CDD WEBINAR

Recorded Live

FAIR (meta)-Data and Assay Annotation

December 8, 2023

9:00 AM (PDT), Noon (EDT), 4:00 PM (GM

FREE - REGISTER NOW!

CDD VAULT®

Complexity Simplified

Featuring these scientists and moderated by Dr. Barry Bunin (Founder & CEO of CDD):

Isabella Feierberg, PhD

Associate Director Cheminformatics
Jnana Therapeutics

As a passionate problem solver & collaborative leader, Isabella is 100% geared towards ideation, strategy & getting things done. She is motivated by the opportunity to make a difference & has compiled great experience in navigating the interface between people, science and data. Having received her PhD at Uppsala University, Isabella has worked in computational chemistry handling (amoung other things) small molecule library design, HTS analysis & data modeling. During her 17-year tenure at AstraZeneca, Isabella became a champion for open innovation and enabling FAIR data.

Ellen Berg, PhD

Scientific Advisor and Consultant
Alto Predict

Ellen specializes in human cell-based translational assays and platforms for drug discovery and chemical safety applications. As such, she works with research groups, technology vendors and startups to advance the development and application of these technologies for better, safer products. With a background in biopharmaceutical research and drug discovery contract research services, Ellen uses her entrepreneurial, operations and commercialization experience to promote technology innovations and data standards within the realm of research data management.

Veronique FRANCOIS-NEWTON, PhD

Project Manager of the In Vitro Pharmacology Working Group - Pistoia Alliance

Veronique has more than 10 years of experience in in vitro research and used to be a senior leader in a CRO providing services to the biopharmaceutical industry. She has directed the clinical and preclinical operations and has participated to the strategic development of the CRO in particular in the areas of innovation and research. Veronique holds an MSc in Immunology and a PhD in Biology from the University of Paris and Institut Pasteur, Paris.

Larry Callahan, PhD

Global Substance Registration System (GSRS)
Office of Health Informatics | Office of Chief Scientist
FDA|DHHS

Larry is a chemist at the FDA responsible for the development of ISO data standards and the Global Substance Registration System (GSRS). He collaborates with the NIH and other leading researchers in industry and academia on standards and ontologies for regulatory submissions.

Have a question to ask our panel?
Open the ZOOM Q&A and type in your question at anytime !

