

The Open PHACTS Foundation

OPF is a not-for-profit membership organisation, supporting the Open PHACTS Discovery Platform:

A sustainable, open, vibrant and interoperable information infrastructure for applied life science research and development.

To reduce the barriers to drug discovery in industry, academia and for small businesses, the **Open PHACTS Discovery Platform** provides tools and services to interact with multiple integrated and publicly available data sources. To integrate this data, extensive cross-referencing of scientific concepts is needed across all databases.

The Open PHACTS Foundation ensures the sustainability of the **Open PHACTS Discovery Platform** infrastructure and acts as a hub for relevant scientific research and development.



ChEMBL



The free chemical database

**DRUGBANK**
Open Data Drug & Drug Target DatabaseWikiPATHWAYS
Pathways for the People

Key Resources

 [Open PHACTS API](#) [Open PHACTS Repository](#)

Subscribe to the Foundation Newsletter

Subscribe

Contact us

Email:
info@openphactsfoundation.org

Twitter: [@Open PHACTS](#)



Open PHACTS Mission:

Integrate Multiple Research Biomedical Data Resources Into A Single **Open & Free Access Point**

CHALLENGE 1: do we need other data to enable phenotypic screening workflows?



HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)



<http://imgs.xkcd.com/comics/standards.png>

We understand our current standards landscape
CHALLENGE 2: would phenotypic screening data bring us into 'raw data'
and instrument/format standards?



Challenge 3: How much manual data handling would be needed for us to handle new types of data for phenotypic screening?



KNIME

Table View - 0:31 - Interactive Table (7 x 6)

Name	Inchi	Activity	Units	Relation	Target
.. Sorafenib	MLDQJTXFUGDVEO-UHFFFAOYSA...	3400	nM	=	Serine/threonine-protein kinase PLK4
.. Sorafenib	MLDQJTXFUGDVEO-UHFFFAOYSA...	250	nM	=	MAP kinase signal-integrating kinase 2
.. Sorafenib	MLDQJTXFUGDVEO-UHFFFAOYSA...	5.4	uM	=	HCT-116 (Colon carcinoma cells)
.. Sorafenib	MLDQJTXFUGDVEO-UHFFFAOYSA...	1700	nM	=	Ephrin type-B receptor 1
.. Sorafenib	MLDQJTXFUGDVEO-UHFFFAOYSA...	3300	nM	=	Dual specificity mitogen-activated protein kinase kin.
.. Sorafenib	MLDQJTXFUGDVEO-UHFFFAOYSA...	6200	nM	=	Cyclin-dependent kinase 5

Workflow Diagram:

```

graph LR
    A[File Reader] --> B[Java Snippet]
    B --> C[Get Name and Inchi]
    C --> D[Get Activity]
    D --> E[Activity Parser]
    E --> F[Column Filter]
    F --> G[Interactive Table]
  
```

Node Descriptions:

- File Reader:** Simply gets the URL [I dont know how to get it to start otherwise!]
- Java Snippet:** Fetch JSON from web
- Get Name and Inchi:** Name & Inchi Grabber
- Get Activity:** Now turn the activity JSON into rows
- Activity Parser:** For each activity row, extract the columns we want
- Column Filter:** Tidy Up: Remove Processing Columns Now
- Interactive Table:** Node 31

Workflow Projects: KNIME_project

- Column Filter (#27)
- Column Rename (#30)
- File Reader (#3)
- Interactive Table (#19)
- Java Snippet (#1)
- Java Snippet (#25)
- Java Snippet (#28)
- JSONArray_2_Rows (#26)
- Split Collection Column (#29)
- OPS_Pharmacology

Node Repository

Challenge 4: do workflow components for phenotypic screening exist?
What are the gaps? How do practitioners actually do this?