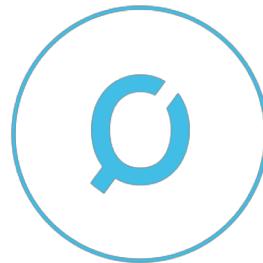


# JAX FAIR Apps Team



# DISQOVER



**The Jackson  
Laboratory**

Leading the search  
for tomorrow's cures



## FAIR Beyond Data – Applications as FAIR Team Members

# How do we become better data stewards?

# We use data standards.

National Bureau of Standards  
(now NIST) is in  
Department of Commerce

# Data standards drive commerce and solutions.

We work in a community and share  
best practices  
and use  
pragmatic solutions to FAIR

## W3C Standards

- Unique and persistent identifiers – ORCIDs, DOIs, etc.
- OWLs -
- DCATs – data catalogs so datasets can be findable
- RDFs – for describing graph networks of your data
- Turtle format (ttl) – terse RDF triple language
- SHACLs – for validation
- ... more

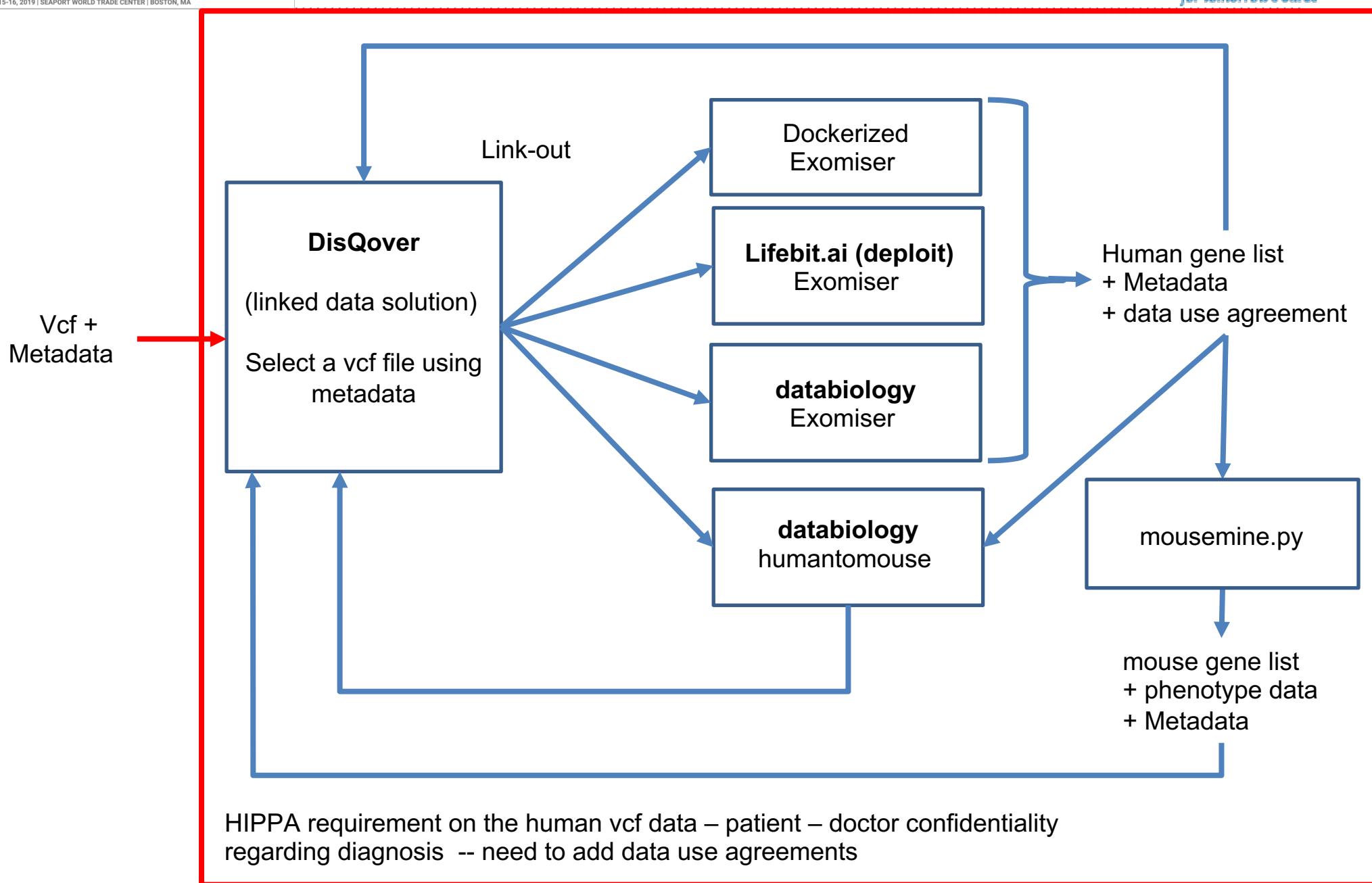
## Tools work with standards (scratching surface)

