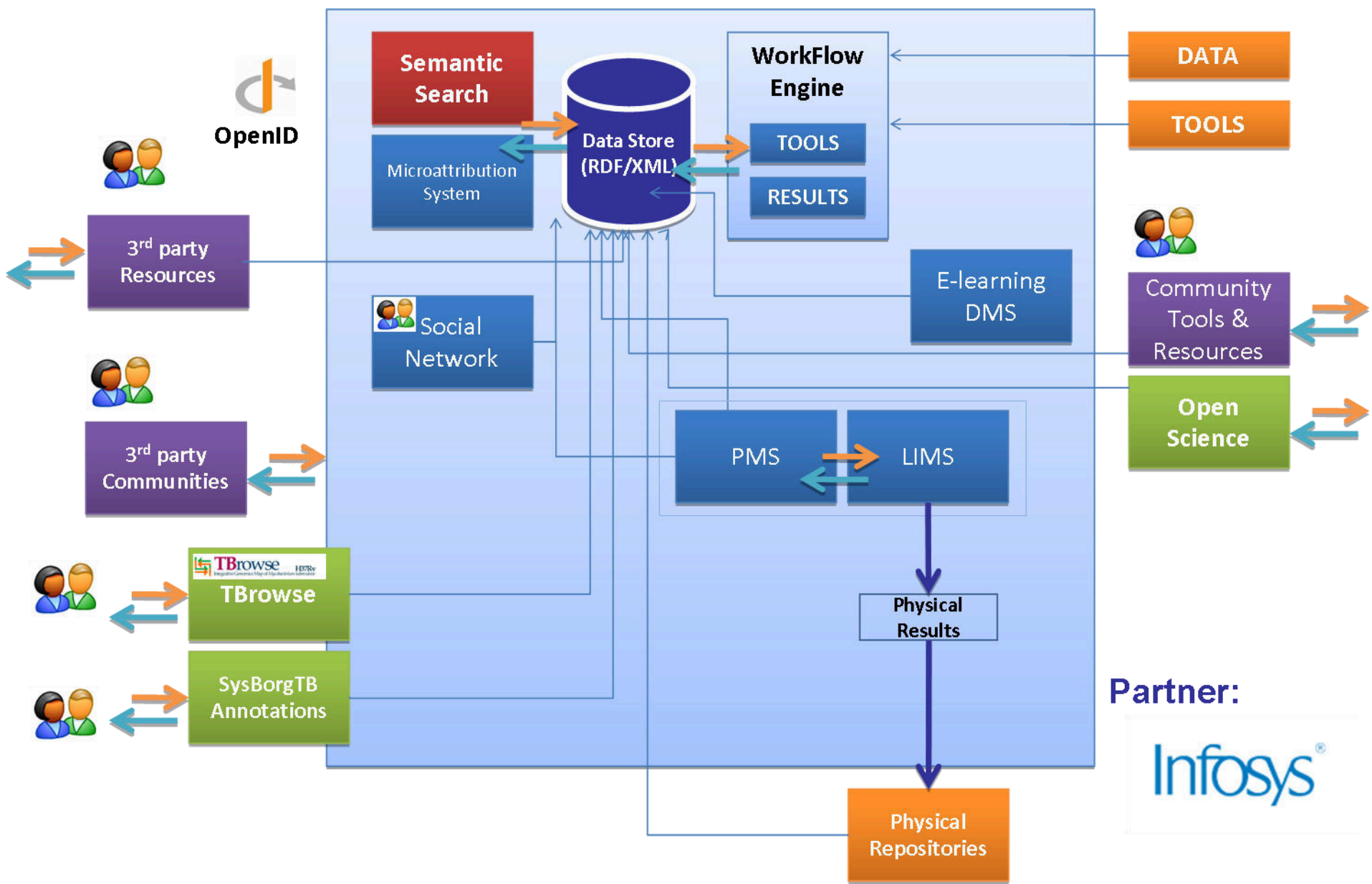
# The OSDD Cycle





# Shaping Science 2.0

## OSDD Semantic Web Architecture



Applications Places System Wed Oct 27, 10:18 AM osdd

SysBorgTB workflows - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://sysborg2.osdd.net:8888/root

Gmail - Inbox - hibeeluck... Facebook SysBorgTB - osdd.net SysBorgTB workflows

**SysBorgTB** Analyze Data Workflow Data Libraries Help User

**Tools**

- Get Data
- Get ENCODE Data
- Send Data
- ENCODE Tools
- Lift-Over
- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- Convert Formats
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Regional Variation
- Multiple regression
- Evolution: HyPhy
- Metagenomic analyses
- Short Read Mapping
- GPSR
- KEGG
- PDB
- EBI
- WSDbfetch
- GetEntry
- ARSA
- Mafft
- Fasta
- Ensembl
- VecScreen
- GO
- NCBIGenomeAnnotation
- TxSearch
- OMIM
- RefSeq
- GIBIS
- GTRFav
- Done

**OSDD**

OSDD is a CSIR-led global initiative with a vision to provide affordable healthcare to the developing world by providing a global platform where the best minds can collaborate & collectively endeavor to solve the complex problems associated with discovering novel therapies for neglected tropical diseases like Malaria, Tuberculosis, Leshmaniasis, etc. It is a concept to collaboratively aggregate the biological and genetic information available to scientists in order to use it to hasten the discovery of drugs. This will provide a unique opportunity for scientists, doctors, technocrats, students and others with diverse expertise to work for a common cause.

**History** Options

refresh | collapse all

Unnamed history

Add tags to history

Your history is empty. Click 'Get Data' on the left pane to start

**Galaxy provides -**

- ✓ Simplified GUI design
- ✓ Ease of integrating modules
- ✓ Fewer components for creating workflows
- ✓ Sharable workflows for better collaboration

**Galactic quickie # 5**

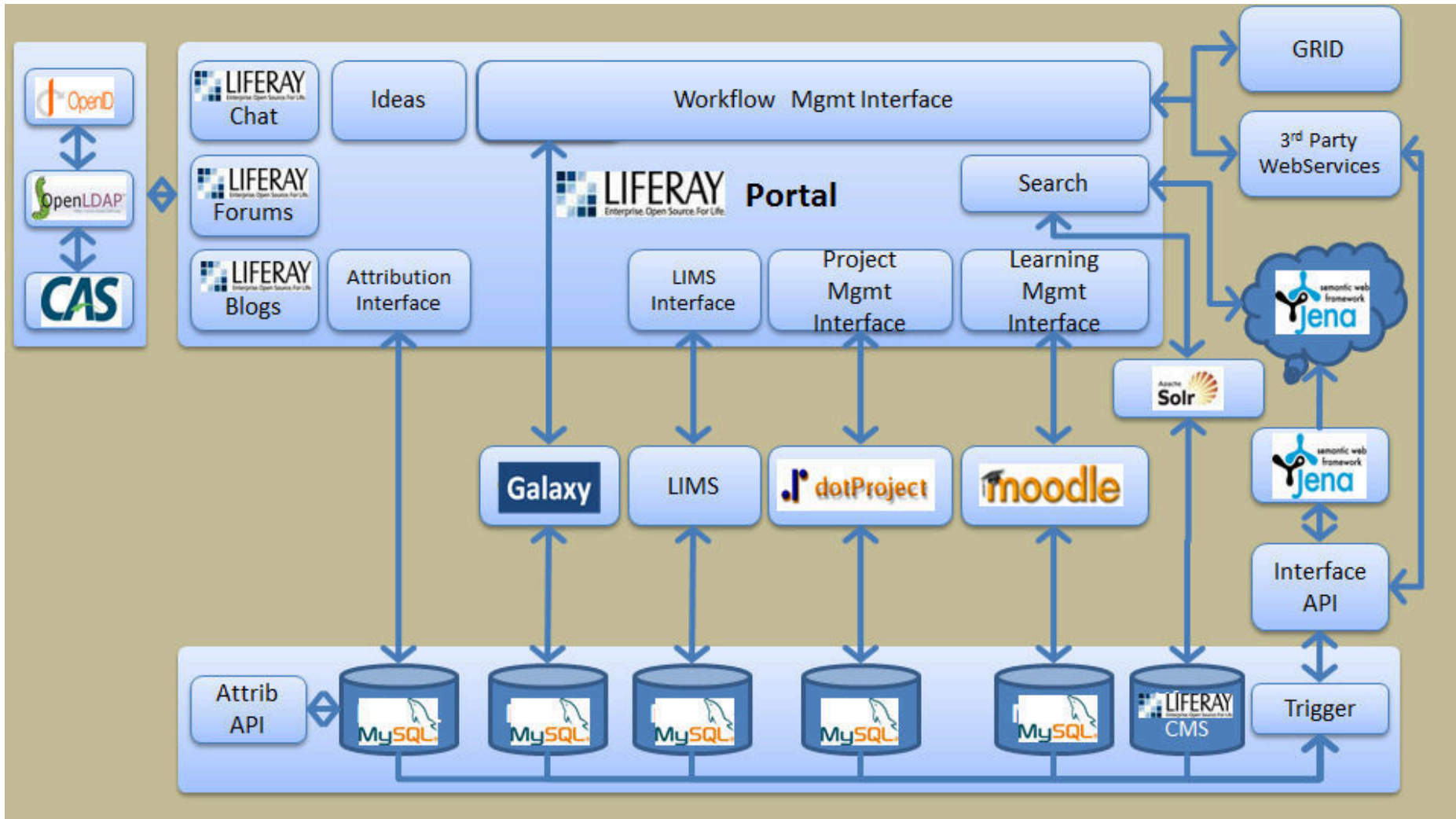
**Previous Quickies**

- Sequences as Tab delimited data  
Galactic quickie # 1
- Grouping  
Galactic quickie # 2
- A word about Interval data  
Galactic quickie # 3
- What's  
Galactic quickie # 4

