PISTOIA ALLIANCE

Pistoia Alliance Portfolio

July 2026, JC Baber & T Escudier

Discussion Overview

01 Semantic Data in Pharma

02 Practical Artificial Intelligence

03 Change, Adoption & Digital Transformation

04 Data Exchange

05 Late Stage R&D

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

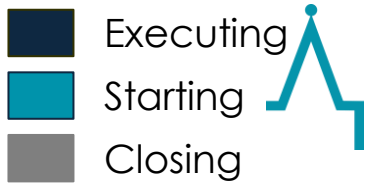
Accelerate Late-Stage R&D

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

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

Data Enrichment
- DataFairy Bioassay Annotation
- FAIR-for-Pharma Community
- Open-source Instrument Data Converter
- Digital Analytical Methods*[2]

Data Integration & Management
- *In Vitro* Pharmacology*[1]
- *In Vivo* Efficacy Data Standards
- Data Governance Community
- GSRS Consortium
- Chemical Exchange Format Community

Harnessing AI to Accelerate R&D

Algorithms
- Benchmarking LLMs for Natural Language Data Mining

Best Practices
- LLM & NLP Use Case Database
- Life Sciences AI Exchange
- Updated IDMP-O Training

Sustainability Driven R&D

- Carbon Footprint for Decentralized Clinical Trials
- *In Vitro* NAMs Data Standards*
- Digital Transformation & Change Management Community
- Minimal *In Vivo* Metadata for Data Reuse & Repurposing

Accelerate Late-Stage R&D

- Social Media & RWE
- Pharmovigilance Case Intake & Processing

Other Activities

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

*[1] Public Private Partnership with the FDA

*[2] In Collaboration with Allotrope

Ideas Aligned with Strategic Priorities

Delivering Data-Driven Value

Ontologies

- Instrument & Equipment Ontology
- Experiment & Assay Ontology

Data Enrichment

- Synthetic Data in Medical Imaging

Data Integration & Management

- Standardized bioinformatics (NGS) computational pipelines
- HELM Format Enhancements
- FAIR Maturity Matrix as a Service
- Knowledge Graph Community

Harnessing AI to Accelerate R&D

Algorithms

- Secure Benchmarking of AI Solutions
- LLM Accelerators for Ontology Development & Adoption
- Federated Foundation AI Models for Pathology and Toxicology

Best Practices

- Utilization of AI/ML results in Regulatory Filings

Sustainability Driven R&D

- Updated IDMP-O Training
- Upcoming Training: Digital Transformation for Labs
- Upcoming Training: AI Ready Data