- SMART APIs – for making services known
- STARDOG – validate RDF graphs with SHACL
- openprovenance.org – used to illustrate the provenance of our data in – application used (who ran it) – data out
- Google dataset search
- ... more

# What's missing?

- Data use agreements
- Instruments generating data to these standards

# JAX FAIR Apps Use Case

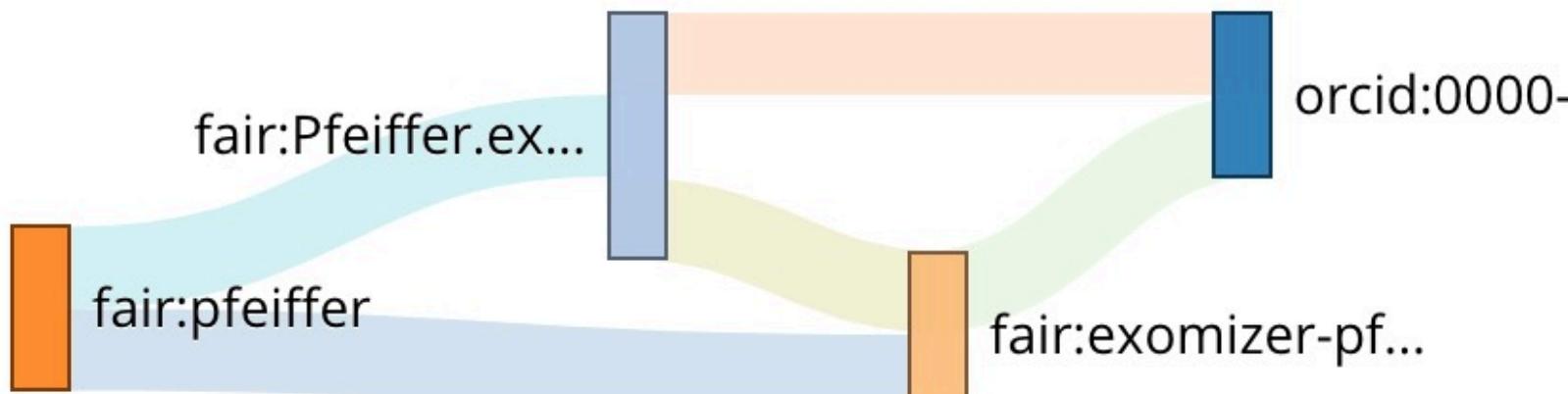


# Example Provenance Document

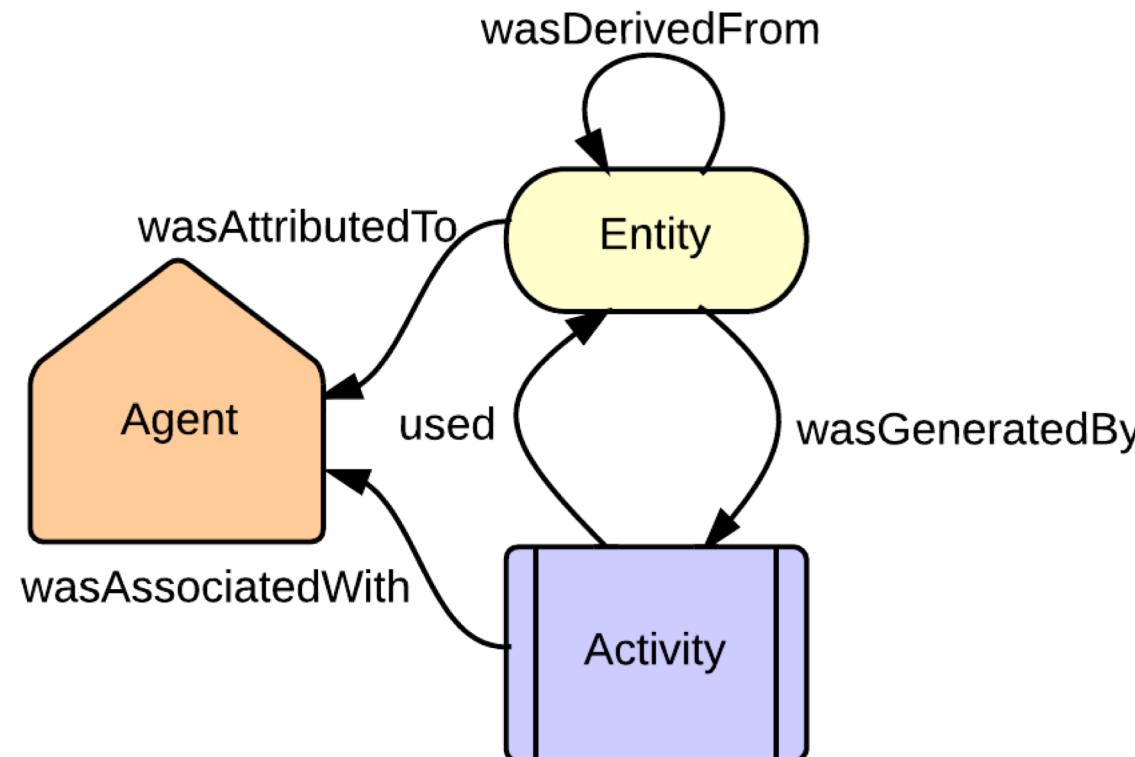
example provenance document

<https://openprovenance.org/store/documents/749>

input on left, result in middle, process is exomiser, and person who ran it is orcid



# Provenance Schema from W3C PROV



http <https://orcid.org/0000-0001-7951-3439> Accept:text/turtle --follow

DISQOVER Deploy DBE | Welcome to Databiology DISQOVER BIT Track 4- Bioinformatics

Not Secure | 54.211.41.249/#visual-analytics

Apps Disqover Ontoforce Customer System dashboard myworkandme - A... All Files | Powered... Timesheet – ONT... Login – Ontoforce... Data Science Trai...