The Galaxy team is a part of EMBL at European Bioinformatics Institute

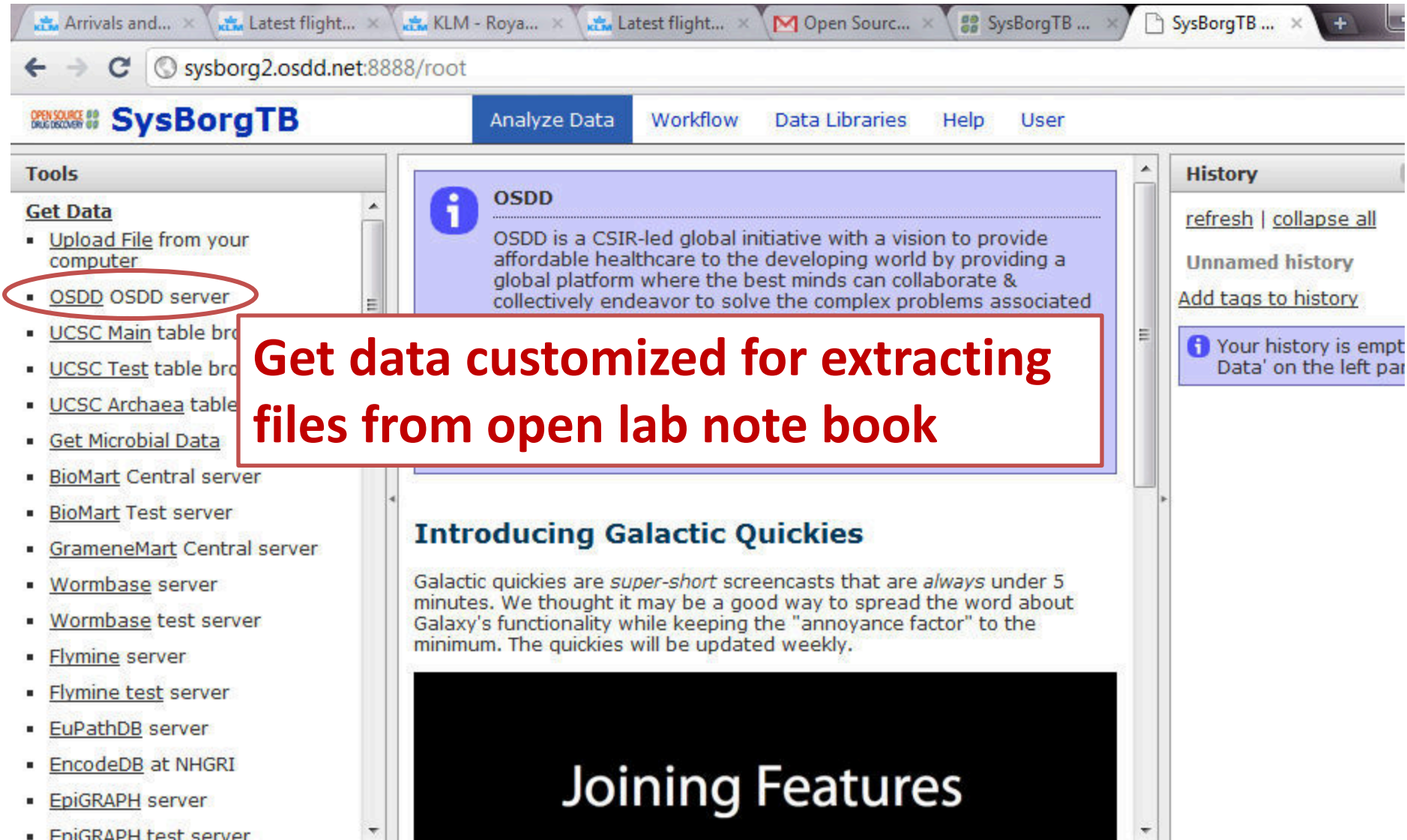
SysBorgTB workflow... PENDRIVE

# System Architecture



Collaborative tools to accelerate neglected diseases research” in the book “Collaborative Computational Technologies for Biomedical Research”. Wiley and Sons. May 2011

# Custom APIs for importing input files from OSDD's open lab note book into Galaxy



The screenshot shows the SysBorgTB web interface. The browser address bar displays 'sysborg2.osdd.net:8888/root'. The interface includes a navigation bar with 'Analyze Data', 'Workflow', 'Data Libraries', 'Help', and 'User'. On the left, the 'Tools' section is expanded to 'Get Data', listing various data sources. 'OSDD OSDD server' is circled in red. The main content area features an 'OSDD' information box and a section titled 'Introducing Galactic Quickies'. A red-bordered box is overlaid on the page, containing text about customizing data extraction. The right sidebar shows a 'History' section with a message indicating an empty history.

**Get Data**

- Upload File from your computer
- OSDD OSDD server**
- UCSC Main table browser
- UCSC Test table browser
- UCSC Archaea table browser
- Get Microbial Data
- BioMart Central server
- BioMart Test server
- GrameneMart Central server
- Wormbase server
- Wormbase test server
- Flymine server
- Flymine test server
- EuPathDB server
- EncodeDB at NHGRI
- EpiGRAPH server
- EpiGRAPH test server

**OSDD**

OSDD is a CSIR-led global initiative with a vision to provide affordable healthcare to the developing world by providing a global platform where the best minds can collaborate & collectively endeavor to solve the complex problems associated

**Introducing Galactic Quickies**

Galactic quickies are *super-short* screencasts that are *always* under 5 minutes. We thought it may be a good way to spread the word about Galaxy's functionality while keeping the "annoyance factor" to the minimum. The quickies will be updated weekly.

**Joining Features**

**History**

[refresh](#) | [collapse all](#)

Unnamed history

[Add tags to history](#)

**i** Your history is empty  
Data' on the left panel



# Custom APIs for exporting results to OSDD's Open lab note book

The screenshot shows the SysBorgTB web interface. The browser address bar displays 'sysborg2.osdd.net:8888/workflow'. The navigation bar includes 'Analyze Data', 'Workflow' (selected), 'Data Libraries', 'Help', and 'User'. The main content area is titled 'Your workflows' and features a table with columns 'Name', '# of Steps', and 'Actions'. A workflow named 'GAP1' with 4 steps is listed. The 'Actions' column for 'GAP1' contains two buttons: 'Export WorkFlow To OSDD Lab Notebook' and 'Export WorkFlow Results To OSDD Lab Notebook'. Both buttons are circled in red. Below this table is a section 'Workflows shared with you by others' with a table showing a workflow named 'Workflow Deblina' owned by 'deblina' with 7 steps. At the bottom, there is an 'Other options' section with a button 'Configure your workflow menu'. Two red arrows originate from the bottom of the slide and point to the two export buttons in the 'GAP1' workflow row.

Name	# of Steps	Actions
GAP1 ▼	4	<a href="#">Export WorkFlow To OSDD Lab Notebook</a> <a href="#">Export WorkFlow Results To OSDD Lab Notebook</a>

Name	Owner	# of Steps
Workflow Deblina ▼	deblina	7

- ✓ Workflows and the result of the workflows are stored as separate lab note books
- ✓ Lab note book has details of the experiments performed
- ✓ Results of one experiment may be invoked for analysis in another experiment
- ✓ All versions of the workflow and the results are stored
- ✓ Flexibility to execute nested workflows

# Our Approach : Data & Tool integration

In addition to access heterogeneous sources of data like BioMart Central/UCSC Table Browser (<http://genome.ucsc.edu/>), Open lab note book of <http://sysborg2.osdd.net> is interfaced with Galaxy



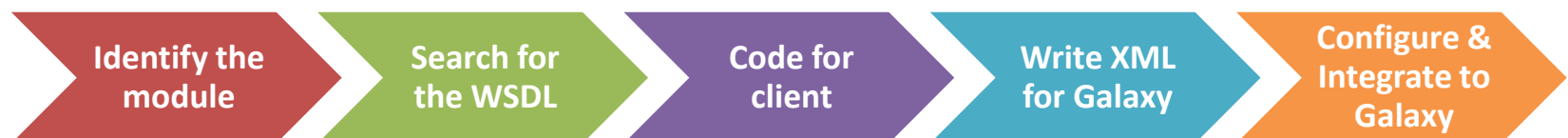
## *Standalone databases and tools*



## *Tools as web services:*

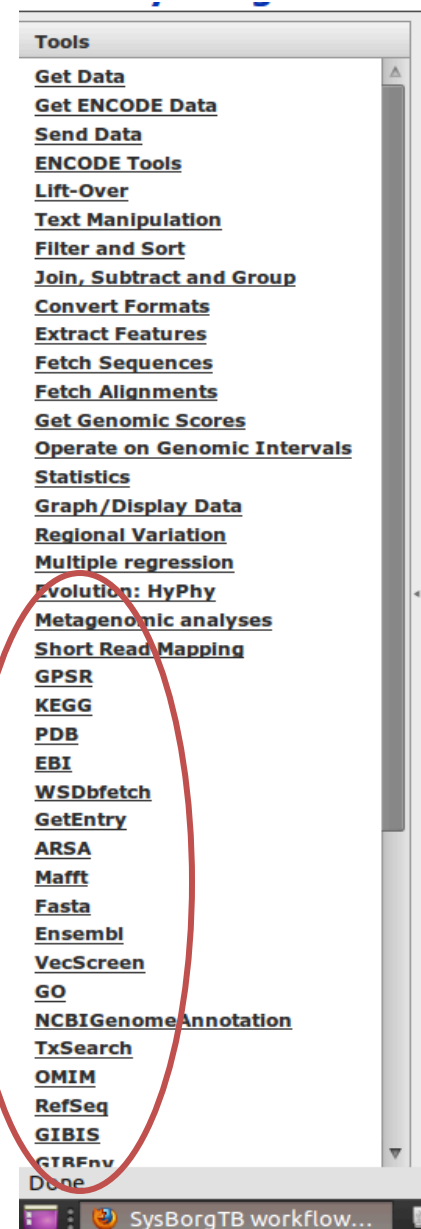
- Web services can be added as tools in Galaxy
- Extends the potential of galaxy workflows

## *The process*



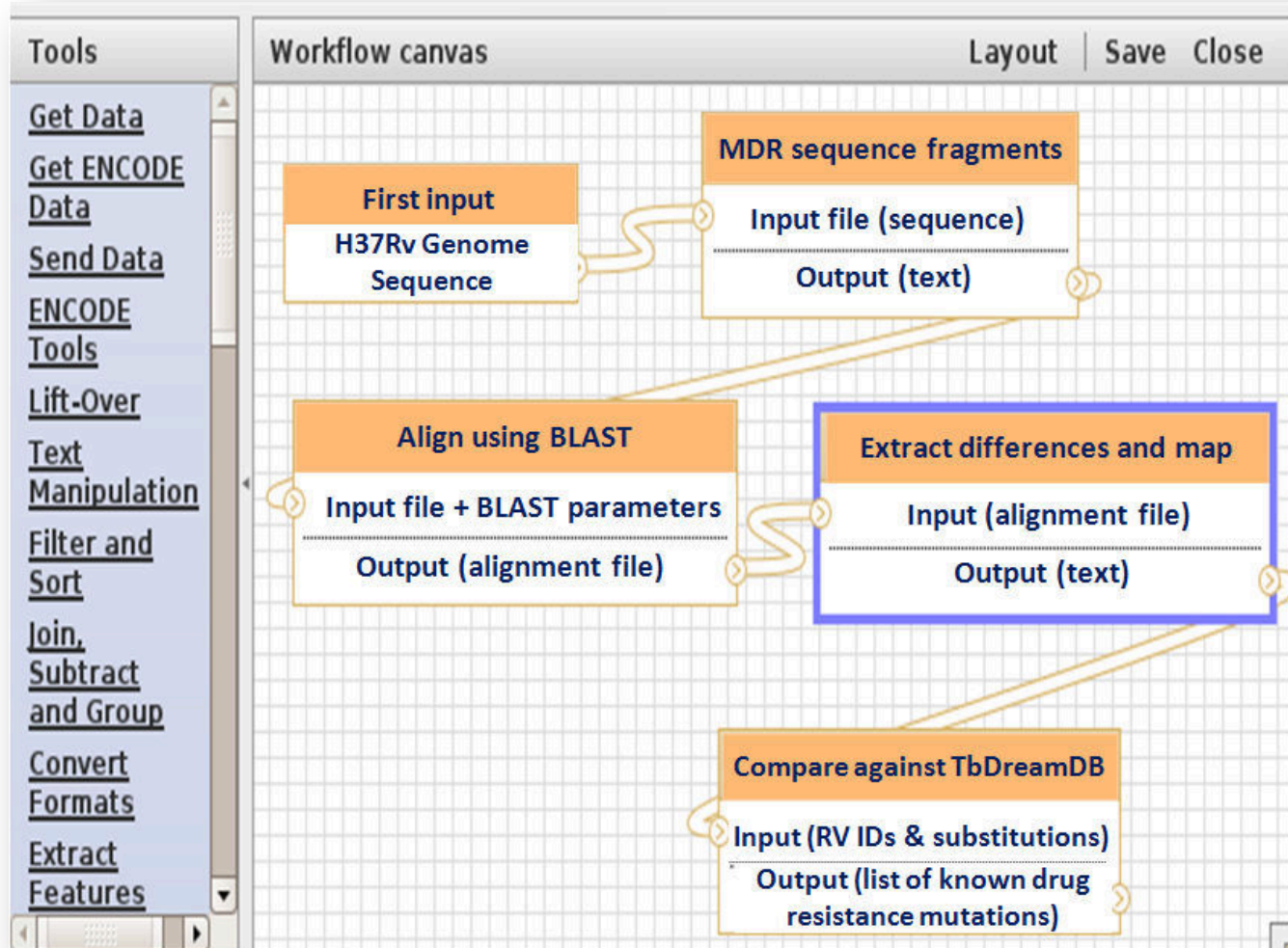
# ChemBio toolkit : Modules integrated by OSDD Community