Accelerate Late-Stage R&D

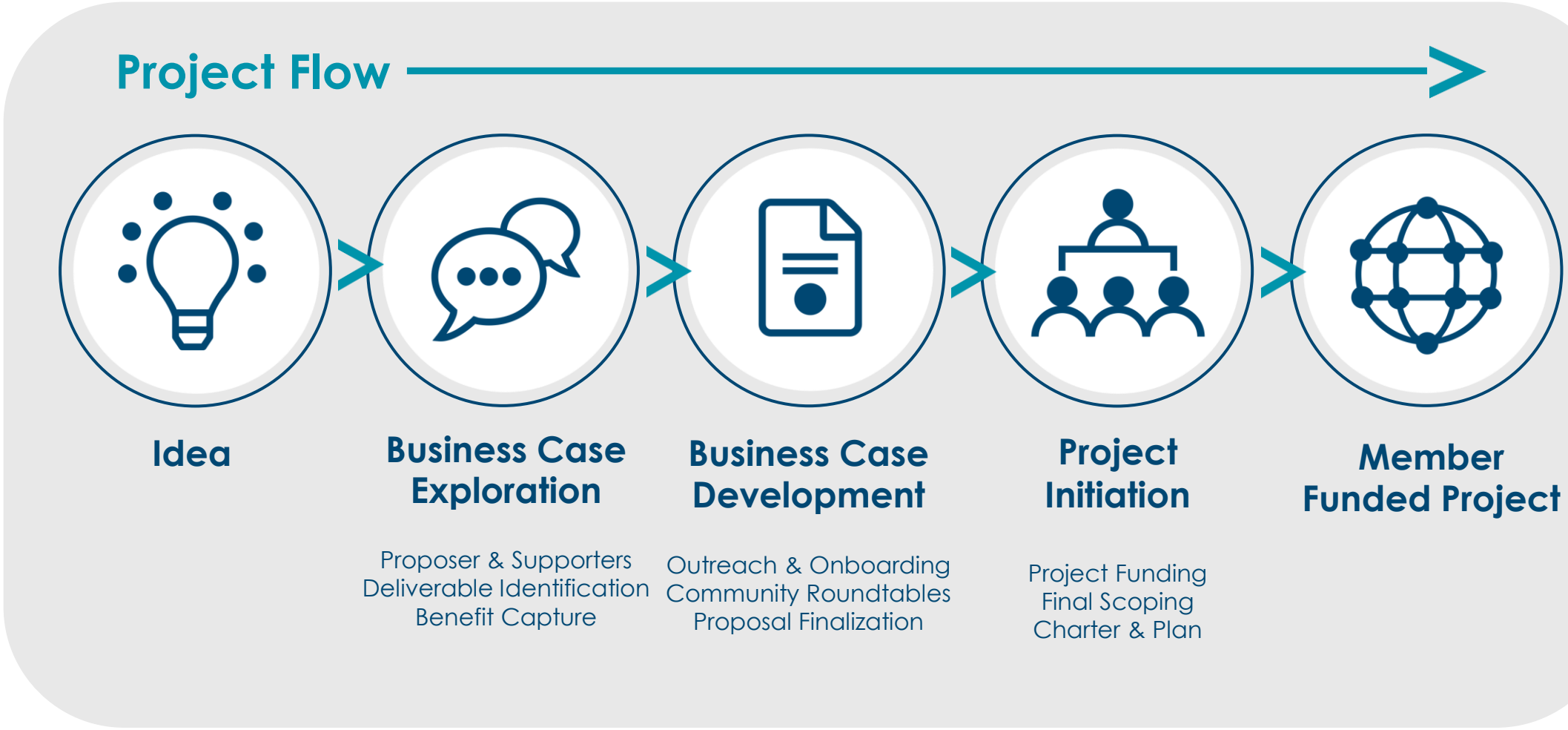
- Safety & PV AI Forum
- Regulatory Submission of Omics Data Collected in Clinical Trials
- Pharmaceutical Quality Knowledge Model
- ClinCilico : AI in Clinical Trials
- Clinical Trial Patient Matching

Other Activities

- Regional Export Restrictions on Shipping & Data Exchange
- Enabling the Citizen Developer
- The Common Corpus: Literature to Knowledge
- EHRs for Early Disease Detection