Saving your Questions to the end

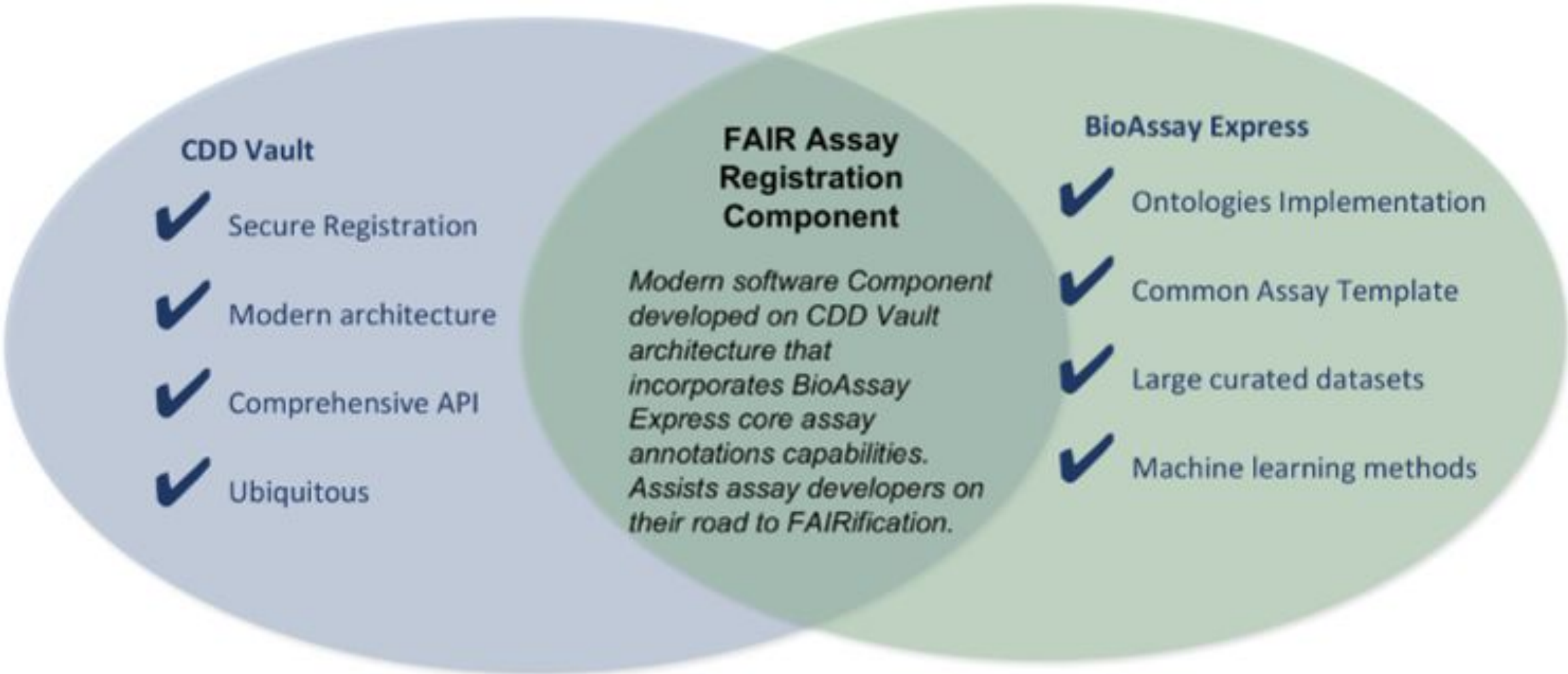
FAIR Assay Registration Application

Live Demo

Kellan Gregory
Director of Product Excellence
Collaborative Drug Discovery

CDD VAULT®
COLLABORATIVE DRUG DISCOVERY

The Genesis of the Bioassay Registration Component



CDD VAULT®
COLLABORATIVE DRUG DISCOVERY

Assay Forms: Configurable, Interactive, Interoperable

API

CDD VAULT®
Complexity Simplified

Protocol Forms