S. No	Resources	Clients
1	KEGG: Kyoto Encyclopedia of Genes and Genomes	60
2	GetEntry: DDBJ sequence search by accessionID	43
3	GPSR : tools	33
4	PDB : Protein Data Bank	30
5	BioModel:mathematical models of biological DB	25
6	Gtps : Gene Trek in Prokaryote Space	8
7	WSDbfetch: retrieve entries from biological dbs using entry identifiers or accession no.	7
8	Gibv: Genome Information Broker for Viruses	7
9	DDBJ :DNA Data bank of Japan	7
10	Mafft: a multiple sequence alignment program	4
11	Fasta:- DDBJ database	4
12	Ensembl : maintains automatic annotation	4
13	VecScreen vector contamination	4
14	OMIM:Online Mendelian Inheritance in man	4
15	Gtop: Gene-product Informatics	3
16	GO: Gene Ontology	3
17	SPS : Splicing Profile based Score	2
18	GIBIS: Genome Information Broker for Insertion Sequence	1
19	RefSeq: database of sequence	1
20	GIB: Genome Information Broker	1
21	GIBEnv- DDBJ database	1
22	TxSearch: Database indexing & searching	1



# Enabling Complex Computational Analysis For Experimental Biologists/Chemists

Q. Find novel genes and mutations & map known drug resistance mutations on genome of an MDR-TB strain





# OSDD Community suggests tools for integration in Galaxy

sysborg2.osdd.net/group/sysborgtb/tools

HOME MyPLACES

OPEN SOURCE DRUG DISCOVERY

**SysBorg 2.0** Systems Biology *Mycobacterium tuberculosis*

Welcome Anshu E

SysBorgTB Attribution eLearning Skills Search Workflows DB Aggregation Reports

### Tools Registration

Register new application/tool here... [View My Approvals](#) [Show Files For Uploading](#)

App Name

App Description

URL

XML file   No file chosen

Client (WS)/Executable   No file chosen

Publication/Abstract

### Tools Search

Keyword

PubMed ID

### Popular Contributors

Contributor Details is not Available

Showing 0 results.

### Popular Tools

Tool	
Cut1	Used in 3 Workflows
get_compounds_by_enzyme	Used in 2 Workflows
get_reactions_by_pathway	Used in 1 Workflows
mapToUCSC	Used in 1 Workflows
get best best neighbors by gene	Used in 1

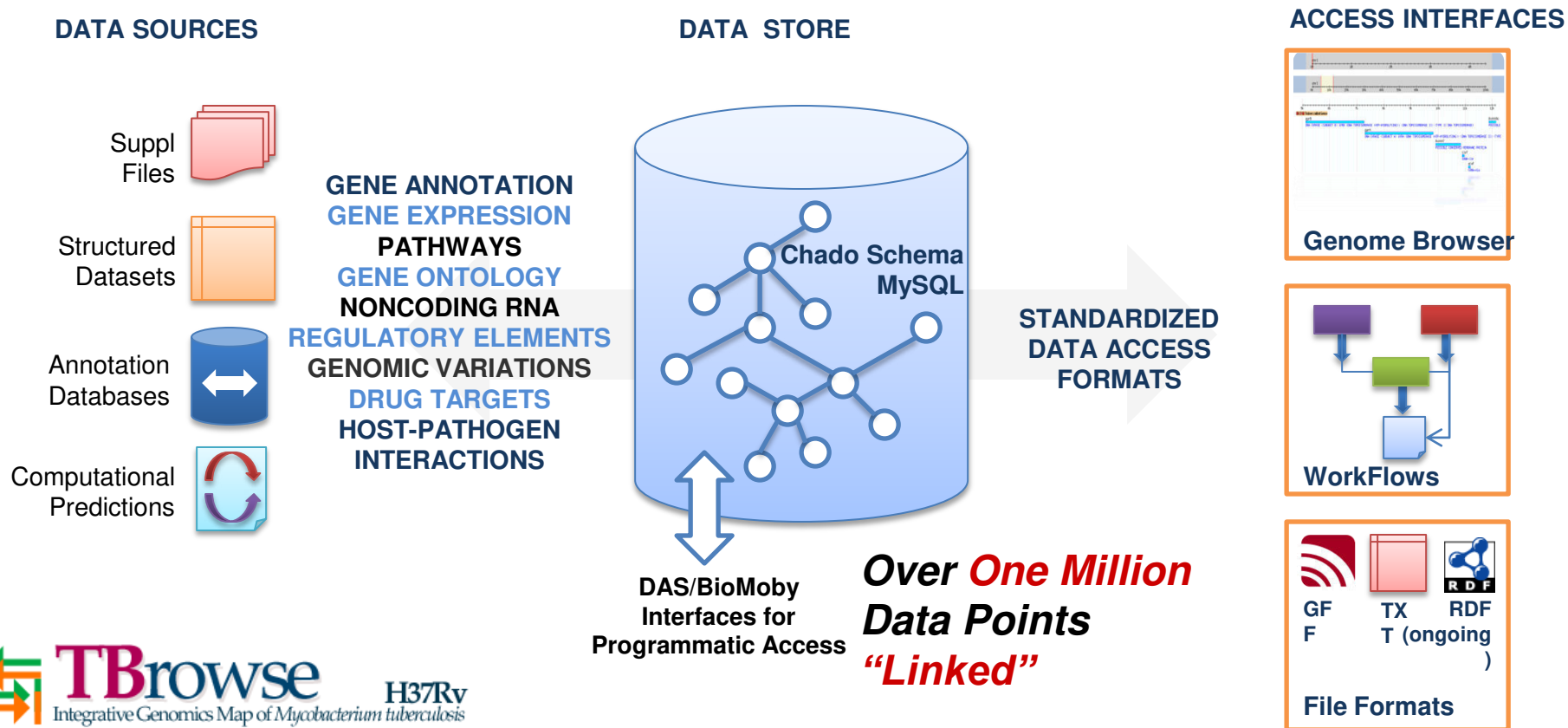
Settings Online Friend

# **OSDD Portal's Galaxy features**

- ✓ **Login customized with OSDD's single sign-on**
- ✓ **Get data customized for extracting files from open lab note book**
- ✓ **Workflows and the result of the workflows are stored as separate lab note books**
- ✓ **Lab note book has details of the experiments performed**
- ✓ **Results of one experiment may be invoked for analysis in another experiment**
- ✓ **All versions of the workflow and the results are stored**
- ✓ **Flexibility to execute nested workflows**

**More than 200 standalone programs & web service clients**

# Integrating post-Genomic Datasets and Resources for *Mycobacterium tuberculosis*



**TBrowse** H37Rv  
Integrative Genomics Map of *Mycobacterium tuberculosis*

<http://tbrowse.osdd.net>

Bhardwaj A, Bhartiya D, Kumar N, Open Source Drug Discovery Consortium, Scaria V  
**TBrowse: an Integrative Genomics Map of *Mycobacterium tuberculosis*.**  
*Tuberculosis* - 17 August 2009

**OPEN SOURCE**  
**DRUG DISCOVERY**

AnshuBhardwaj

DeekshaBhartiya

## Publication in Tuberculosis 2009

NitinKumar

VinodScaria

Tuberculosis



GENERAL

### TBrowse: An integrative genomics map of *Mycobacterium tuberculosis*

Anshu Bhardwaj<sup>a,b</sup>, Deeksha Bhartiya<sup>a</sup>, Nitin Kumar<sup>b</sup>,  
Open Source Drug Discovery Consortium<sup>b</sup>, Vinod Scaria<sup>a,b,\*</sup>

<sup>a</sup> GN Ramachandran  
<sup>b</sup> The Open Source

its lists available at ScienceDirect  
Tuberculosis  
journal homepage: [www.elsevierhealth.com/locate/tube](http://www.elsevierhealth.com/locate/tube)



Deeksha Bhartiya



Nitin Kumar


## OpenLabNoteBook on SysBorgTB

<http://sysborgtb.osdd.net/bin/view/OpenLabNotebook/TBMapDataset>

### ExperimentReportForm

Title	TBMap Dataset in standard format (GFF) for <i>Mycobacterium tuberculosis</i>
<a href="#">LinkedToProject</a>	GenomeBrowser?
Materials	Genome-level information extracted from individual databases and supplementary information of publications
Methods	Genomic datam is systematically culled from online resource and publications and is organized into more than ninety tracks. The resource is built based on the Generic Model Organism Database Genome Browser. This ensures interoperability with other genome browser installations.The datasets were then uniformly transformed, tagged and reformatted confirming the Sequence Ontology annotation ontologies, wherever applicable.
Results	The database hosts more than half a million data points encompassing different genomic features and computational predictions encompassing over ninety datasets.
Discussion	TBMap aims to provide a standard and integrative resource to facilitate integrative annotation of the tuberculosis genome.

