



Pistoia Alliance Portfolio

Christian Baber & Thierry Escudier
July 2025



Our Strategic Priorities



Delivering Data-Driven Value at Scale

Data is the foundation for AI and digital strategies, but we struggle to manage, maintain and consume data



Harnessing AI to Expedite R&D

83% of Life Sciences Execs say they will not achieve growth goals without AI*

* Accenture Report



Accelerate Late-Stage R&D

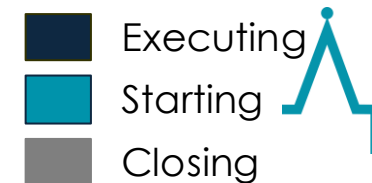
Increase the effectiveness of the most expensive phases of R&D to benefit patients



Sustainability Driven R&D

ESG is a priority for companies from a growth and public health perspective. We must ensure that the benefits of our therapies are not offset by our operations

Activities Aligned with Strategic Priorities



Delivering Data-Driven Value at Scale

Ontologies

- IDMP-O
- Pharma CMC Process Ontology
- Pharma General Ontology

Data Enrichment

- DataFairy Bioassay Annotation
- FAIR-for-Pharma Community

Integration and Management of Data, Ontologies and Standards

- Methods HUB
- In-vitro Pharmacology
- Data Governance Community
- GSRS Consortium

Harnessing AI to Accelerate R&D

Algorithms

- Large Language Models in Life Sciences

Best Practices

- LLM NLP Use Case Database
- AI/ML Expert Community

Sustainability Driven R&D

- Carbon Footprint for Decentralized Clinical Trials
- In Vitro NAMs Data Standards
- Digital Transformation & Change Management Community

Accelerate Late-Stage R&D

- Social Media & RWE
- PS² - Case Intake

Other Projects And Communities

- Microbiome-Mediated Drug Metabolism Database
- Controlled Substance Compliance & Shipping Community (CSCS)
- User Experience in the Life Sciences Community (UXLS)
- Future Lab Evolution Community
- Quantum Computing Community

Ideas Aligned with Strategic Priorities



Delivering Data-Driven Value at Scale

Ontologies

- Instrument & Equipment Ontology
- Experiment & Assay Ontology (BAO)

Data Enrichment

- Open-Source Allotropy
- Synthetic Data in Medical Imaging

Integration and Management of Data, Ontologies and Standards

- Chemical Exchange File Format Community
- Standardized bioinformatics (NGS) computational pipelines
- Standard API for Semantic Data
- HELM Format Enhancements

Harnessing AI to Accelerate R&D

Algorithms

- Real World Benchmarking of AI Solutions
- AI Agent Specification and Communication Protocol
- Benchmarking Natural Language Data Mining with LLMs