- Protocol Fields
- Ontology terms
- Run Fields

GUI

Interactive Forms include:

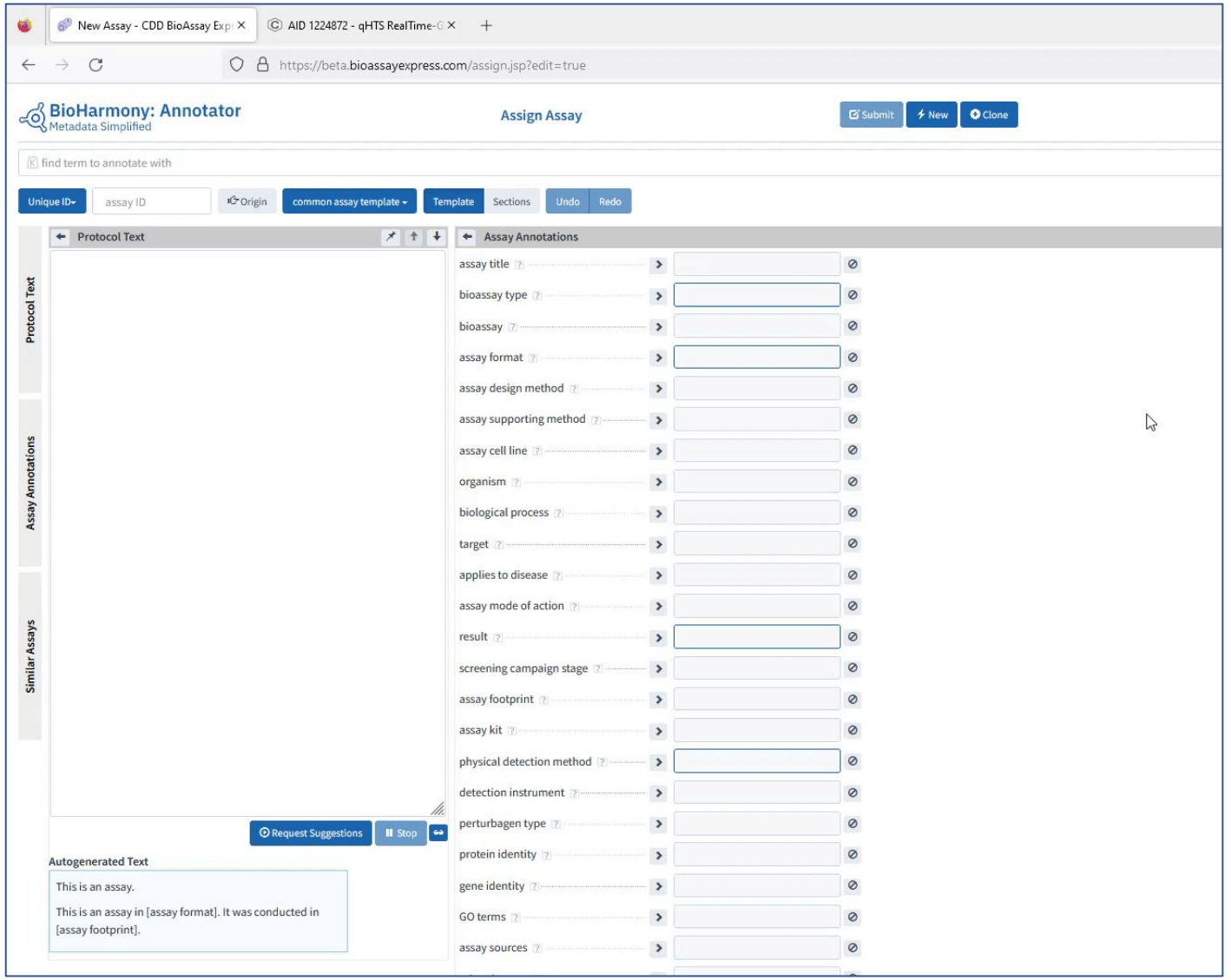
- Assay Fields
  - Vault Admin defined
  - Custom to your organization
- Ontology-based Fields
  - Public Ontologies, including BAO, CELL, GO, Drug Target
  - Fields from Common Assay Template as defined in
- Run Fields
  - Experiment specific
  - May change over time
  - Avoid data duplication across all rows of data (readouts)