# Comparison of Browsers

s.no.	Source	Tracks
1	UCSC Genome Browser on Mycobacterium tuberculosis H37Rv 06/20/1998 Assembly	6
2	WebTb	Operon Map
3	Argo Genome Browser	not web based
4	PGBrowser: Pathogen Genome Browser	3
5	BioHealthBase	16
6	Ensembl	~15
7	 <b>TBrowse</b> <small>Integrative Genomics Map of <i>Mycobacterium tuberculosis</i> H37Rv</small>	100

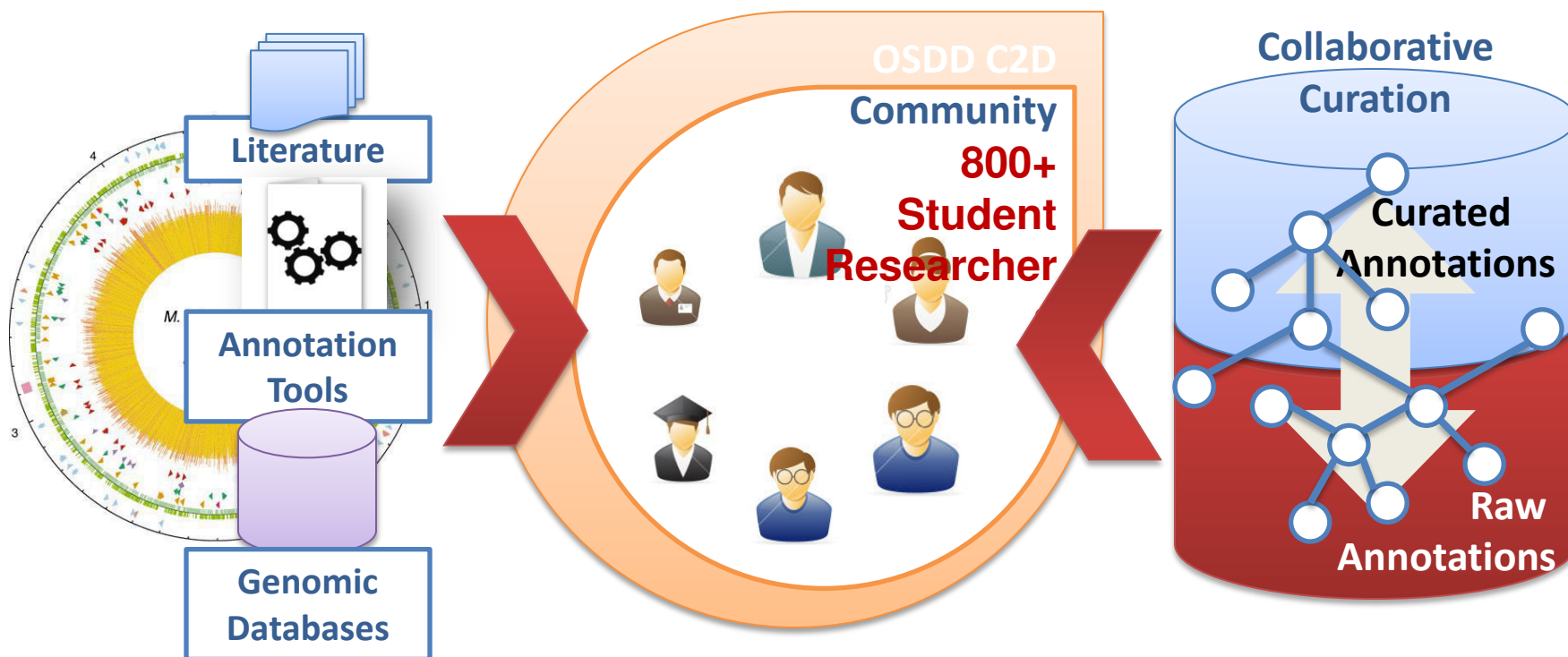


# The “Connect to Decode” Programme

OPEN SOURCE  
DRUG DISCOVERY

2

CONNECT to DECODE  
SUMMER RESEARCH PROGRAMME '10



Pathway/Interactome | Gene Ontology | Protein  
Structure/Fold | Glycomics | Immunome



IGIB  
INSTITUTE OF GENOMICS  
& INTEGRATIVE BIOLOGY

SBI

The  
Systems  
Biology  
Institute



IMTECH  
Chandigarh

# OSDD Community Effort to further the understanding of the biology of Mtb



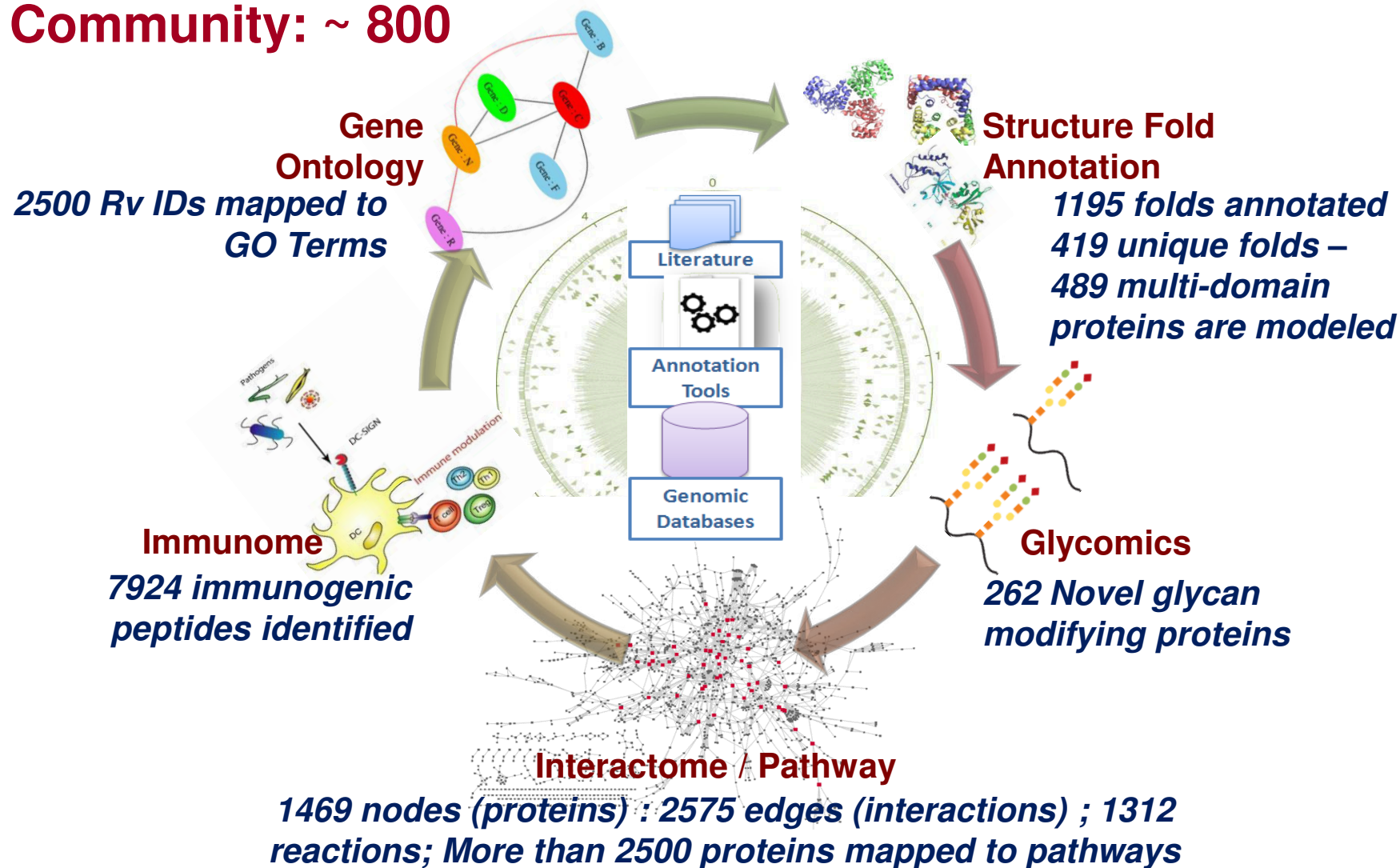
OPEN SOURCE  
DRUG DISCOVERY



CONNECT to DECODE  
SUMMER RESEARCH PROGRAMME

'10

**Community: ~ 800**



Within weeks, 830 volunteered to re-annotate the entire *M. tuberculosis* genome. The work started in December 2009 and was completed by April 2010, **packing nearly 300 man-years into 4 months!**

***Source: Munos B. Can Open-Source Drug R&D  
Repower Pharmaceutical Innovation?  
Clin Pharmacol Ther 2010;87:534–536***

Clinical Pharmacology  
& Therapeutics



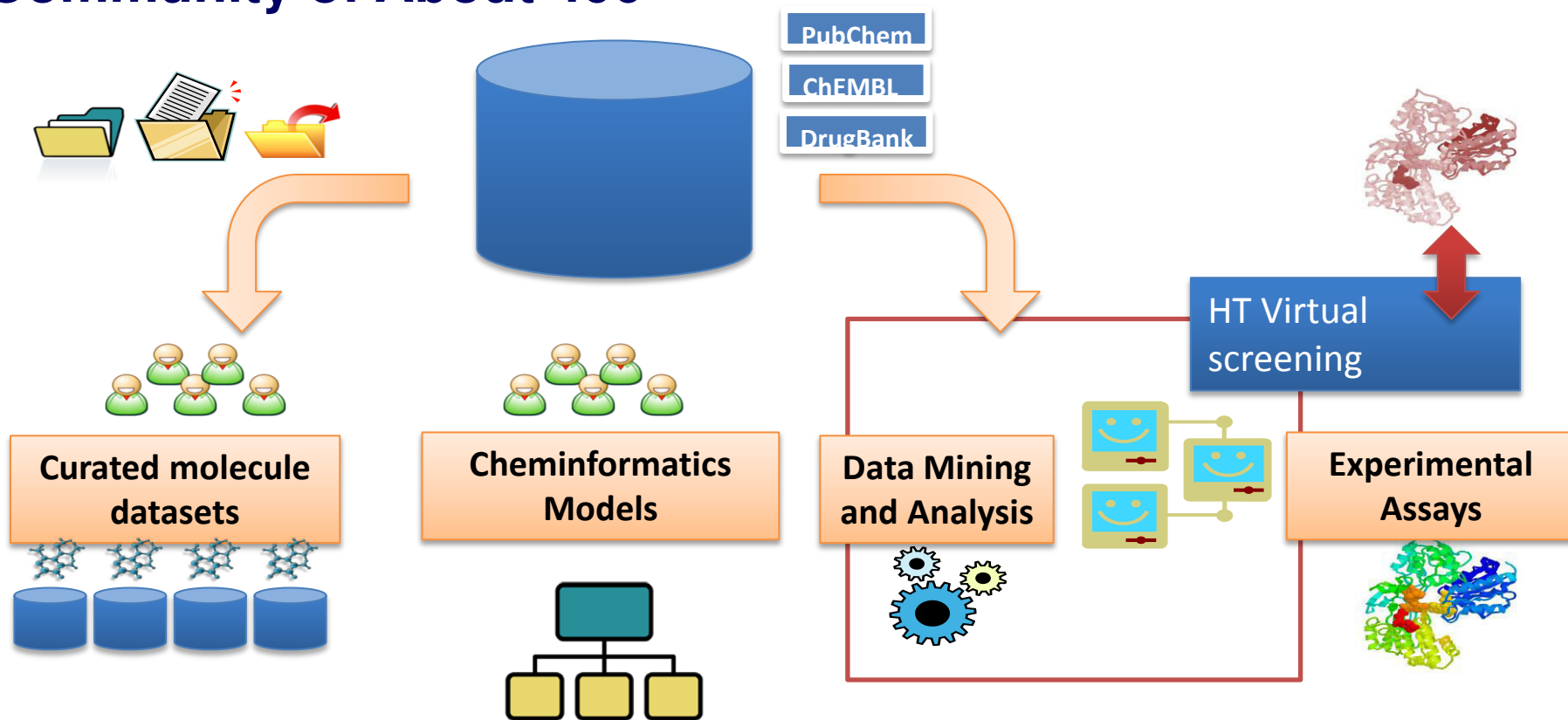
The Open-Source Drug Discovery (OSDD), which is dedicated to discovering treatments for diseases that plague the developing world, surprised many when more than 400 of its volunteer researchers reannotated the tuberculosis bacterium genome, wiki style, in just 4 months—record time for such an endeavor