DISCOVERIES VISUAL ANALYTICS DATA SOURCES Everything niels@ontoforce.com

Person

Start Person +

Search strategy Count Property list

1 Select all instances in Person. Filters: Sources: Hackathon People Preferred label: Nichols, Nolan

Restricted to local instances Preferred label: Clark, Alex

Affiliation

Affiliation

Count

15

Affiliation

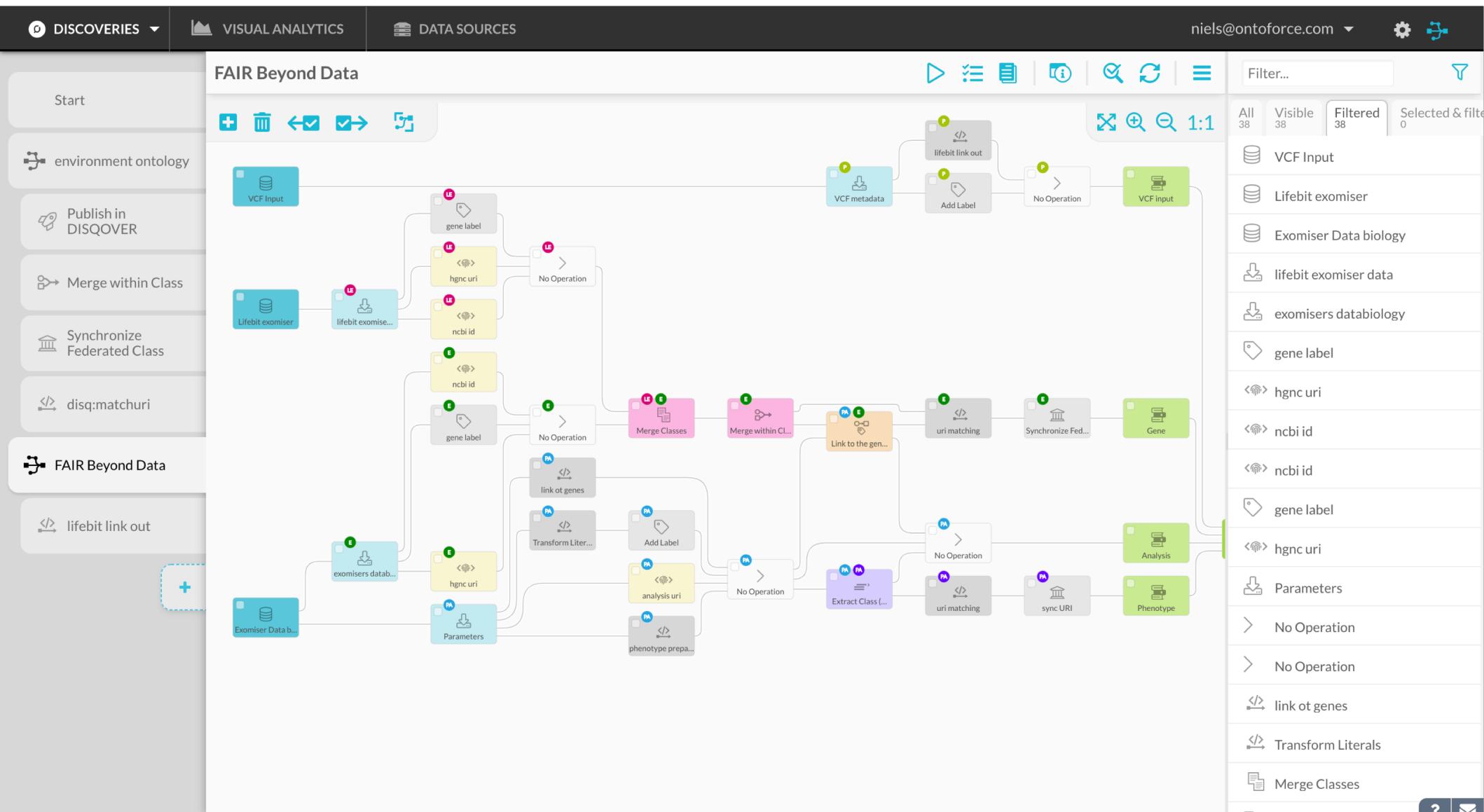
Johns Hopkins ... Boehringer Ing... Carleton Univer... Celera Corp Central Michig... Centro De Inve... Collaborative D... Cray Inc Doe Joint Gen... Genentech Inc Genetics Comp... Genstruct Georgetown U... Instituto Tecnol... Institute For Sy... Jackson Labor... Keygene Nv Knowledge Sy... Lewis And Clar... Maastricht Univ... McGill University Merck Sharp A... Molecular Mole... New Mexico St... Paradigm Gen... Pfizer Global R... San Diego Med...

Data missing: 6

Preferred label: Madduri, Ravi Preferred label: Caballero, Juan Preferred label: Cuomo, Christina A(\*) Preferred label: Slater, Ted Preferred label: 0000-0001-7951-3439 Preferred label: Thomas, Ashleigh Preferred label: Burks, Jody Preferred label: Nissenbaum, Brad Preferred label: Gnimpieba, Etienne Z(\*) Preferred label: Salomon, Hugh Preferred label: Wagner, Rick Preferred label: Dumontier, Michel Preferred label: Jeschonek, Samantha

+ ?