NLP Technology + Expert Review = “AI in the Loop”

- ✔ FAIR
- ✔ High Quality
- ✔ Efficient
- ✔ Expandable
- ✔ Exportable

assay.biometadata.com

+ Ontologies
+ Standard Templates
+ NLP Technology
  + machine learning
+ Expert Review



Procedure:

1. Select template
2. Paste text
3. NLP suggestions
4. Review suggestions
5. Save Annotations

Built by Collaborative Drug Discovery

CDD VAULT® Complexity Simplified

CDD VAULT®
COLLABORATIVE DRUG DISCOVERY

FAIR Biological Assay Registration: Why?

Problem

Bioassay Protocols: Unstructured, non-standard language, difficult to digest, compare, or replicate.

Most Databases: Free-text, non-normalised fields → hard to find and understand what has been done previously which leads to data duplication & redundant screens.

Solution

Normalize Text: Align with industry-standard public ontologies → Enhanced data findability & reusability.

Assay Registration System: Essential for FAIR (Findable, Accessible, Interoperable, Reusable) data.

Pistoia Alliance

Home | Join the Alliance | Join a Project | Join a Community | Submit an Idea | Events | News

PRESS RELEASE

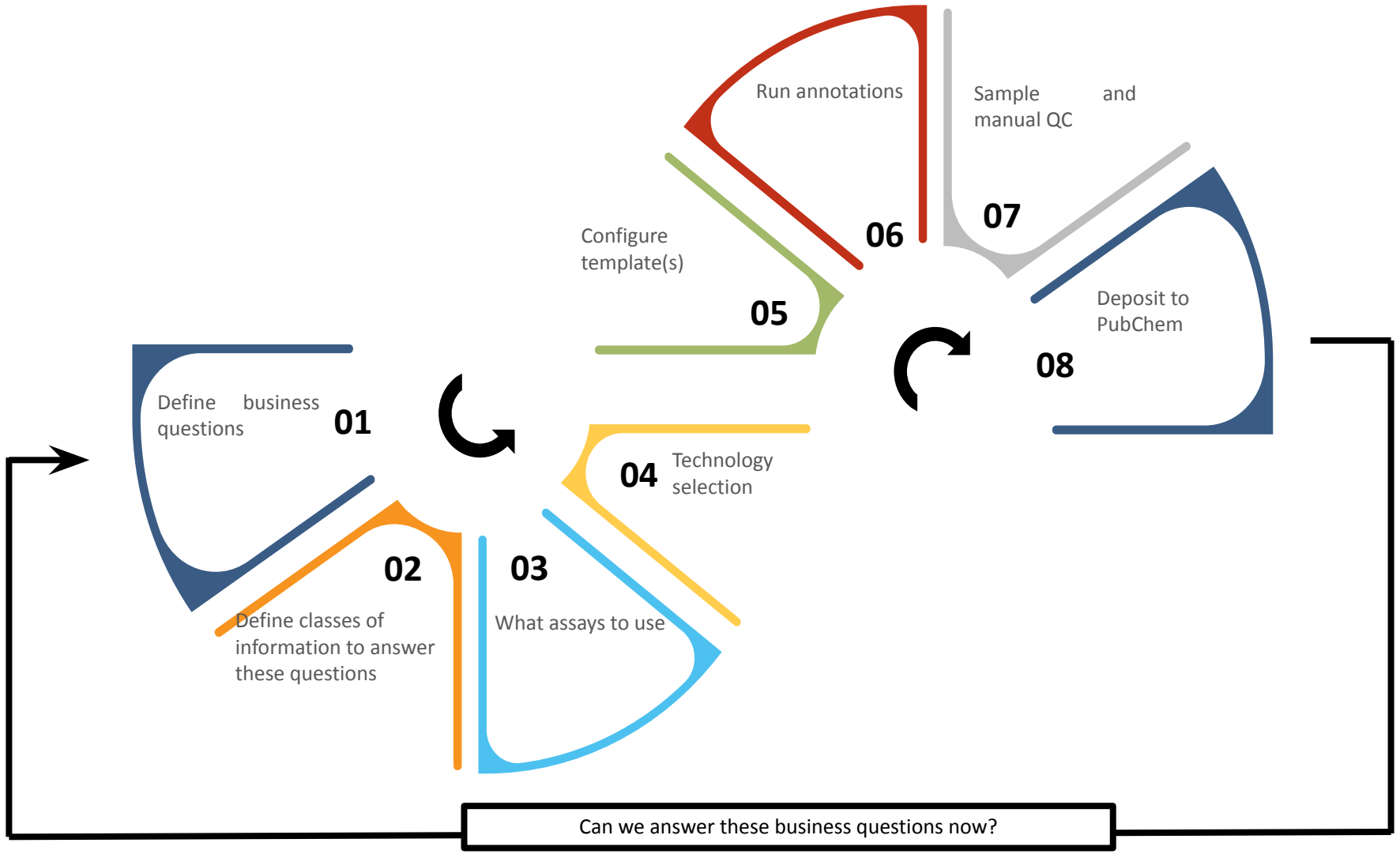
Pistoia Alliance and FDA Collaborate to Establish *In Vitro* Pharmacology Working Group to Accelerate Regulatory Submissions

FDA will participate in Pistoia Alliance's working group as it develops a standardized and shared data template for pre-clinical pharmacology data submissions

CDD

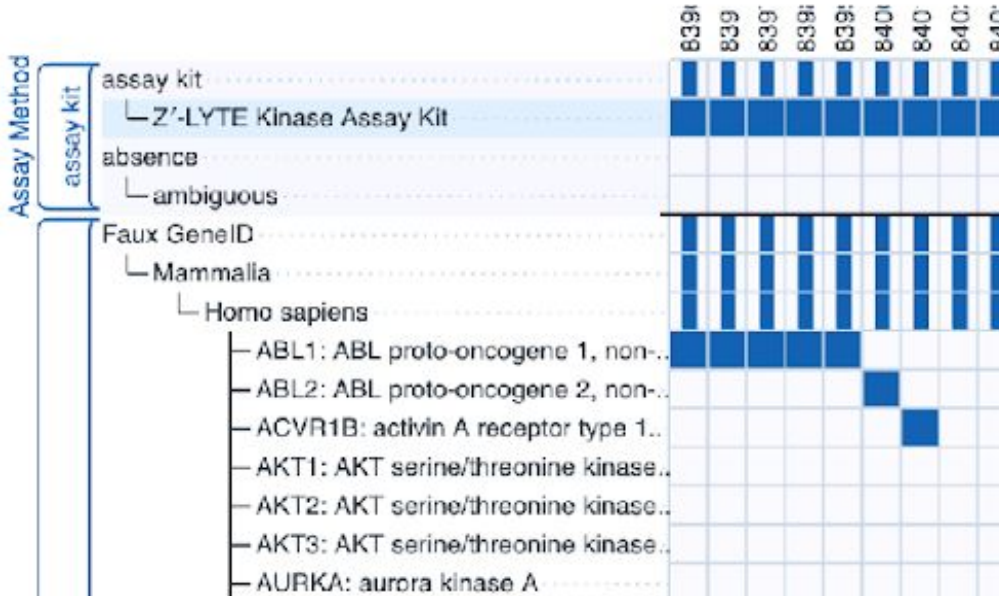
Pistoia Alliance “DataFAIRy” Biological Assay Metadata FAIR Annotation Project