Best Practices

- Utilization of AI/ML results in Regulatory Filings

Sustainability Driven R&D

- Minimal *In Vivo* Metadata for Data Reuse & Repurposing

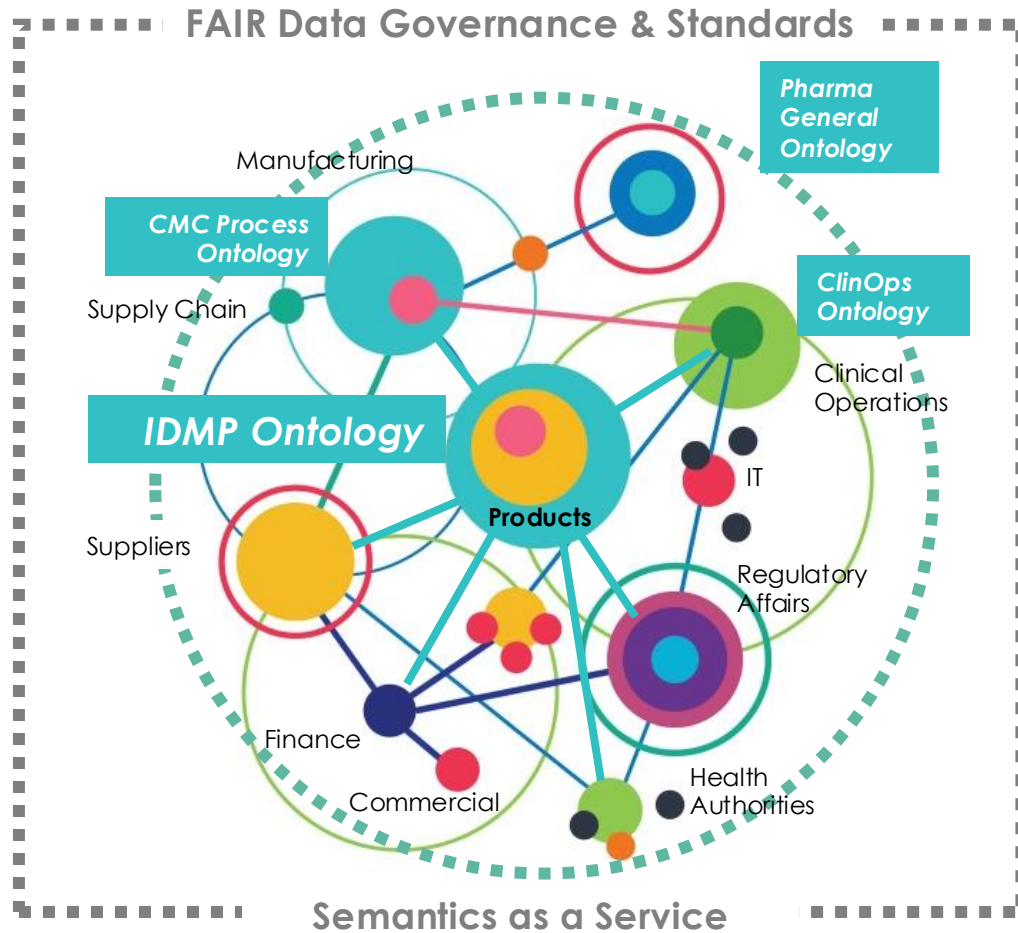
Accelerate Late-Stage R&D

- Safety & PV AI Forum

Other Projects And Communities

- Regional Restrictions on Shipping & Data Exchange

Pistoia Alliance Semantic Network



A sustainable portfolio of actionable life sciences' ontologies & semantic resources coordinated by an in-house Ontologist to ensure interoperability.

IDMP Ontology acts as the Backbone, Pharma General Ontology provides a Rosetta Stone and FAIR is the guiding principle for the industry

Coming Soon: Pistoia Alliance Ontology Design Guidelines

Training



Completed Training

Now available on demand

General Ontology Training

Webinars

Introduction and Business Value of Ontologies

From Taxonomies to Ontologies and Knowledge Graphs

Usage of Ontologies: Use Cases from the Pharmaceutical Industry

Ontologies for Pharma and Life Sciences Leaders

Increasing Adoption to Realize Value

Speakers

Christine Memmott, Vertex

Rajaram Kaliyaperumal, J&J

Martin Romacker, Roche

Birgit Meldal, Pfizer

Sheila Elz, Bayer

Joshua Valdez, Novo Nordisk

Jane Lomax, Scibite

Andrea Splendiani, IQVIA

Berenice Wulbrecht, Ontoforce

65 Graduates

IDMP-O Training

Webinars

Introduction to IDMP (in collaboration with CTADHL)

IDMP-O overview and benefits

Getting started with IDMP-O

Use Cases

IDMP-O Dependencies and Implementation Guide

Speakers

Frits Stulp, CTADHL

Jean-Gonzague Fontaine, GSK

Sheila Elz, Bayer

Joerg Stueben, BI

Ciby Abraham, AZ

Rajaram Kaliyaperumal, J&J

Norman Schmuff, FDA

26 Graduates



AI/ML in Pharma Training

Ongoing Q2 2025

Goal: Provide participants with a comprehensive understanding of how advanced AI and ML technologies are transforming the pharma landscape.

Topics

Introduction to AI/ ML and how it is used in pharma
AI for drug discovery
Generative AI for Drug Development
AI in computational chemistry
Enhancing Drug Development through Multimodal Deep Learning
LLMs for information extraction
Computer vision
Integrating Knowledge Graphs for Enhanced Data Understanding
Digital pharmaceutical manufacturing

Speakers from:

- Ipsen,
 - Parabilis Medicines,
 - AbbVie,
 - Recursion,
 - Merck & Co
 - Novo Nordisk,
- and other industry professionals

Format:

- Nine weekly live 1-hour online sessions
- Followed by on-demand access

Target Audience:

- | | |
|---|--|
| • Pharmaceutical Researchers and Scientists | • Pharmaceutical R&D Managers |
| • Data Scientists and Analysts in Pharma | • Product Development & Innovation Leaders |

Each session includes both an Industry Speaker and an Academic one from Queen Mary University or the University of Dundee

FAIR Data Governance

Coming Q4 2025

Goal: Develop participants' understanding of FAIR data principles and their application in pharma, empowering them to enhance data governance, maximize data value, and build relevant skills through practical insights and real-world examples.