# DisQover



## FAIR Analysis Solution



Search &  
Find data



5

Ingest back data to  
Ontoforce

3 API Call to Deploit  
to run analysis



5

Link to any  
system/UI  
through API



Nextflow

BDS

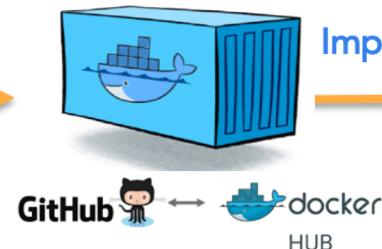
bcbio

R

scripts written  
in other  
languages

1

Containerise



2

Import to Deploit



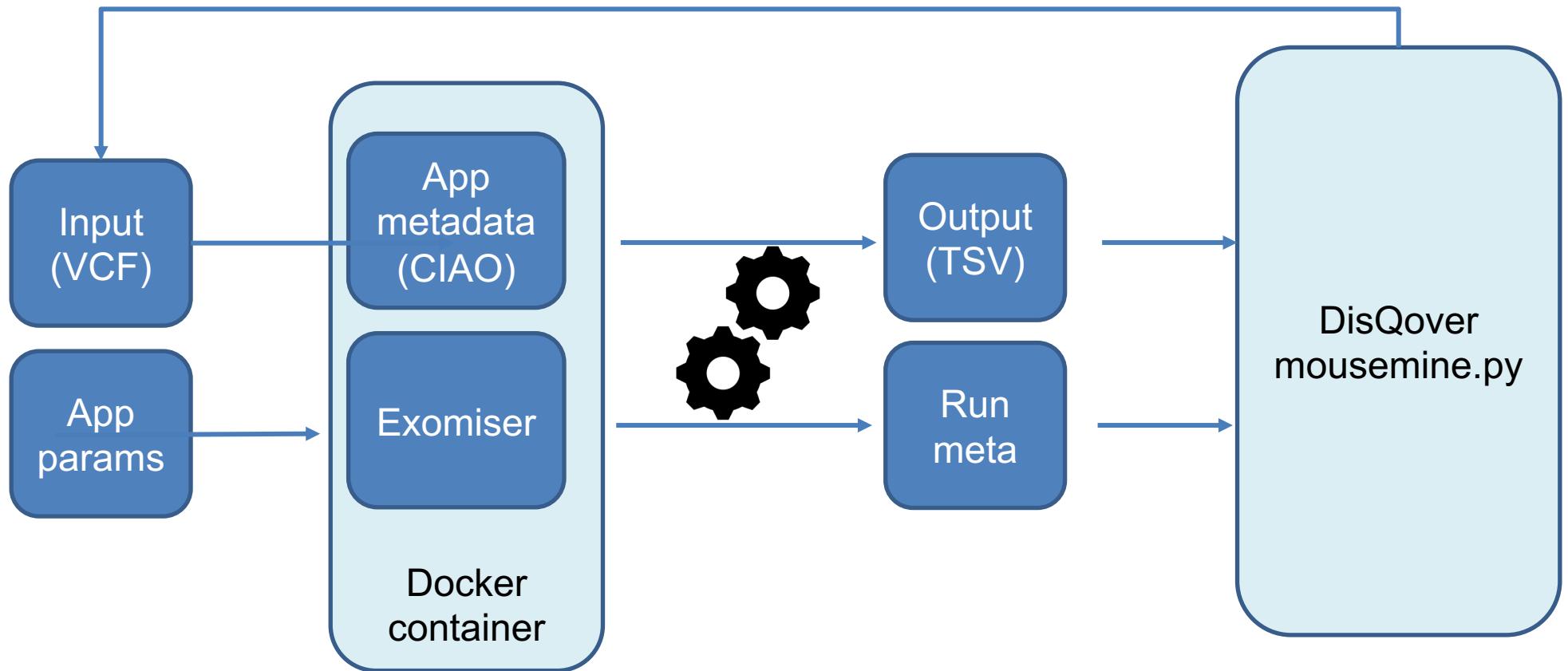
4

Monitor & Share  
analysis progress



# Databiology app

<https://bitbucket.org/databiology/app-dbio-exomiser/>



# Citations/Attributions

[Home](#)[Menu ▾](#)[About](#)[MyMPD ▾](#)[Search](#)[Examples](#)

## Citing / Publications

### Citing specific data found in MPD

From any project / data set page in MPD, you can see associated publication citation(s). Or, you can cite the MPD web page itself by clicking on the project's "About | Contact" button then find the "Formatted Citation" for copy-paste.

### Citing this resource

Please use RRID:SCR\_003212 when citing MPD.