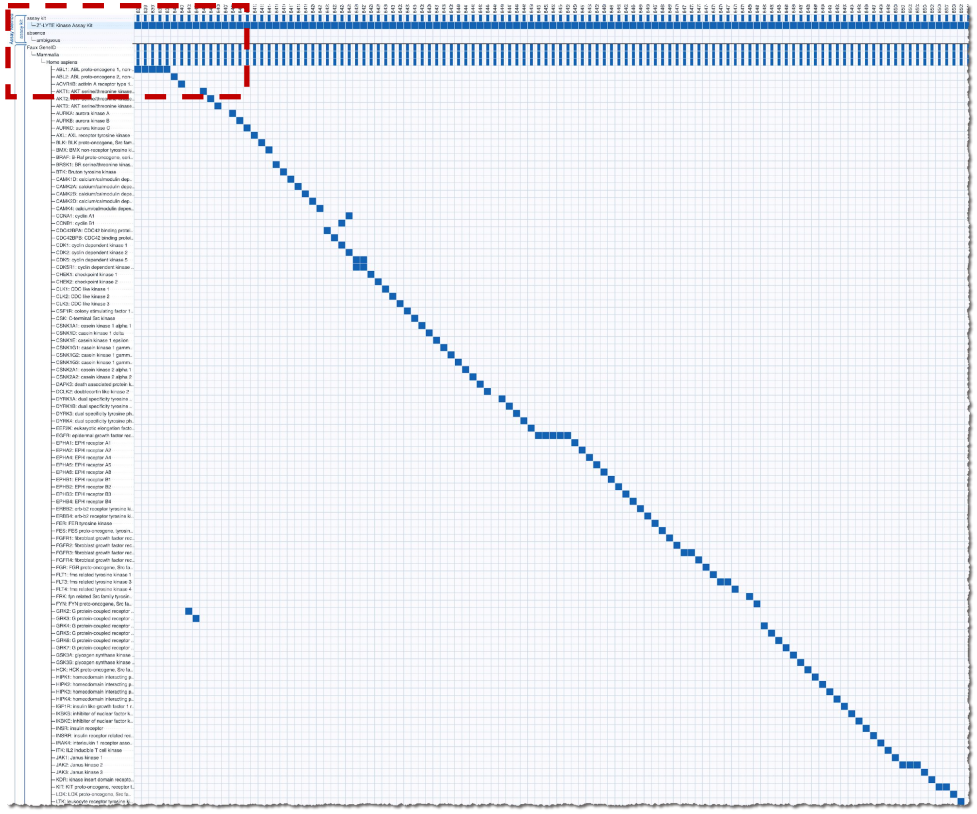
Closing the Loop



Closing the Loop

Q: *How many and what targets has this assay kit been used for?*
A: *All of these, see list*