Potential Topics

Introduction to FAIR Principles

FAIR Implementation Frameworks

FAIR Maturity indicators

FAIR Data Strategy and Architecture

Technical Infrastructure for FAIR

FAIR in Practice (Case Studies)

Governance of FAIR Data

FAIR roles in cultural Change Management

Business Value of FAIR

FAIR for AI, ML & Digital Transformation

Target Audience:

- Pharmaceutical Researchers and Scientists
- Data Scientists and Analysts in Pharma
- Data and R&D Managers
- Pharmaceutical Technology Leaders
- Decision Makers and Budget Holders

Format:

- A range of 1-hour webinars presented weekly with the suggestion that individuals join only the sessions relevant to them
- Followed by on-demand access



**Looking
for
Speakers**



Communities



Controlled Substance Compliance & Shipping Expert Community



PM: Birthe Nielsen

Start: 1-Jan-2025

End: 31-Dec-2025

Next Phase: 1-Jan-2026

The Challenge

Legislation relating to pharma R&D continually changes, and each change adds complexity and potentially widened controls. Maintaining compliance is challenging and consequences for breaches can be severe.

Shipping of pharmaceutical product including customs, cold chains, dangerous goods and infectious material legislation is particularly complex.

Share Best Practices & Expertise to Develop Industry-wide Positions on Legislation and Support Compliance

Purpose:

Focuses on tackling common challenges, developing solutions, communicating with regulators, and learning from each other to stay current and compliant with legislations governing controlled substance activities.

CSCS Conference planned for 2025

Completed deliverables:

- Communication with Regulators (UK and US)
- Development of two commercial compliance solutions

Community Supporters



In Person Event

PISTOIA ALLIANCE



1st European Controlled Substance Compliance &
Shipping Conference

Novartis Campus, Basel, Switzerland

**October 2 & 3
2025**

User Experience in Life Sciences – UXLS (2025)



PM: Farah Egby

Start: 1-Jan-2025

End: 31-Dec-2025

Next Phase: 1-Jan-2026

The Challenge

User Experience (UX) is a powerful, evidence-based design process that puts user behaviors and needs at its core. While the benefits of UX are well-established in sectors like retail, its adoption in more complex fields, such as the life sciences, has been comparatively slow.

Bringing UX to life sciences poses a unique set of challenges: existing best practices require significant adaptation to fit the scale and complexity of life science applications.

Bringing User Experience to the Life Sciences

The Pistoia Alliance UXLS community empowers life sciences organizations to drive digital transformation by embedding user experience principles and supporting adoption of best practices.

2025 workstreams: **AI for UX and UX for AI** considers AI workflows and UX design for innovation, user satisfaction, with leadership in AI strategy and governance. **Best Practices for UX in Life Sciences** demonstrates and supports UX value with frameworks, case studies, and best practices.

The UXLS community organizes a unique **conference** enabling UX experts and practitioners to share best practices and industry trends.

Community Supporters





In Person Event



User Experience for Life Science (UXLS) Conference
September 22 & 23, 2025, Rahway, New Jersey



Projects



Pharmaceutical CMC Process Ontology (2025)

PM: Birthe Nielsen

Start: 1-Sep-2024

End: 31-Dec-2025



Goals

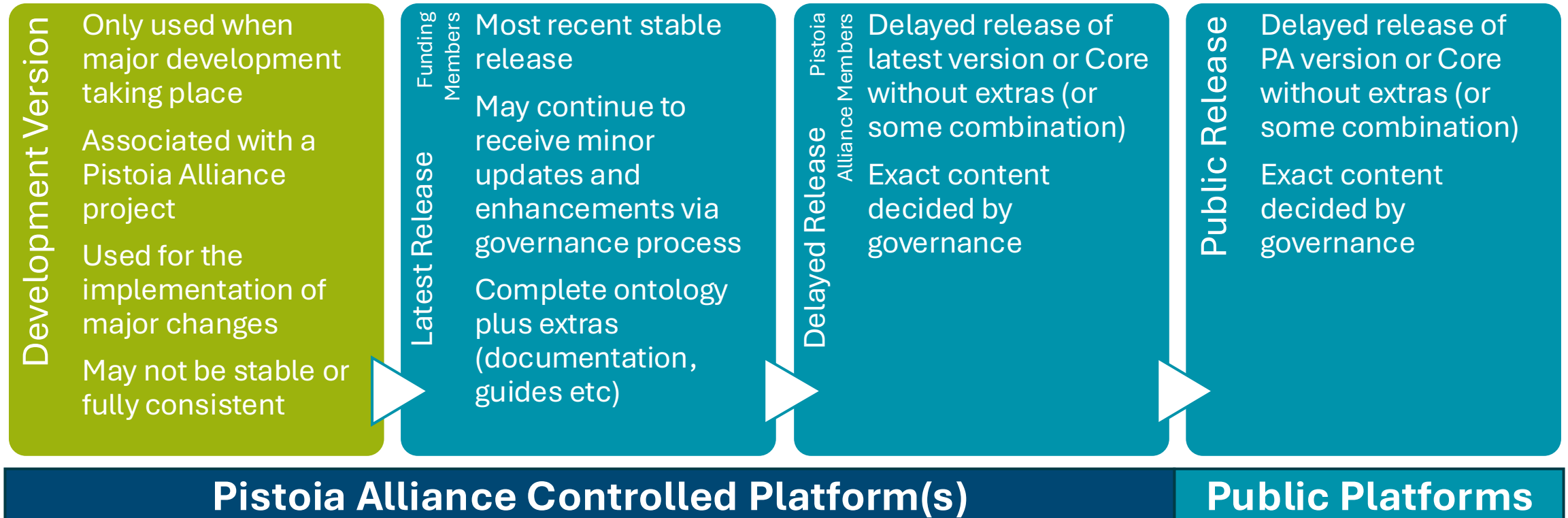
- Developing integration approaches with other ontologies used in the CMC domain.
- Extending beyond small and large molecules and adding domain-specific relationships to create flexibility for other ontologies to connect to the domain lexicons.
- Assessing the feasibility of extending the ontological design to Analytical domain.
- Defining a governance structure for ontology sustainability.