### Publications about MPD

Bogue MA, Grubb SC, Walton DO, Philip VM, Kolishovski G, Stearns T, Dunn MH, Skelly DA, Kadakkuzha B, TeHennepe G, Kunde-Ramamoorthy G, Chesler EJ. Mouse Phenome Database: an integrative database and analysis suite for curated empirical phenotype data from laboratory mice. Nucleic Acids Res. 2018 Jan 4;46(D1):D843-D850. doi: 10.1093/nar/gkx1082. [PubMed 29136208](#)

Bogue MA, Churchill GA, Chesler EJ. Collaborative Cross and Diversity Outbred data resources in the Mouse Phenome Database. Mamm Genome. 2015 Oct;26(9-10):511-20. doi: 10.1007/s00335-015-9595-6. Epub 2015 Aug 19. [PubMed 26286858](#) [FullText](#)

Bogue MA, Peters LL, Paigen B, Korstanje R, Yuan R, Ackert-Bicknell C, Grubb SC, Churchill GA, Chesler EJ. Accessing Data Resources in the Mouse Phenome Database for Genetic Analysis of Murine Life Span and Health Span. J Gerontol A Biol Sci Med Sci. 2016 Feb;71(2):170-7. doi: 10.1093/gerona/glu223. Epub 2014 Dec 22. [PubMed 25533306](#) [FullText](#)

## Exomiser



OVERVIEW

MANUAL ▾

FAQ

ABOUT

RESOURCES ▾

# About

The Exomiser software suite

- [License and Citation](#)
- [Contact](#)
- [Publications](#)

## License and Citation

Published under the [GNU Affero General Public License \(AGPL\), version 3](#)

Please cite the Exomiser using:

**Improved exome prioritization of disease genes through cross-species phenotype comparison.** Robinson PN, Köhler S, Oelrich A, Sanger Mouse Genetics Project, Wang K, Mungall CJ, Lewis SE, Washington N, Bauer S, Seelow D, Krawitz P, Gilissen C, Haendel M and Smedley Genome research 2014;24;2;340-8

# Citations/Attributions

Not Secure | www.informatics.jax.org/mgihome/projects/aboutMouseMine.shtml

Apps DISQOVER SmartAPI Events FAIR\_apps/READ... Nextflow enables... Data lifecycle | S... ClientAxcess Platform for Scien... Other Bookmarks

MGI About Help FAQ

Home Genes Phenotypes Human Disease Expression Recombinases Function Strains / SNPs Homology Pathways Tumors

Search Download More Resources Submit Data Find Mice (IMSR) Analysis Tools Contact Us Browsers

Keywords, Symbols, or IDs Quick Search

ALLIANCE OF GENOME RESOURCES FOUNDING MEMBER

## About the MouseMine data warehouse resource

MouseMine is a data warehouse for accessing integrated mouse data from Mouse Genome Informatics (MGI) and other web resources. MouseMine is a member of an extensive community of online data resources based on the [InterMine](#) software framework. It provides multiple, diverse, query templates and supports custom and iterative querying. MouseMine gives you the ability to upload and manipulate lists of features, save and combine results from different queries, and its web services API offers robust and comprehensive programmatic access to MGI data. For more information on MouseMine data see its [Data Categories](#).

For help using MouseMine, see MouseMine [help](#) and/or [Contact](#) MGI User Support.

See also: Motenko H, Neuhauser SB, O'Keefe M, Richardson JE. 2015. MouseMine: a new data warehouse for MGI. [Mamm Genome](#). 2015 Aug;26(7):325-30. [PubMed](#)

---

Contributing Projects: Mouse Genome Database (MGD), Gene Expression Database (GXD), Mouse Tumor Biology (MTB), Gene Ontology (GO), MouseCyc

[Citing These Resources](#)  
[Funding Information](#)  
[Warranty Disclaimer & Copyright Notice](#)  
Send questions and comments to [User Support](#).

last database update  
04/09/2019  
MGI 6.13





About Help FAQ

Home Genes Phenotypes Human Disease Expression Recombinases Function Strains / SNPs Homology Pathways Tumors

Search ▾ Download ▾ More Resources ▾ Submit Data Find Mice (IMSR) Analysis Tools Contact Us Browsers

ALLIANCE OF GENOME RESOURCES FOUNDING MEMBER

## Citing MGI Resources

Publications to cite in your manuscripts

Please use the following citation formats when referring to resources available from the Mouse Genome Informatics Web Site in your publications.