***Source: Margie Patlak. Open-Source Science Makes Headway  
J Natl Cancer Inst. 2010 Aug 18;102(16):1221-3***



# Ongoing: Cheminformatics

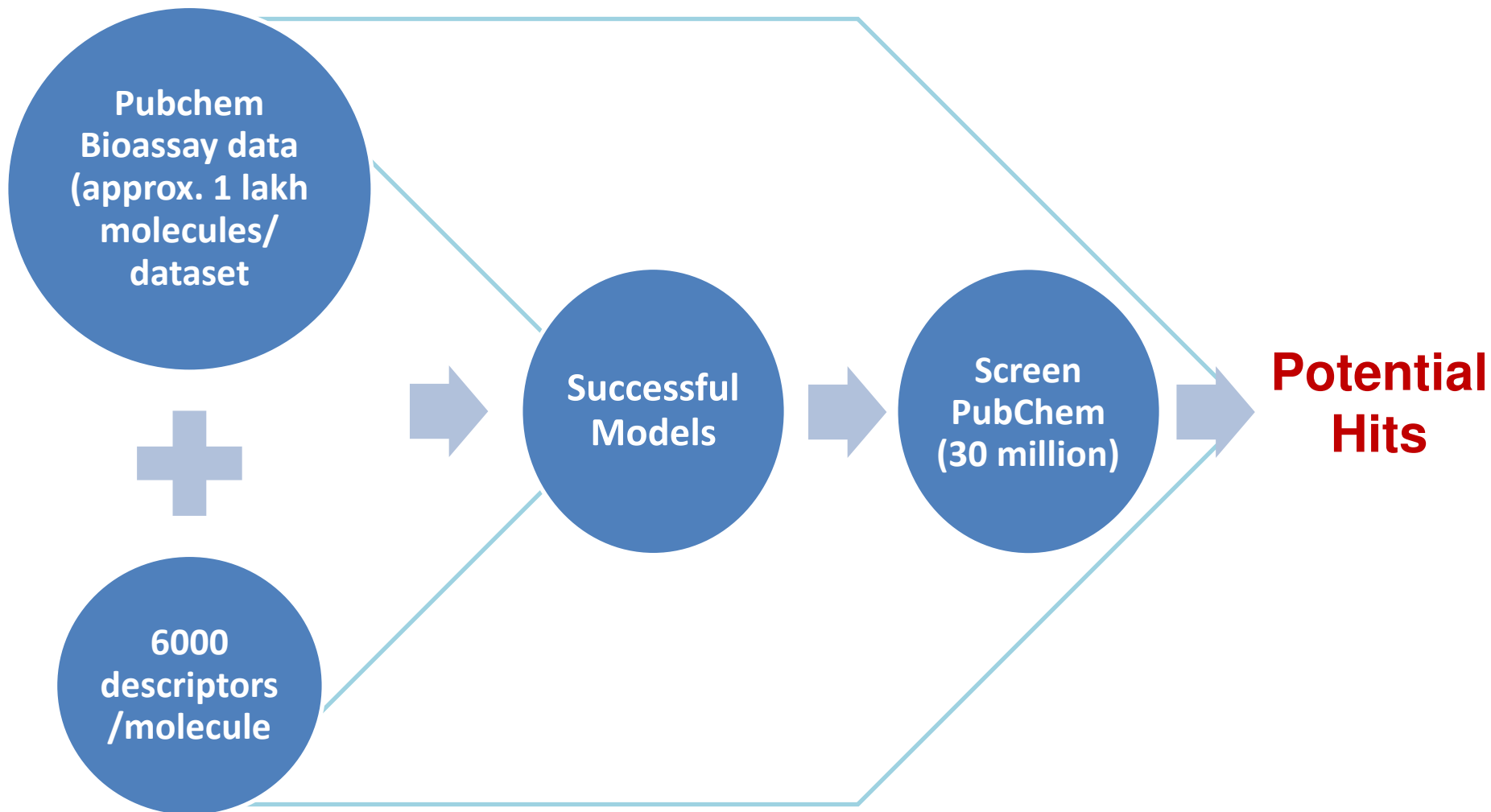
Community of About 400



Other Active Communities:

- OSDD Women Scientists Forum
- OSDD Junior Scientists Forum

# Data amplification in Cheminformatics



- Down sizing and random validation require multiple calculation for validation of results
- Cross validation up to 50+ time for each experiment

# **cheminformatics**

## **Project details**

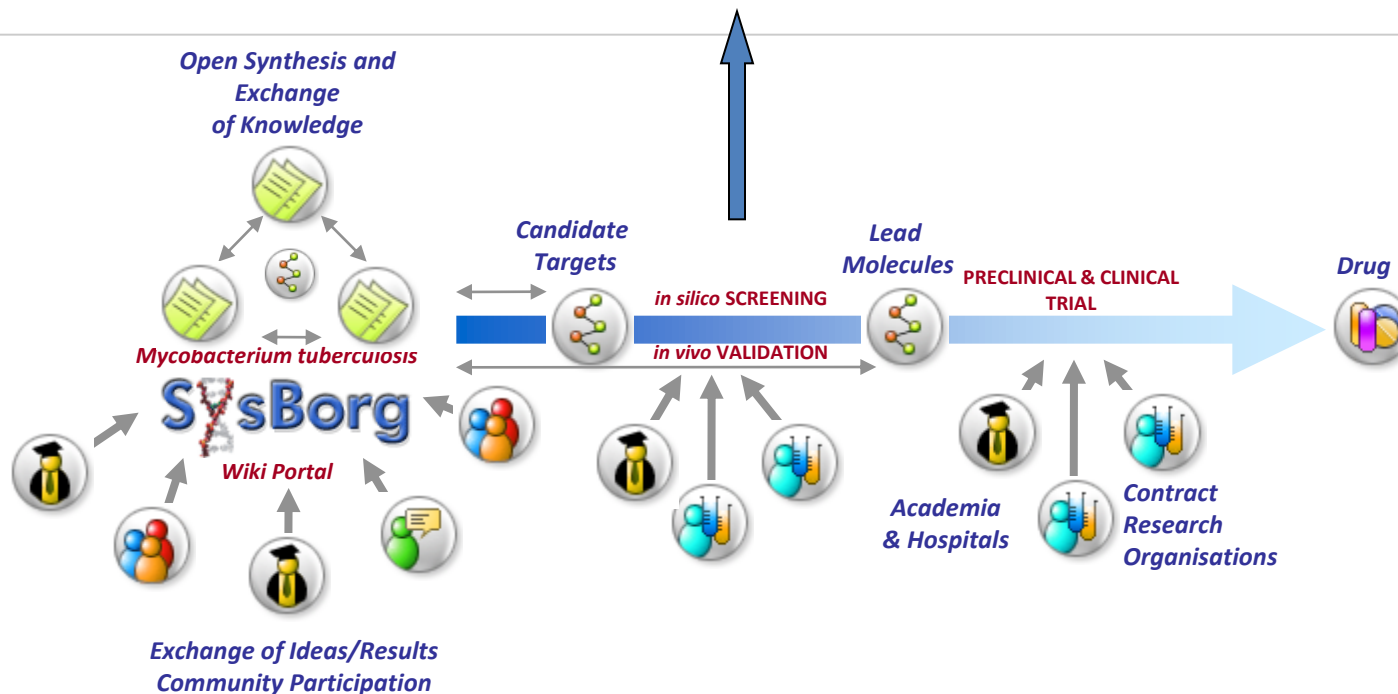


- **Project started June 2010**
- **Number of participants enrolled 400+**
- **Number participants still active 75+**
- **12 groups each with a Team Leader**

# The Problem



# Role of Garuda Grid



## PARTNER ORGANISATIONS



Institute of  
Microbial  
Technology



Jawaharlal  
Nehru  
University



ICMR-JALMA



CDFD



AU-KBC



CDRI



invent



Sun  
microsystems



LEADINVENT  
Inventing Leads



Institute of  
Life  
Sciences



Infosys



TCG  
Lifesciences  
Enabling Translational Medicine



CAMBIA



INDIA800  
FOUNDATION  
Building One India



premas  
biotech  
partners in discovery



SBI  
The  
Systems  
Biology  
Institute



National Institute  
of Immunology



National Institute  
of Immunology



MACMILLAN  
SCIENTIFIC COMMUNICATIONS  
An exclusive partner of nature publishing group