Build a Release Ready Ontology covering all common modalities

- Complete capture of process steps (PSOA Process Scale-Up and Optimization) definitions for chemical, mAb and CAR-T
- Recommended PSOA designation of process step terms for each process supported by key process parameter definitions
- Define ontology integration strategy with key related ontologies such as analytical, equipment, UoM etc.
- Pilot process recipe implementation in vendor software
- Define the process recipe standard (eg. in JSON or XML) for process recipe and process batch data exchange.
- Extend SKOS vocabularies for parameter and recipe component kinds distinguishing configurations, specifications & measurements.
- Demonstrate use cases based on representative data samples and prioritizing those with biggest impact on business efficiency



IDMP-O Sustainability Releases



Pistoia Alliance IDMP-O Sustainability Components

	Pistoia Members Funding Development	Pistoia Members Funding Maintenance	Non-Funding Pistoia Members	Non-Members of Pistoia Alliance
Access to Development Version	✓	✗	✗	✗
Governance of Development ^{*1}	✓	✗	✗	✗
Access to Latest Release	✓	✓	After - 6 months	After - 12 months
Part of Maintenance Governance Team ^{*2}	✓	✓	✗	✗
Basic Technical Support	✓	✓	✗	✗
Implementation Guide (including Best Practices)	✓	✓	✗	✗
SPOR Knowledge Graphs (latest)	✓	✓	✗	✗
GSRS Knowledge Graphs (latest)	✓	✓	✗	✗
Access to Pistoia Alliance IDMP-O Training	✓	✓	Paid	Paid

^{*1} Including selection of use cases for development

^{*2} Including prioritization of bug fixes and enhancements



The Challenge

The GSRS is an open-source registration system for the ingredients in medicinal products. The creation of the GSRS was a collaborative effort between the FDA, NIH and EMA, and driven by the need to accurately track the global supply chain of ingredients in regulated products.

In practice, the system serves many use cases and there is limited stakeholder governance to guide ongoing and future system development, and to provide financial stability.

A Stakeholder Consortium to manage and govern development and support of the open-source GSRS system

- A Consortium acting as the governing body for the GSRS. The mission of the Consortium would be to develop a shared vision for GSRS, develop and enact a business plan to attain the vision, and guide ongoing development and maintenance of the system.
- Development of a business case that meets the needs of all primary stakeholders with a viable funding model for development and maintenance
 - Agreement on a governance approach that balances immediate and long-term needs and delivers value to all stakeholders.
 - Plan to implement a governance, development model and system architecture that works with in today's fast changing business, regulatory and technology environments.

Project Supporters



Large Language Models (LLMs) in Life Sciences Phase 1 (2024)



PM: Vladimir Makarov

Start: 1-Dec-2023

End: 30-Apr-2025

Next Phase: 1-May-2025

The Challenge

There is an interest in LLMs and their use in drug discovery R&D, but the technology is poorly understood.

The main technical challenge is translation from natural language queries into structured queries suitable for Knowledge Graphs. Naïve LLMs fail at this task and new technical solutions are sought.

Learn Best Practices for Structured Scientific Queries with LLMs

- Key issue: translation from natural language to structured queries
- Is important for "scientific chat" in a natural language and for "AI co-scientist" use cases
- Experimented with many technologies
- Target discovery was the practice use case
- Stayed away from training of custom LLMs or using proprietary data
- LLM agents are the best technology for natural language querying
- Drafted a paper that reports the results submitted
- Phase 2 announced: objective is to create benchmark(s) for every step in the scientific chat process

Project Funders



abbvie

Large Language Models (LLMs) in Life Sciences - Benchmarking



PM: Vladimir Makarov

Start: 31-Dec-2024

End: 31-Dec-2026

Next Phase: TBD

The Challenge

We are exploring the best practices for use of LLMs as scientific assistants in translation of natural language questions to structured queries for data mining in a natural language.