**MGD:**

Smith CL, Blake JA, Kadin JA, Richardson JE, Bult CJ, the Mouse Genome Database Group. 2018. Mouse Genome Database (MGD)-2018: knowledgebase for the laboratory mouse. Nucleic Acids Res. 2018 Jan. 4;46 (D1): D836-D842.

**GXD:**

Finger JH, Smith CM, Hayamizu TF, McCright IJ, Xu J, Law M, Shaw DR, Baldarelli RM, Beal JS, Blodgett O, Campbell JW, Corbani LE, Lewis JR, Forthofer KL, Frost PJ, Giannatto SC, Hutchins LN, Miers DB, Motenko H, Stone KR, Eppig JT, Kadin JA, Richardson JE, Ringwald M. 2017. The mouse Gene Expression Database (GXD): 2017 update. Nucleic Acids Res. 2017 Jan. 4;45 (D1): D730-D736.

**MTB:**

Bult CJ, Krupke DM, Begley DA, Richardson JE, Neuhauser SB, Sundberg JP, Eppig JT. Mouse Tumor Biology (MTB): a database of mouse models for human cancer. Nucleic Acids Res. 2015 Jan 28;43(Database issue):D818-24.

See [Mouse Genome Informatics Publications](#) for an expanded list of resources available at the Mouse Genome Informatics website.

To cite specific database projects, use a format similar to that shown in the following examples:

Mouse Genome Database (MGD) at the [Mouse Genome Informatics](#) website, The Jackson Laboratory, Bar Harbor, Maine. World Wide Web (URL: <http://www.informatics.jax.org>). [Type in date (month, yr) when you retrieve data cited].

[Gene Expression Database \(GXD\)](#), Mouse Genome Informatics Web Site. World Wide Web (URL: <http://www.informatics.jax.org>). [Type in date (month, yr) when you retrieve data cited.]

[Mouse Tumor Biology Database \(MTB\)](#), Mouse Genome Informatics Web Site, The Jackson Laboratory, Bar Harbor, Maine. World Wide Web (URL: <http://tumor.informatics.jax.org/mtbwi/index.do>). [Type in date (month, yr) when you retrieved data cited].

[MouseMine](#), Mouse Genome Informatics Web Site, The Jackson Laboratory, Bar Harbor, Maine. World Wide Web (URL: <http://www.mousemine.org/>). [Type in date (month, yr) when you retrieved data cited].

To cite a specific data area or display, use a format similar to that shown in the following examples:

Some phenotype data for this paper were retrieved from the Mouse Genome Database (MGD), Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, Maine. World Wide Web (URL: <http://www.informatics.jax.org>). (June, 2013 [i.e., the date you retrieved the data cited]).

Davission MT, Cook SA, Eicher EM. The first spontaneous mutation in the mouse Herc2 gene, MGI Direct Data Submission to Mouse Genome Database (MGD), MGI:1349786, (URL: <http://www.informatics.jax.org>). (1999).

# Citations/Attributions

https://www.ontoforce.com

Apps DISQOVER SmartAPI Events FAIR\_apps/READ... Nextflow enables... Data lifecycle | S... ClientAxcess Platform for Scien... Other Bo

LOGIN TRY IT FOR FREE

ONTOFORCE PRODUCT TECHNOLOGY CUSTOMERS PRICING COMPANY BLOG CONTACT

We're hiring! [APPLY NOW](#)