## Lead Organization

Council of Scientific and  
Industrial Research (CSIR), India

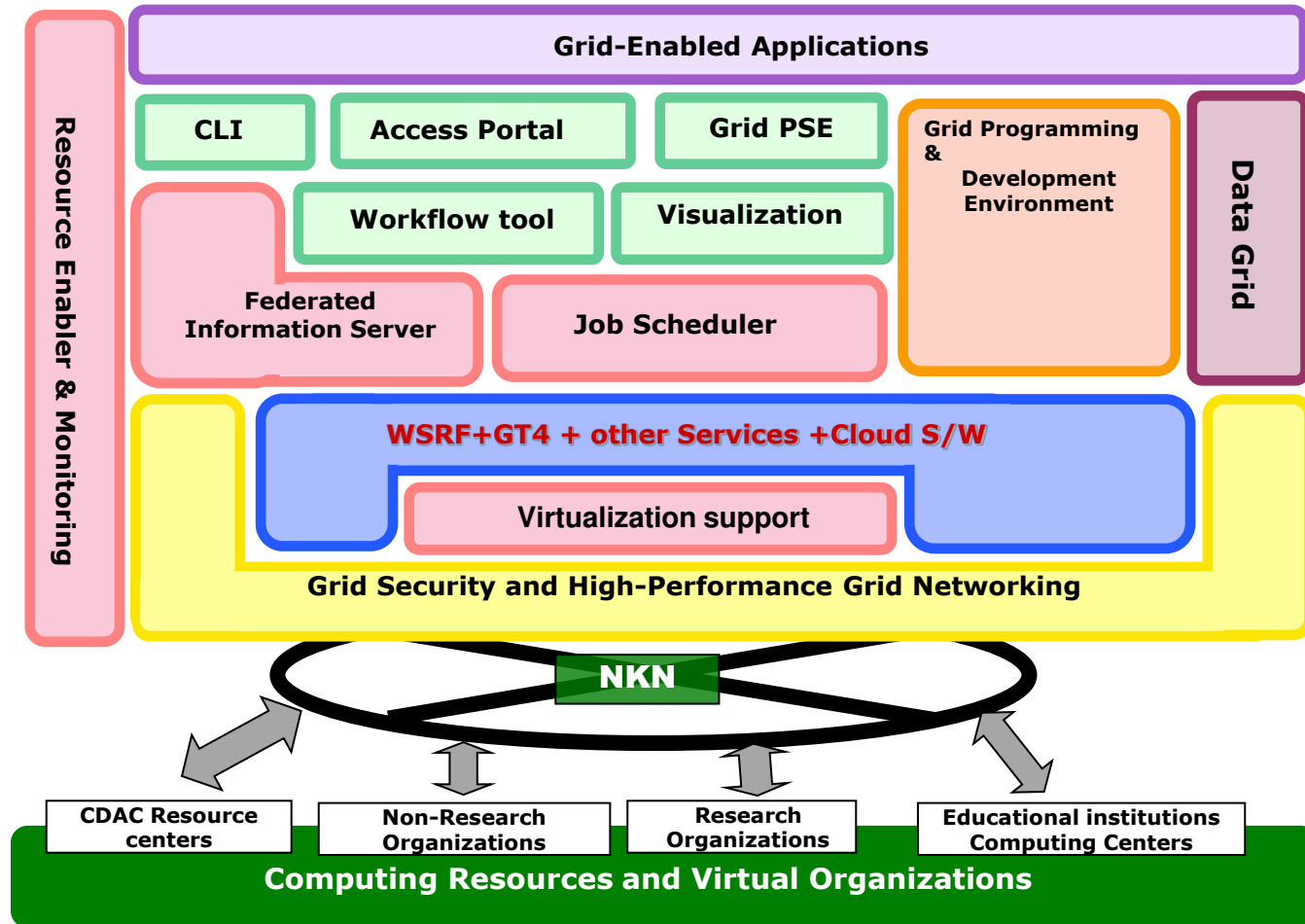


# C-DAC's Garuda Grid – Indian Grid Computing Initiative

- C-DAC is R&D organization under Ministry of Communication & Information Technology, India
- C-DAC's Garuda Grid is targeted at providing **a facility for the scientific community**, which would enable them to seamlessly access the distributed resources.
- **Compute Power of GARUDA: ~ 70TFs (6000 CPUs)**
- Currently there are **55 Garuda Partners**
- Has **NKN** (National Knowledge Network) connectivity at 10Gbps



# GARUDA Grid: Architecture



# Features:

## Customized Galaxy on GARUDA

- **Integrated with Grid Authentication mechanism - Indian Grid Certificate Authority (IGCA)**
- **Integrated with Gridway Metascheduler - Job scheduling and management**
- **Integrated OSDD tools - Weka (for data mining) and Autodock (Virtual screening).**
- **Provided support to upload multiple input files as tar file**
- **Data libraries of OSDD community are uploaded and are shared by all users**
- **Integrated with PostgreSQL**



# OSDD-Garuda Interface



Internet/NKN



Results



## GARUDA GRID

Garuda Middleware Stack

LRM- Torque

GGHYD Cluster

Yuva Cluster

JNU Cluster

Other OSDD Cluster

OSDD Tools – weka, cdk, ...