We discovered that there are **no accepted community benchmarks** for this process.

This project will serve as an extension of the 2023-2024 project in exploration of the best practices for LLM use in life sciences.

Create Benchmarks for Natural Language Scientific Queries with LLMs

- Benchmark(s) for every step in the natural language data mining process:
 - Understanding the question
 - Recognition of named entities, synonyms, and disambiguation of terms
 - Building the structured query
 - Assessment of the overall answer quality
- Value: allow pharmaceutical and biotechnology research organizations to make better tool use decisions based on the objective evaluation of technologies, and technology vendors to be able to plan product improvements.

Project Supporters





The Challenge

Pharmaceutical companies are developing isolated AI applications that cannot communicate, creating data silos and missed synergy opportunities.

Challenges include decentralised development, regulatory compliance, data security, role conflicts, and workflow orchestration complexity.

Without standards, agents cannot share insights or coordinate effectively.

A Standardised AI Agent-Agent Communication Protocol & Best Practice Guide

Solution

Creation of standards for processing and data packages used to participate in Agentic Workflows and best practices guides for how these should be consumed and orchestrated.

Key Benefits

- Time Acceleration:** Automation accelerates drug development timelines
- Cost Reduction:** Eliminates redundant processes and minimizes manual errors
- Quality Enhancement:** Standardised comms ensures higher-quality outputs
- Scalable Integration:** Connect AI applications without costly overhauls
- Competitive Edge:** Allows Separation of Processing from Business Logic

Project Supporters



LLM (NLP) Use Case Database Phase 2

PM: Vladimir Makarov

Start: 1-June-2025

End: 31-Dec-2026

Next Phase: TBD



The Challenge

Previously we successfully built a **Natural Language Processing (NLP) Use Case Database** capturing real-world implementations across the pharmaceutical industry.

Given the rapid rise of **Large Language Models (LLMs)**, we are now expanding the database to include LLM-specific use cases.

The Value Proposition: Analyze LLM Use Cases

A **Use Case Database** for real-world pharma applications of LLMs.
Benefits:

- Informed decision-making
- Faster adoption
- Risk mitigation

Discover:

- Trends in **what works and scales**
- Where **LLM projects struggle or fail**
- The underlying factors driving these outcomes

Project Supporters



Late-Stage R&D and Sustainability



Change Management for Digital Transformation CoE



PM: Anca Ciobanu

Start: 1-Jan-2025

End: 31-Dec-2025

Next Phase: 1-Jan-2026

The Challenge

Organizations capture less than one-third of the value that respondents expected to see from recent digital transformations and initiatives. To meet business objectives and be effective with available resources, there is a need to use the most appropriate approaches and tools to maximize the return on investments that is made in digital transformations.

Support and Leverage Change Management Best Practices in Digital Transformation

The objective is to bring significant value improvement for members in increasing impact and efficiency with successful transformations

This new Community of Experts aims to :

1. Forge a community to foster dialogue and share best practices in change management for digital transformation
2. Raise awareness of the value of change management in life sciences
3. Create tools and collateral that can support Pistoia Alliance members in their journey
4. Evaluate new ideas

Community Supporters





In Vitro Pharmacology (2025)

PM: Veronique Francois **Start:** 1-Jan-2025

End: 31-Dec-2025

Next Phase: 1-Jan-2026

The Challenge

FDA review of IND in-vitro pharmacology assay data is arduous due to lack of common data structure for safety and secondary pharmacology in-vitro assays.

A format exists for *In Vivo* submission data (CDISC-SEND) but not for critical *In Vitro* pharmacology results generated at pharma companies and CROs.

Existing Pistoia Alliance DataFAIRy project provides a starting point to address the issue.

Regulators, Pharma and CROs Collaborate to Develop Data Standards for IVP Assays to Streamline Review

Public-private partnership with FDA and Pistoia Alliance with working groups focused on:

- Data content transmitted from CROs to pharma to regulators
- Data formats and transmission methods

Deliverables include standardized:

- Template for critical information required for review by regulators
- Assay protocol registration & repository of assays

AbbVie donated their assay protocol registration system (Tabascode) to serve as the open-source assay repository

Project Supporters

abbvie

Zifo

In Vitro Novel Alternative Methods (NAMs)

PM: Veronique Francois Start: 1-Oct-2024

End: 31-Dec-2025

Next Phase : 1-Jan-26



The Challenge

- *In Vitro* Novel Alternative Methods (NAMs) are **human cell-based assays and systems** used in biomedical research to represent human physiology, predict clinical outcomes and reduce, replace or refine the requirements for animal testing in drug discovery.
- Many organizations are validating these platforms but are facing numerous hurdles which include the **lack of standardization** in performance qualification, method ontology, validation standards, and data management but also poor annotation of assay metadata.