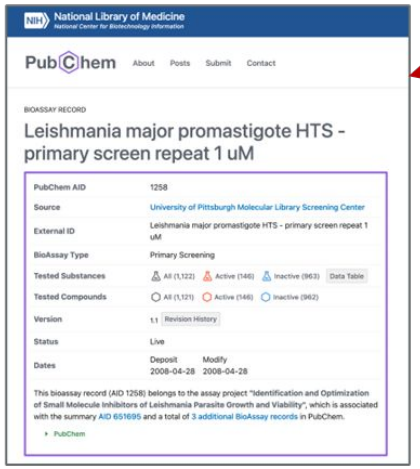


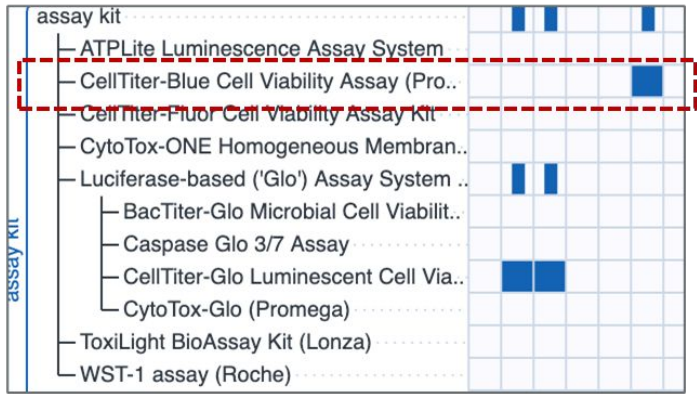


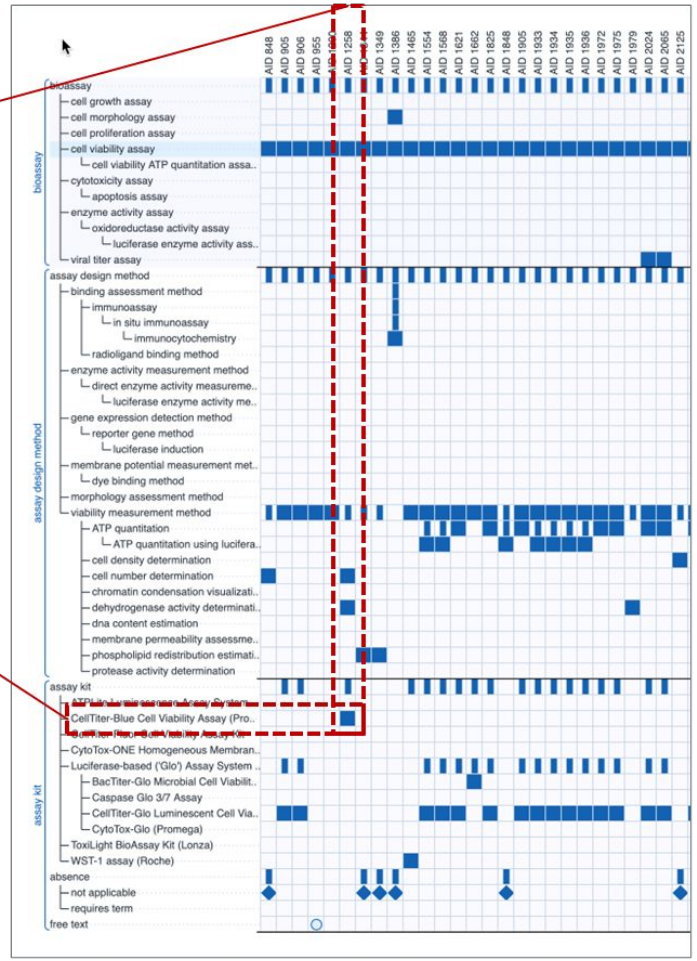
Closing the Loop

Q: *Can I find tool compounds to develop a cell viability assay that doesn't use Luciferase?*

A: *A cell viability assay that doesn't use luciferase, from this paper with compounds in PubChem*