NKN



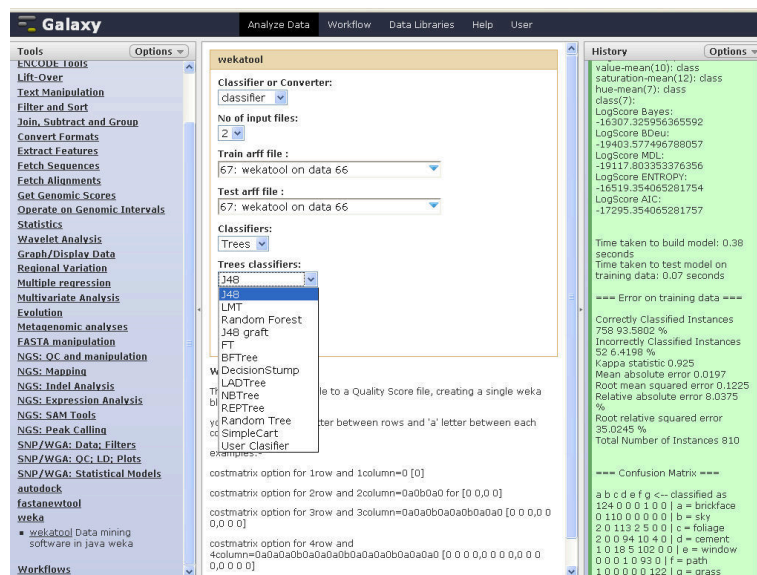
## OSDD HeadNode

Garuda Middleware Stack, login service, Gridway Metascheduler

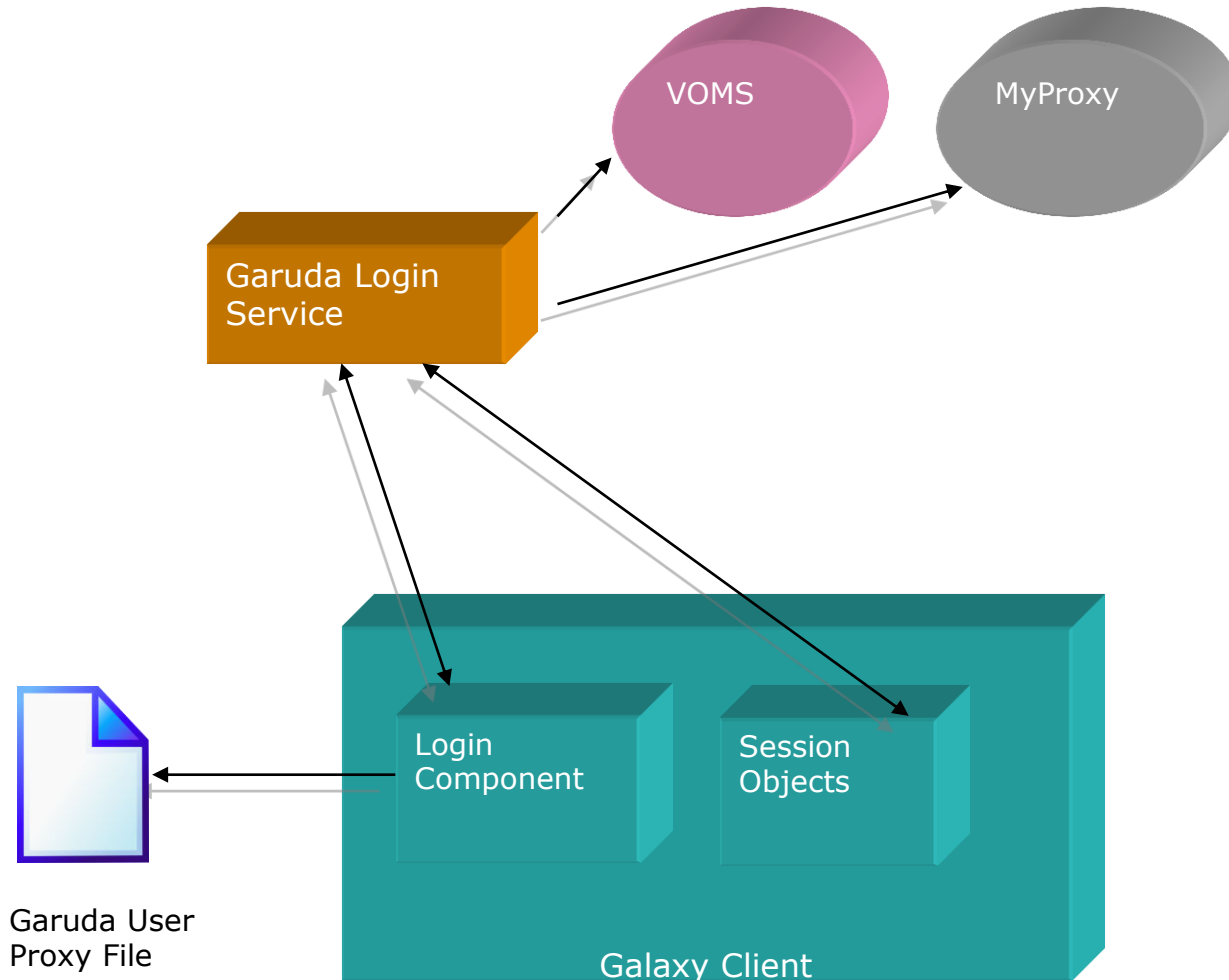
OSDD Customized Galaxy



Figure 1: Login page for OSDD-GARUDA Interface



# Garuda Login - Components involved



# Garuda Login - Flow

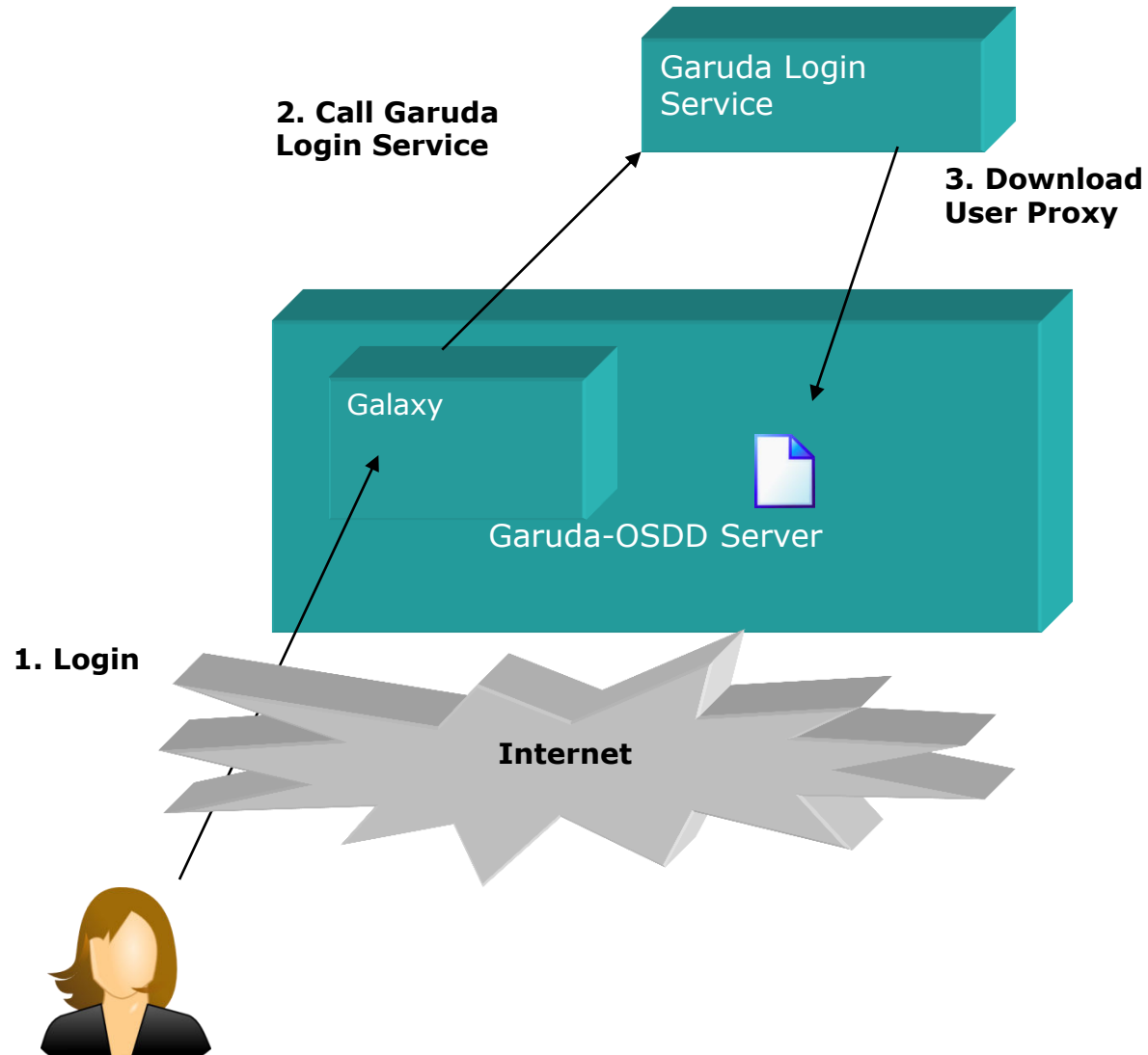




Figure 1: Login page for OSDD-GARUDA Interface

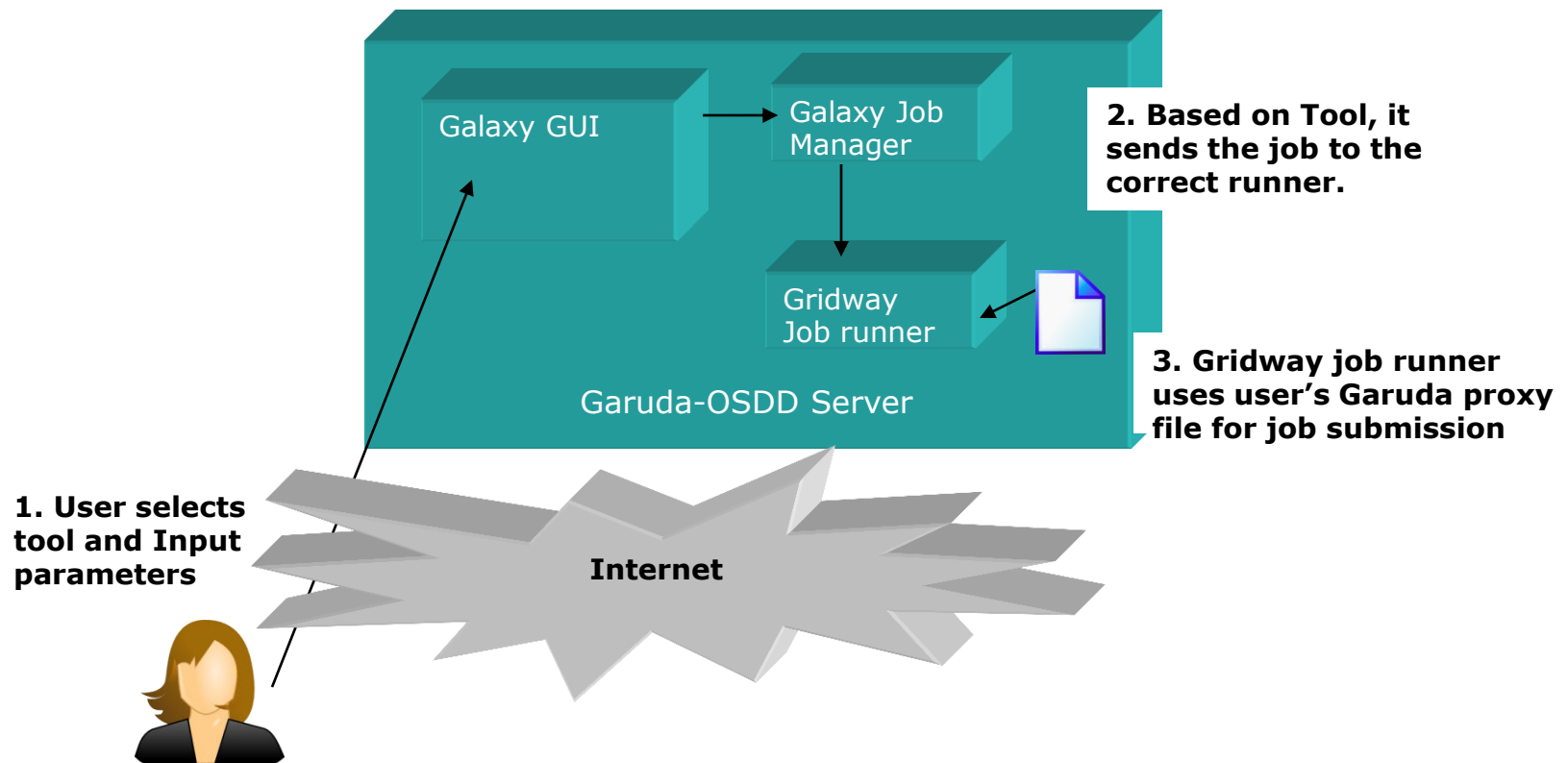
✓ You are now logged in as janaki@cdacb.ernet.in.

Your Proxy for Garuda usage is valid for 238:55:16 (9.9 days)

You will be redirected to the [home page](#) in 12 seconds.

**Page showing proxy validity of 9.9 days**

# Garuda- Galaxy Job Submission - Flow



# Weka in Galaxy

The screenshot displays the Galaxy web interface with the **wekatool** tool selected. A red arrow points to the **weka** tool in the left-hand Tools menu.

**wekatool Configuration:**

- Classifier or Converter:** classifier
- No of input files:** 2
- Train arff file:** 67: wekatool on data 66
- Test arff file:** 67: wekatool on data 66
- Classifiers:** Trees
- Trees classifiers:** J48 (selected from a dropdown menu showing options: J48, LMT, Random Forest, J48 graft, FT, BFTree, DecisionStump, LADTree, NBTree, REPTree, Random Tree, SimpleCart, User Classifier)

**History Output:**

```
value-mean(10): class
saturation-mean(12): class
hue-mean(7): class
class(7):
LogScore Bayes:
-16307.325956365592
LogScore BDeu:
-19403.577496788057
LogScore MDL:
-19117.803353376356
LogScore ENTROPY:
-16519.354065281754
LogScore AIC:
-17295.354065281757

Time taken to build model: 0.38
seconds
Time taken to test model on
training data: 0.07 seconds

=== Error on training data ===

Correctly Classified Instances
758 93.5802 %
Incorrectly Classified Instances
52 6.4198 %
Kappa statistic 0.925
Mean absolute error 0.0197
Root mean squared error 0.1225
Relative absolute error 8.0375
%
Root relative squared error
35.0245 %
Total Number of Instances 810

=== Confusion Matrix ===

a b c d e f g <-- classified as
124 0 0 0 1 0 0 | a = brickface
0 110 0 0 0 0 0 | b = sky
2 0 113 2 5 0 0 | c = foliage
2 0 0 94 10 4 0 | d = cement
1 0 18 5 102 0 0 | e = window
0 0 0 1 0 93 0 | f = path
1 0 0 0 0 0 122 | g = grass
```

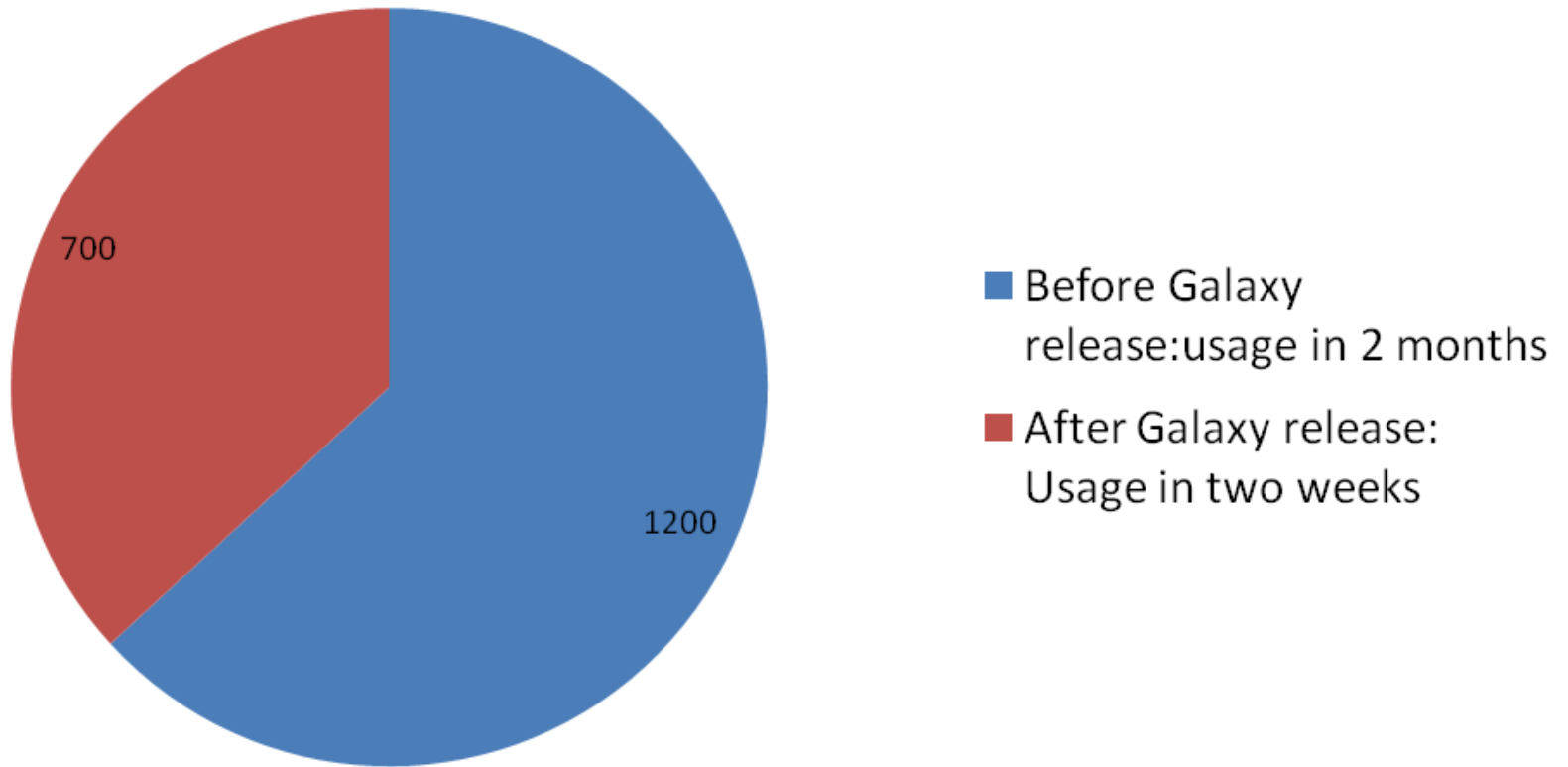
# How Grid is helping?

- Can manage any number of Concurrent user jobs
- Before Grid-enabling – Running Weka on stand-alone is time consuming, machine needs to be up for long period, only one job at a time
- After integrating in Galaxy on Garuda: User can submit any number of jobs at one time..need not worry about power problems or maintenance.