The Value Proposition: Harmonize data standards

Public-private partnership with FDA and Pistoia Alliance with working groups focused on:

- Standardizing data structure through utilization of new or existing ontologies for the description of *In Vitro* NAM assays and endpoints commonly used in targeted research areas.
- Harmonizing the provenance of assay metadata and standardizing the measurement of assay performance.
- Output will be human and computer readable standard data, increasing capacity for provenance and attribute connection for insight and analysis.
- This initiative would help promote understanding and adoption of *in vitro* NAMs, to incorporate human-relevant tests to reduce or complement preclinical *In Vivo* studies and would also help facilitate the integration of structured data sets for use in AI/ML applications.

Project Supporters



abbvie

In Vivo Minimal Metadata Set

PM: Veronique Francois Start: 1-April-2025

End: 31-Dec-2025

Next Phase : 1-Jan-26



The Challenge

In the field of biomedical ***in vivo* research, data sharing and repurposing** are not commonly practiced. One critical element for enabling data sharing and repurposing is the **provision of metadata that well describes the raw or primary data.**

There is currently no established, domain-agnostic, minimal metadata set (MNMS) that can be used across various biomedical research fields.

Solving this problem could bring new scientific insights, enable opportunities to **reduce or replace animal experimentation through data reuse & virtual control groups, and bring gains in research efficiencies.**

Define a Minimal Metadata Set (MNMs) for repurposing non-clinical *in vivo* data for biomedical research

Pistoia Alliance will address the challenges of non-harmonized, non-standardized data sources that prohibit repurposing of biomedical research data from animal studies by specifying a minimal metadata set (MNMS) designed to enable the repurposing of *in vivo* data.

The MNMS will align with existing validated guidelines for reporting *in vivo* data (e.g. ARRIVE 2.0) and contribute to making *in vivo* data FAIR-compliant with a focus on the reusability aspect to reduce the unnecessary consumption of animals.

Project Supporters



abbvie

 **TECNIPLAST**
innovation through passionSM

Impact of Digital on Clinical Trial Carbon Emission



PM: Thierry Escudier

Start: 1-March-2025

End: 31-Dec-2025

Next Phase: 1-Jan-2026

The Challenge

Following the 'Carbon Footprint of Decentralized Clinical Trials' project run by Pistoia Alliance members in collaboration with Sustainable Healthcare Coalition members in the industry Low Carbon Clinical Trials (iLCCT) consortium, and the release of the Clinical Trial Carbon Calculator, there is a need to get a better understanding of the impact of digital : this project extension will aim to focus data collection of the digital tools used in clinical trials including digital trial delivery (such as eConsent, DCT & tokenization) and digital health solutions (such as devices, apps, BYOD)

Define a Common Methodology to Measure the Climate Impact of DCT and Digital Components

The objective is to bring significant value improvement to members in using the calculator in increasing the understanding on how to reduce Co2 emissions in designing new clinical trials when using digital :

Collect case studies and collect key insights from members on use of digital tools

Collect data from members on emission impact from digital tools

Contribute to perform analysis on use of digital in clinical trials

Project Supporters



Social Media Listening

PM: Aditya Tyagi

Start: March-2022

End: Dec-2025

Next Phase: TBD



The Challenge

Social media listening is a potentially valuable RWD source for patient insights.

However, there are no comprehensive regulatory guidelines governing the use of social media as a source of RWD for patient-focused drug development

Develop Best Practices & Guidelines for the Use of Social Media Analysis as RWD for Patient-Focused Drug Development

Guidelines produced will reduce uncertainty in the use of social media as RWE to support product-focused drug development and contribute to create methodological best practices that pave the way for sponsors to collect and submit evidence based on social media as part of their integrated evidence generation and regulatory submission activities. The first deliverable was achieved to explain the current state, the challenges and the opportunities. Publication published in **Frontiers in Medicine** in March, 2024.

Additional project deliverables/activities will be:

- Define best practices on Data collection/Data analysis
- Strategy discussion with regulators for establishing guidelines
- Understand patient perspective on social media listening through survey/interviews
- Additional publication

Project Supporters



Pharmacovigilance Systems & Processes Standards (PS²) - Case Intake



PM: Marc Graber

Start: May-2025

End: Dec-2025

Next Phase: Jan-2026

The Challenge

Today's pharmacovigilance (PV) solutions lack interoperability, making it difficult for pharmaceutical companies to adopt best-of-breed solutions, thereby limiting innovation and automation.

Although PV system requirements are broadly similar across organizations, each company independently defines specifications for procurement and implementation, driving up costs and complexity for both buyers and vendors.