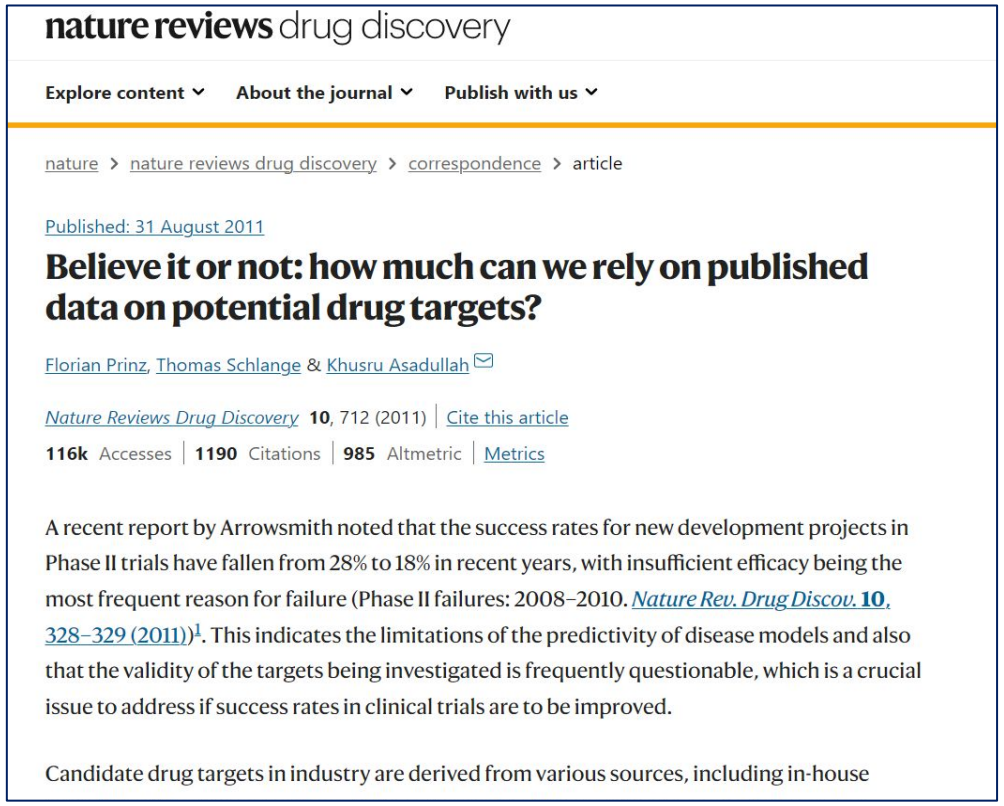


Poll question: Who Benefits from FAIR Assay Annotation?

1. Folks Running Assays (data generators)
2. Folks Consuming Assay Data
3. Project Managers
4. All of the above

Better Assay Metadata Higher Research Quality

Main Point



nature reviews drug discovery

Explore content ▾ About the journal ▾ Publish with us ▾

nature › nature reviews drug discovery › correspondence › article

Published: 31 August 2011

Believe it or not: how much can we rely on published data on potential drug targets?

Florian Prinz, Thomas Schlange & Khusru Asadullah

Nature Reviews Drug Discovery **10**, 712 (2011) | Cite this article

116k Accesses | **1190** Citations | **985** Altmetric | Metrics

A recent report by Arrowsmith noted that the success rates for new development projects in Phase II trials have fallen from 28% to 18% in recent years, with insufficient efficacy being the most frequent reason for failure (Phase II failures: 2008–2010. *Nature Rev. Drug Discov.* **10**, 328–329 (2011))[1]. This indicates the limitations of the predictivity of disease models and also that the validity of the targets being investigated is frequently questionable, which is a crucial issue to address if success rates in clinical trials are to be improved.

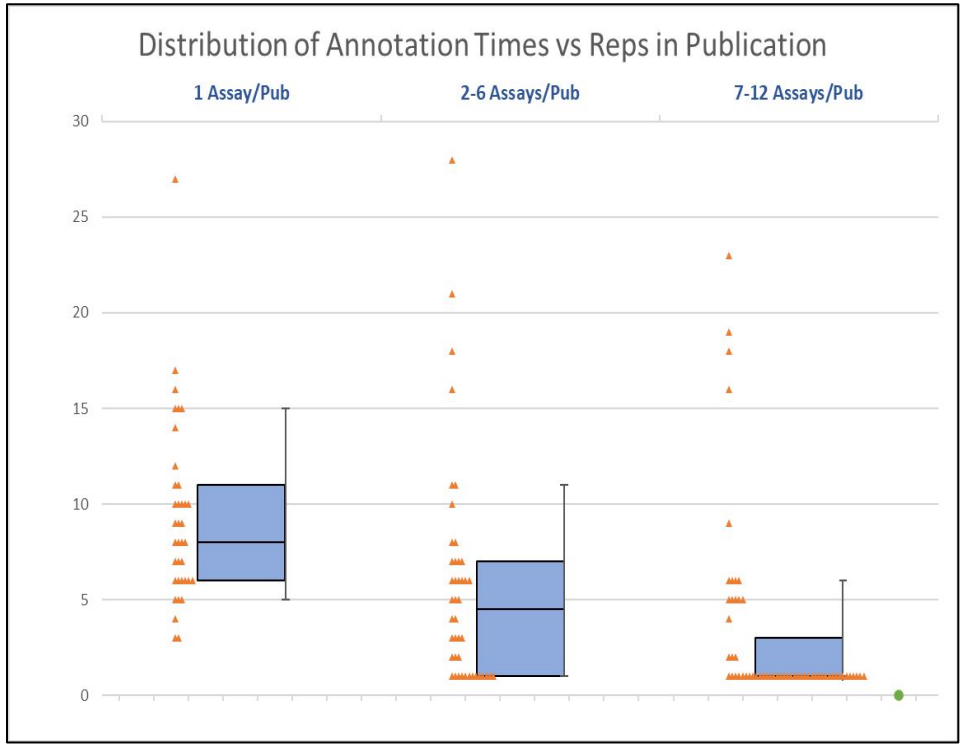
Candidate drug targets in industry are derived from various sources, including in-house

"**A quarter** of Bayer's in-house findings replicated the original results. But the analysis of Bayer's results found that the results that did replicate could often be successfully used for clinical applications."