# Garuda Usage by OSDD: Job Accounting

OSDD Community Usage



# **Kudos to Galaxy Team: From OSDD Community**

- ***“Galaxy is killing beautiful, sleek and agile.. Its super superb Fabulous Beauty; Amazing Work, Kudos!!!” -***
- ***“Galaxy is very nice to use, surely it should be a stepping stone to achieve higher altitudes of the world of science.”***
- ***“The weka in Galaxy is a great and easy even simpler than weka GUI , even the cost setting is very simple and the UI is very user friendly.”***

# An Open Source Integrated Computational Resource for the Analysis of the Structural Interactome to predict Off-Site Interactions of Drug Candidates

Nagasuma Chandra, R. Sowdhamini, N. Srinivasan & Tom Blundell



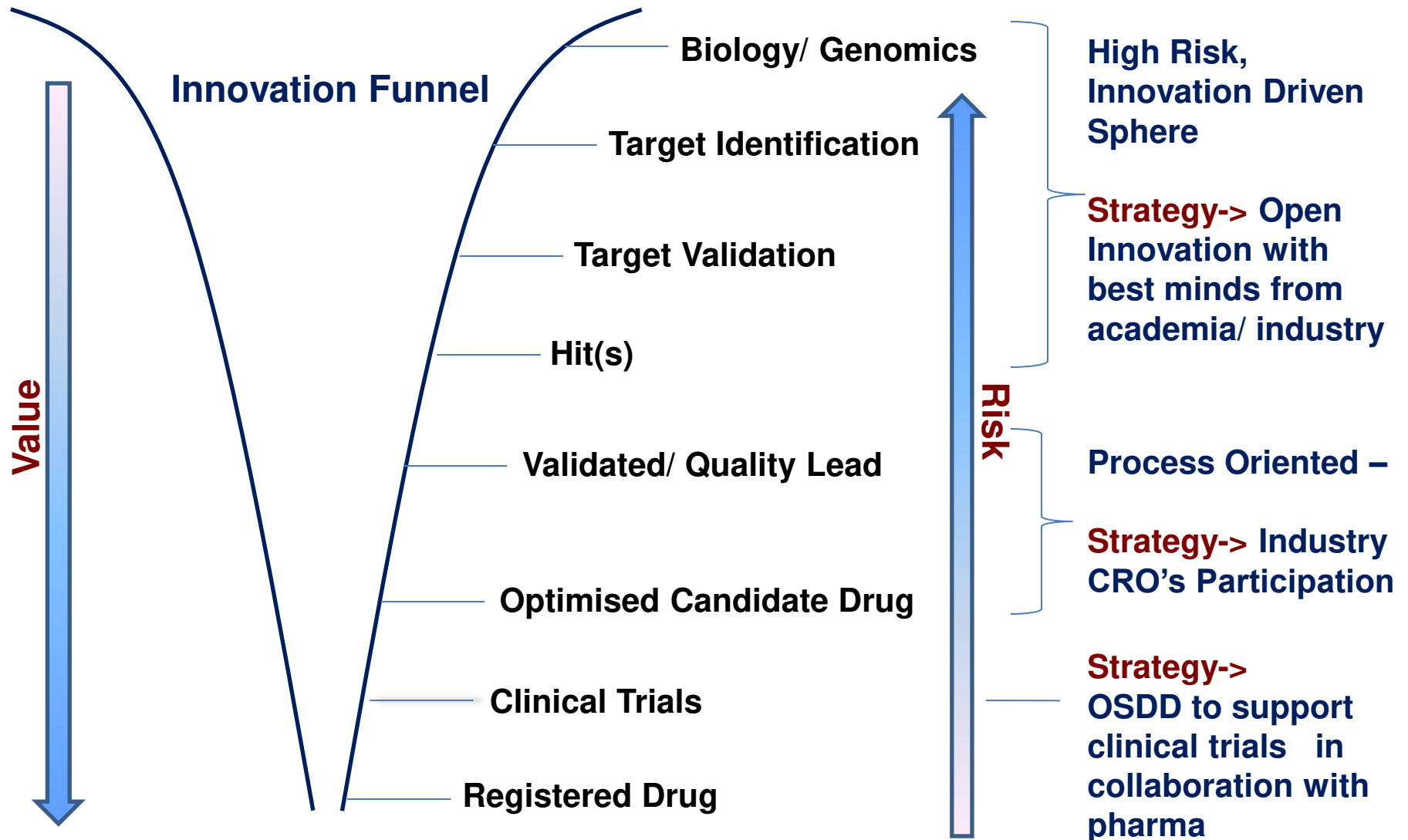
## Broad Objectives:

- To develop a resource *SInCRe CamBan* that can generate a structural interactome and thus give a structural dimension to the identification of off-target interactions of drug molecules
- To develop databases of 'pocketomes' and of '*interactophores*' that contain binding-site signatures of protein molecules from human and Mtb proteins
- To devise approaches to compare three-dimensional binding sites in different proteins and to identify off-target drug binding.
- To develop a databases that use homology to extend structural information in order to provide descriptors of protein-small-molecule, protein-protein and protein-nucleic acid in humans and in Mtb, and between host and parasite.

# Requirements in OSDD

- Real time collaborative annotation on the web with multiple collaborators sharing the same screen and feature control
- Easy integration of applications used for protein structure analysis like docking, binding site comparisons, etc which needs visualization of 3D structures
- Naming convention for I/O files needs to be flexible
- Should be able to share individual jobs

# An Innovative Approach to Drug Discovery: A New Paradigm



**License the drug on non exclusive basis to ensure access and affordability**

# Acknowledgement

**OSDD : A Global Community - 4511 members from more than 130 countries**

Albania	Costa Rica	Iran	Monaco	Sri Lanka
Algeria	Croatia	Iraq	Morocco	Suriname
Andorra	Cuba	Ireland	Mozambique	Sweden
Angola	Cyprus	Israel	Namibia	Switzerland
Argentina	Czech Republic	Italy	Nepal	Syria
Aruba	Denmark	Jamaica	Netherlands	Taiwan
Australia	Dominican Republic	Japan	New Zealand	Tajikistan
Austria	Ecuador	Kazakstan	Nicaragua	Thailand
Azerbaijan	Egypt	Kenya	Nigeria	Trinidad and Tobago
Bahrain	El Salvador	Korea	Norway	Tunisia
Bangladesh	Estonia	Kyrgyz Republic	Oman	Turkey
Belarus	Ethiopia	Laos	Pakistan	Turkmenistan
Belgium	Fiji	Latvia	Panama	Uganda
Belize	Finland	Lebanon	Paraguay	Ukraine
Benin	France	Liberia	Peru	United Arab Emirates
Bhutan	Georgia	Libya	Philippines	United Kingdom
Bosnia and Hercegovina	Germany	Liechtenstein	Poland	United States of America
Brazil	Gibraltar	Lithuania	Portugal	Uruguay
Bulgaria	Greece	Luxembourg	Romania	Uzbekistan
Burundi	Guam	Macao	Russia	Vatican City
Cambodia	Guatemala	Macedonia	Saudi Arabia	Venezuela
Cameroon	Haiti	Malawi	Senegal	Vietnam
Canada	Hong Kong	Malaysia	Singapore	Yemen
Chile	Hungary	Malta	Slovakia	Yugoslavia
China	Iceland	Mauritius	Slovenia	Zaire
Colombia	India	Mexico	South Africa	Zambia
Congo	Indonesia	Moldova	Spain	Zimbabwe



*Together we can ...*

*.. and **we should !***

<http://www.osdd.net>

<http://c2d.osdd.net>

Email: **info@osdd.net**  
**anshu@osdd.net**  
**janaki@cdacb.ernet.in**

Skype: **anshu.bhardwaj**  
**janakich**



# OSDD: Attribution and IP

- All contributions on the OSDD portal attributed to the authors with date and time stamp
- Real time data sharing
- Click wrap license agreement
  - All contributions treated as Protected Collective Information
    - mandates sharing,
    - attribution,
    - contribute back

# OSDD View on Patents

- Two patent applied molecules in hit to lead phase
- **Patent only** to ensure that:
  - Quality assurance in downstream processes
  - Subsequent innovations remain in open source
  - Affordability : through non exclusive licenses

# OSDD Open Access Resources



OPEN SOURCE DRUG DISCOVERY  
Open Access Repository

**Assembly line for drug discovery**

## **I Biological Repository**

- i. Open access clinical strains repository**
- ii. Open access clone repository**
- iii. Open access protein repository**

## **II Chemical Repository**

- i. Open access small molecule repository**

## **III Open Screening Facility**

- I. Submit your compounds for anti-tuberculosis screening**

# Community peer review

## Open Funding Review

Irish He... x Transcri... x SmallMol... x AllProje... x AntiMyc... x Results... x Results... x Potenti... x Genomic... x

borgtb.osdd.net/bin/view/OpenProjectSpace/TranscriptProfilingofMacrophagesExposedtoTBDrugsandDrugMicroparticlesDiscussions

**SysBorgTB** Systems Biology of *Mycobacterium tuberculosis*

Jump Search

SysBorg > OpenProjectSpace Web > TranscriptProfilingofMacrophagesExposedtoTBDrugsandDrugMicroparticlesDiscussions (17 Jun 2009, AmitCDRI)

[Edit](#) [Attach](#)

The current proposal submitted by Dr. Amit Misra from CDRI, Lucknow is aimed at validating the results which have been obtained before through microarray gene expression data on human monocyte like THP-1 cell lines under different conditions of experimets like treatment with Isoniazid and Rifabutin, microparticles etc. This project comes under Work Package WP5.

The project is recommended for its merits-----Charu Sharma, Scientist, IMTech, Chandigarh

[CharuSharma](#) 04 Mar 2009 - 20:4

The research proposal submitted by Dr. Amit Misra, CDRI, Lucknow is aimed at providing microarray gene expression data on human monocyte line THP-1 under different conditions of soluble drugs (isoniazid and rifabutin), microparticles along with appropriate controls. This proposed work is based on previously carried out work, in which they observed induction of pathways related to killing intracellular bacteria with a concomitant down regulation of pathways favouring bacterial survival. The results from this project will come under the Work Package WP5: Take best inhibitors and do microarray gene expression for human cells and tissues, under the specific deliverables D5.1: Systematically collected transcriptomics data serving as resource on selected inhibitors and D5.2: Transcript Foot Prints. The results will more closely lead to providing transcript footprints. Recommended -- Ramachandran

[SrinivasanRamachandran](#) 05 Mar 2009 - 16:0

Thank you for your generous comments, Ramu and Charu. BTW, Meenakshi Anurag's comment does NOT appear here! I hope it wasn't something that trashed my proposal! What now? Two positive reviews, no recommendations for change-- sounds like this proposal made the grade. Or should we wait for some deeper criticism? Like why rifabutin? Why not rifampicin? Why microparticles? Why H37Rv? Why not either drug alone? In case any of you are planning on asking these questions, I have the answers up my sleeve. Its the stuff that our group has MISSED considering that I was hoping to get from you people. Keep 'em coming... Amit

[AmitCDRI](#) 05 Mar 2009 - 20:0

# Challenges

- ✓ **Most of the biologists and chemists do not use computational workflows for their analysis**
- ✓ **Awareness about the advantages of using such workflow engines**
- ✓ **The Community needs to be trained for using the workflows**
- ✓ **The Community needs to be trained for integrating applications**
- ✓ **Web services vs standalone applications – each have their own set of advantages and limitations**
- ✓ **Developers of algorithms should be encouraged to report results in globally accepted standard formats with standard ontologies**