Create Case Intake Solution Standards

Goals: Establish common Case Intake solution requirements & interoperability standards and an operational approach to sustain those standards & certify software solutions against them.

Pharmaceutical Company Benefits:

- Lower costs: Requirements re-used across sponsor companies, less customizations
- Easier and faster solution evaluation via certification process
- More innovation: Ability to integrate best-in-class solutions, PV resources focused on innovation, vendors incentivized to differentiate through innovation

Vendor benefits:

- Lower costs: Re-use answers to RFPs, no need to elicit standard requirements
- Higher requirements maturity

Deliverables: Case Intake standard solution requirements, data and interoperability standards, certification process

Project Supporters

abbvie



Safety and PV AI Forum

PM: Thierry Escudier

Start: May-2025

End: Dec-2025

Next Phase: Jan-2026



The Challenge

The business problem is the lack of a unified platform for industry professionals to share insights, best practices, and challenges related to the emerging field of AI in PV, as well as to develop a harmonized understanding of regulatory expectations and requirements. This gap hinders the advancement and implementation of AI technologies, which are crucial for improving the efficiency and accuracy of safety and PV processes.

Enhance the Use of AI in PV

Goals: A forum to address the challenge of effectively integrating AI into PV practices by driving innovation, reducing duplication of efforts, and ensuring that best practices are widely adopted, benefiting the entire industry and the patients.

Industry Benefits:

- Share use cases and tools
- Learn from others
- Define best practices for PV industry
- Participate to communications and get recognized

Deliverables: The Safety and PV AI Industry Forum project would produce several key deliverables to ensure its success and value to its members. These deliverables include: collaborative projects to work on best practices, industry position, tool kit, metrics, communication and building a knowledge repository and resources