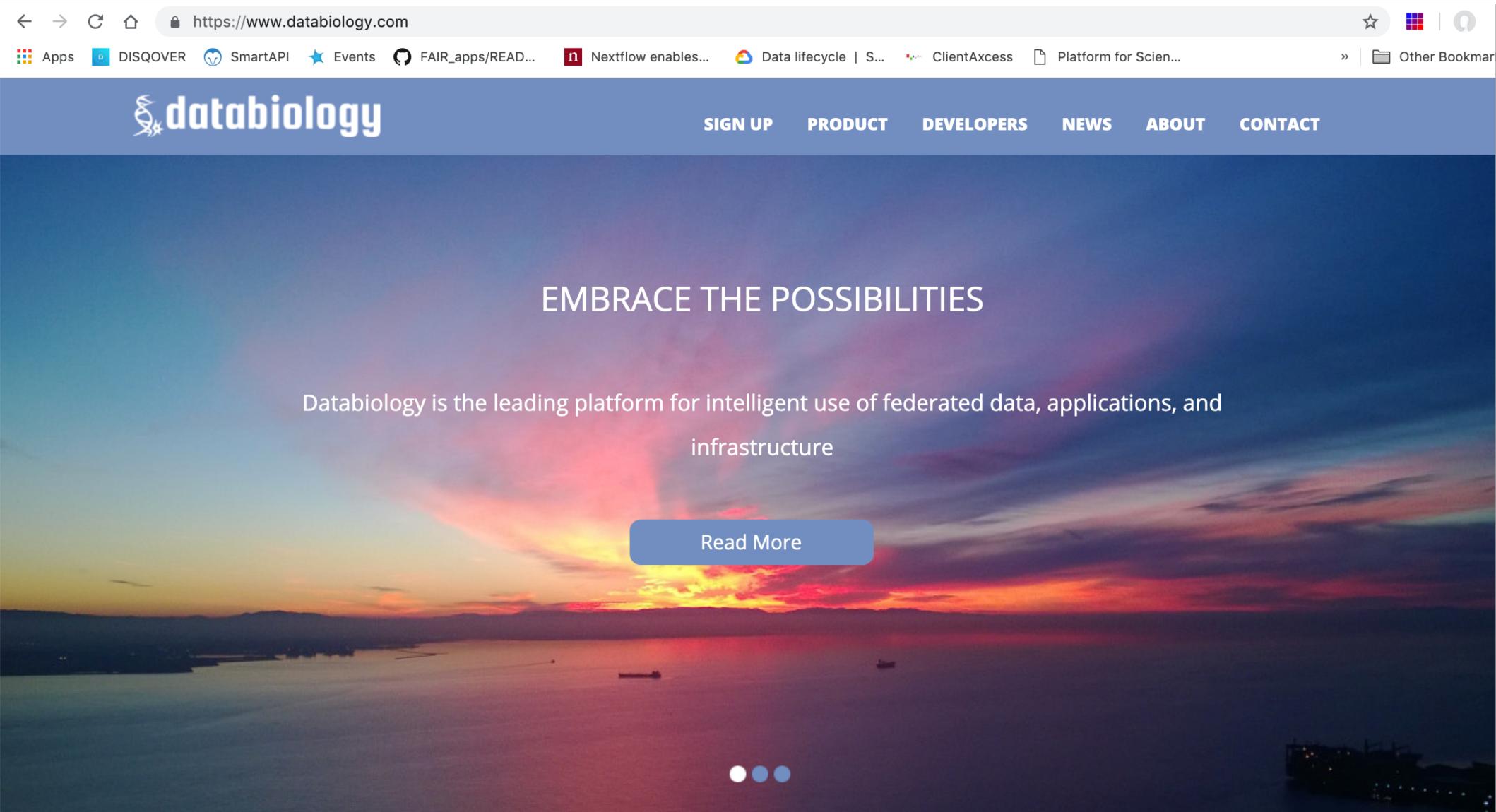
Find data faster, smarter and simpler with DISQOVER  
Our semantic linked data and visualization platform DISQOVER powers citizen data science.

4/17/19

# Citations/Attributions

The screenshot shows the homepage of Lifebit.ai. At the top, there is a navigation bar with links to Apps, DISQOVER, SmartAPI, Events, FAIR\_apps/READ..., Nextflow enables..., Data lifecycle | S..., ClientAxcess, and Platform for Scien... A "Try for free" button is located on the left side. In the center, there is a large call-to-action text: "Automate your multi-omics & big data HPC/Cloud deployment. Leverage AI for breakthrough insights generation." To the right, a laptop screen displays a user interface for managing recent jobs and projects, showing metrics like 56 completed jobs and 564 computed hours. The overall design features a light blue color scheme and abstract circular patterns.

# Citations/Attributions



https://www.databiology.com

Apps DISCOVER SmartAPI Events FAIR\_apps/READ... Nextflow enables... Data lifecycle | S... ClientAxcess Platform for Scien...

SIGN UP PRODUCT DEVELOPERS NEWS ABOUT CONTACT

## EMBRACE THE POSSIBILITIES

Databiology is the leading platform for intelligent use of federated data, applications, and infrastructure

Read More

• • •