Nature Reviews Drug Discovery, 2011, https://www.nature.com/articles/nrd3439-c1

Main Lessons Learned in the POC Phase

- Not all assay descriptions are equal Attention publishers!
  - Vendor-supplied instructions are best – way better than peer-reviewed papers
  - Publications with many assay protocols are better quality than those with one or few
- NLP + Expert in the Loop = an extremely efficient process
  - Mean time/assay: 5.8 minutes, but on average, need 2 rounds of QC
  - Final rate of critical errors 2.5%



Why Stop There?

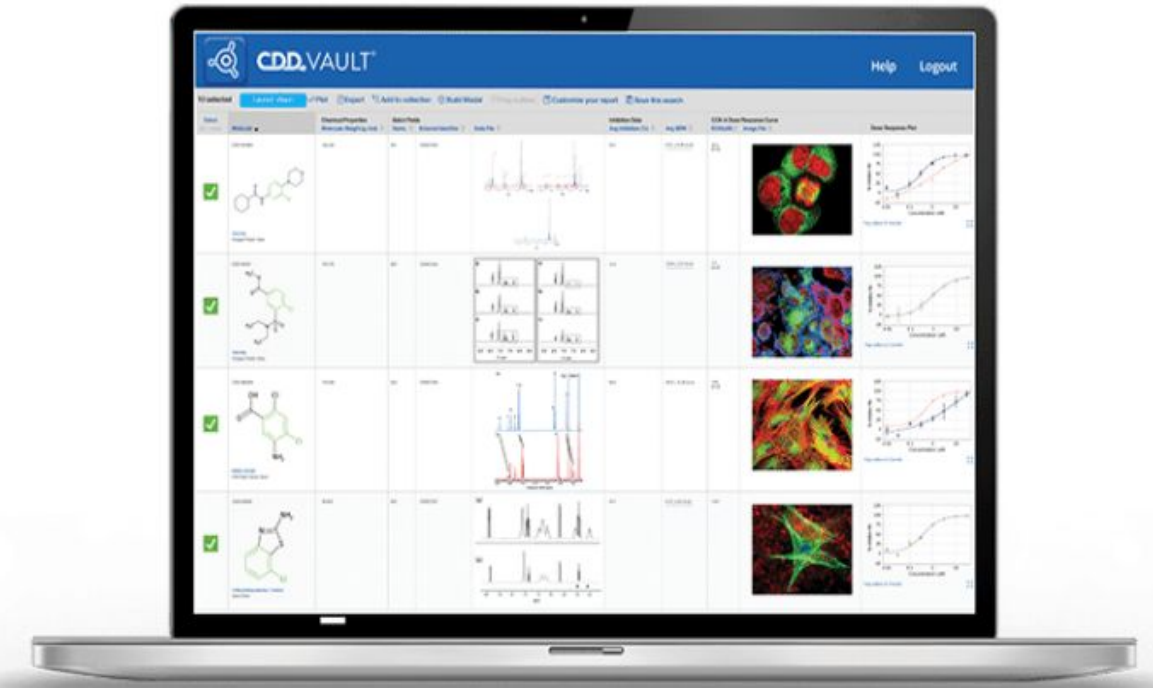
- FAIR standards for scientific data across the board
- Ideally, all information published in a paper should be accompanied by the ready-to-use publicly accessible digital content
  - Raw data – re-use for meta-analysis and comparison
  - Methods – load into your lab robot and reproduce the research
  - Computational Models – load and reproduce and/or find mistakes
  - Results – ready for import into KGs, systems biology models, etc

Open-Ended Questions

- What are the costs and benefits to organizations of using these meta-data technologies?
- Share the Data FAIRy in-depth business analysis, POC results, and lessons learning from community annotation?
- Describe the PA & FDA Collaboration on IVP Standards & Shared data templates for pre-clinical submissions?
- Where was this field in the past (say a decade ago)…& where do you see it going in the future?

Activity & Registration
Store and organize your data

Inventory
Keep Track of Compounds



ELN
Document all your research

Visualization
Plot datasets and mine them

Learn More:
https://www.collaborativedrug.com

Personalized Demo & Free Trial:
info@collaborativedrug.com