Community Supporters



PISTON
ALLIANCE

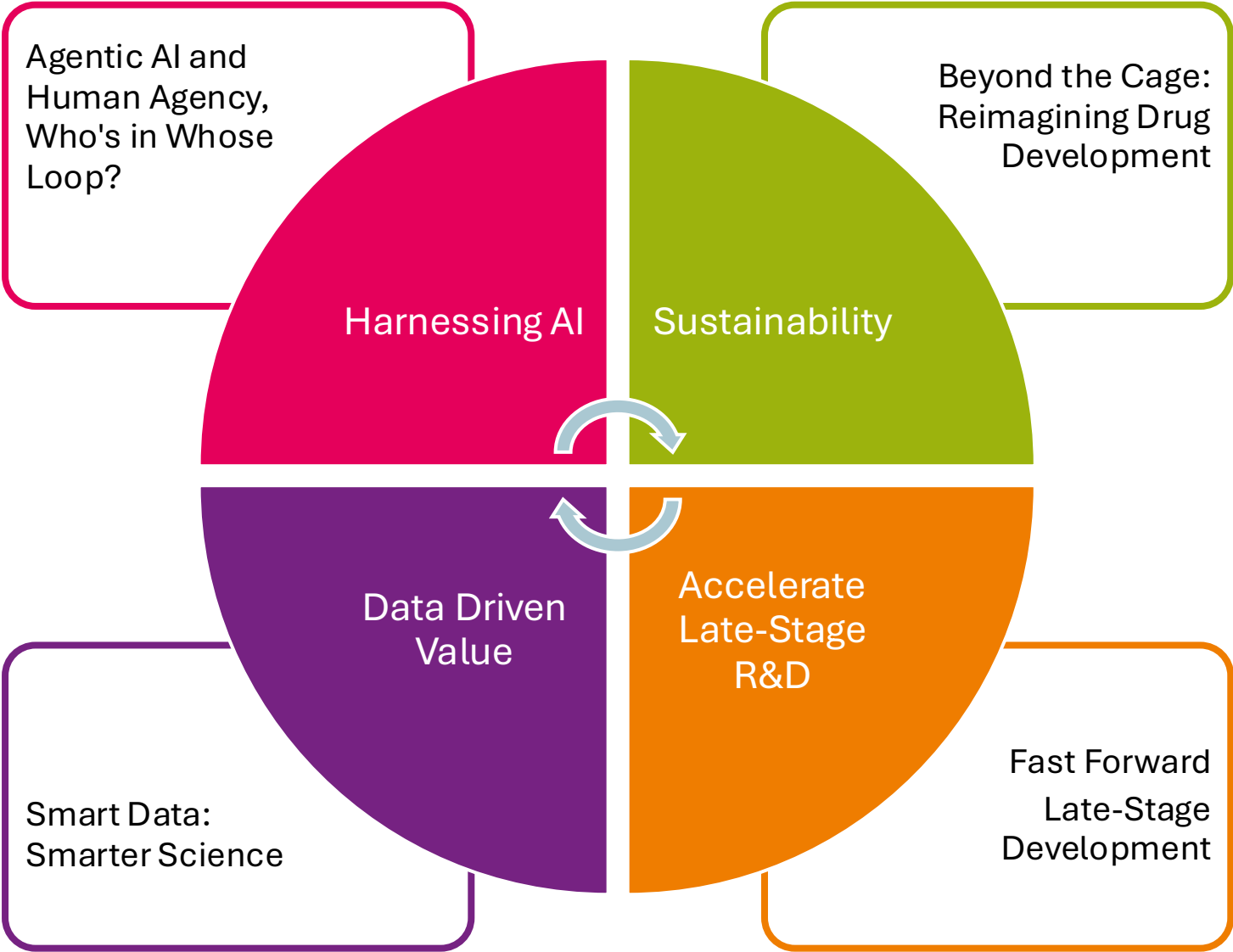
Annual Member 2 Day Conference



SAVE THE DATE for Boston, November 11 & 12, 2025
data driven innovations



Pistoia Alliance USA Conference 2025 Streams



Questions



2025 FAIR-for-Pharma



Funders / Steering Committee

abbvie



Active Members Include

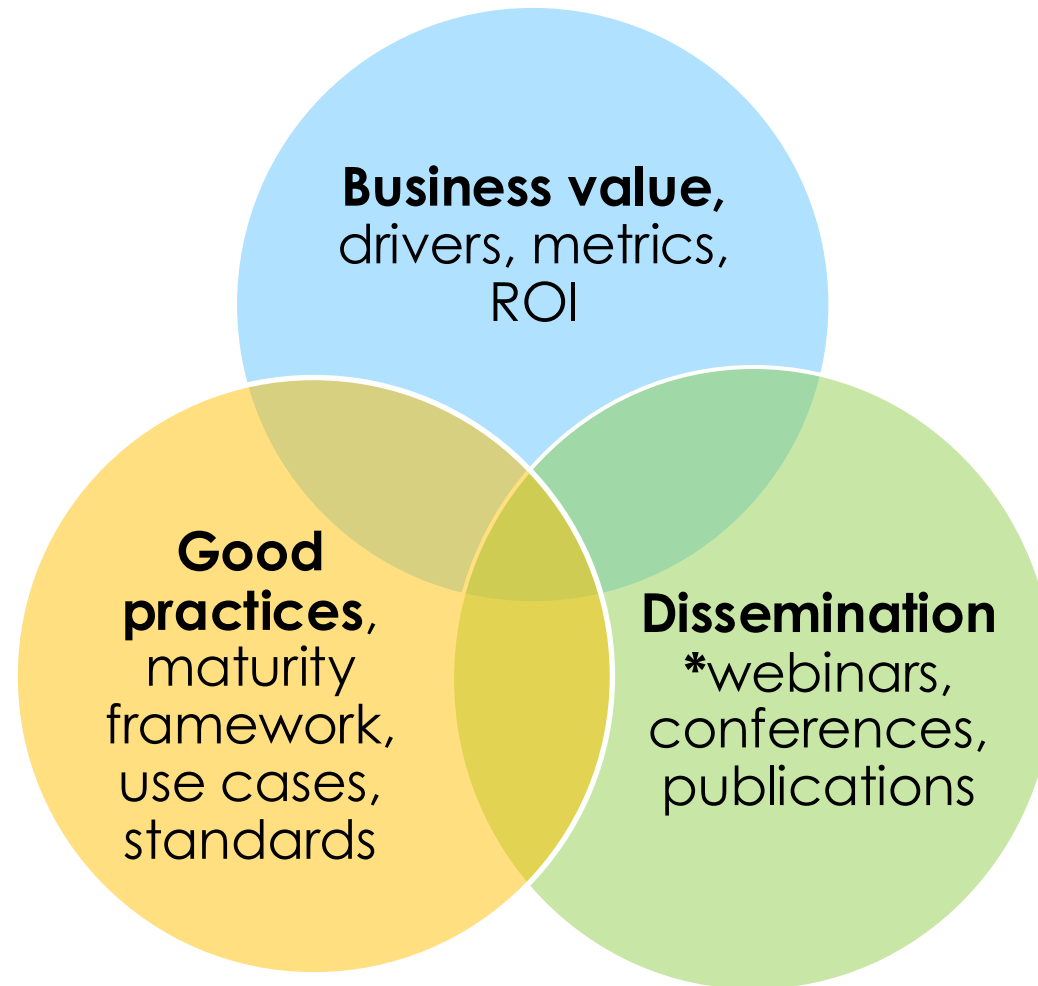


CSL Behring



FAIR-for-Pharma Community of Experts 2025

What do we do?



*not a formal working group