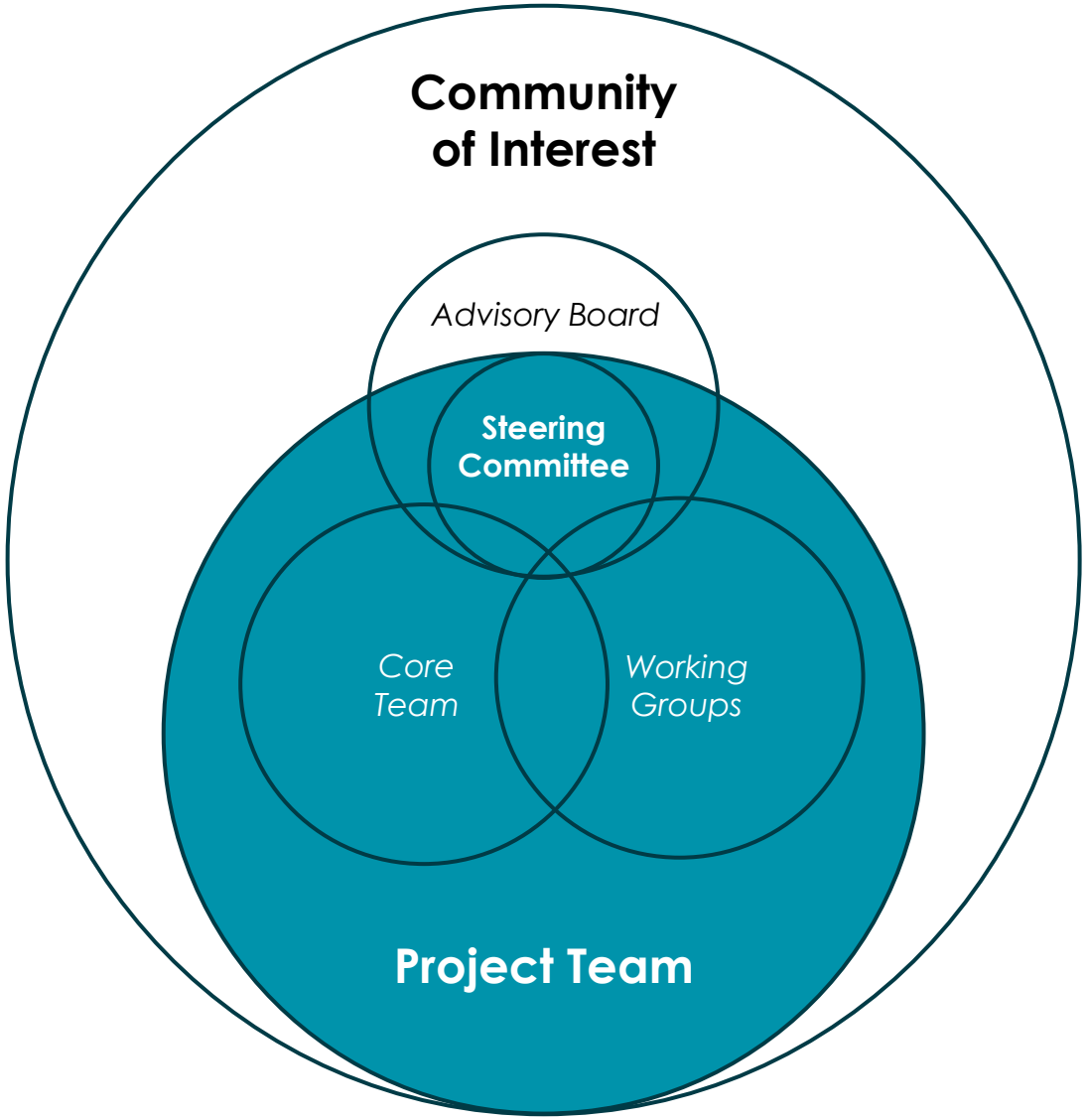
Idea Development

How an idea becomes a project at Pistoia Alliance



Submit Your Ideas via Our Website: [Submit an Idea - Pistoia Alliance]

Pistoia Alliance Project Team Structure



Pistoia Alliance Members

Wider Community

Optional teams in *italics*

PISTOIA ALLIANCE

Semantic Data in Pharma

Standards Make Life Easier

Why is it so difficult to get people to adopt them?

IDMP-Ontology
Trajectory & Outlook

2021
Initiation

2022 (Phase 1)
IDMP-O MVP

2023 (Phase 2)
IDMP-O v1.0 Build

- Suitable for Production Use

2024 (Phase 3)
Pharma Implementation & Industry Adoption

- V1.0 Public Release
- Production Implementation
- Collaboration with Health Authorities
- Interoperability with other domains
- Introduction of Open-Core Model

2025
Sustainable Release & Focused Development

- Separate Maintenance and Development
- Complete shortages use case (EMA-PMS alignment)
- Develop CMC/PQ use cases (batches)
- Production implementation and industry adoption

2026-2029
Maintained Ontology with Incremental Enhancements

- Focus on implementation and adoption
- On going bug fixing and enhancement
- Extension beyond ISO standard
- Refactoring for usability
- 3-year partnership with ACCURIDS

1st IDMP-O User Group Meeting
After Pistoia Alliance Conference (Nov 2028)
Hosted by Amgen, Boston

CMC Process Ontology

A vendor-agnostic semantic architecture to standardize laboratory and plant production process recipes with integration points to analytics, materials, equipment and quality, facilitating tech transfers and interoperability.

PISTOIA ALLIANCE

Industry Challenge

1. There is currently no machine-readable representation of CMC processes and process development.
2. CMC data is fragmented and non-interoperable across sites, partners, and systems.
3. As a consequence, process knowledge is rarely accessible and reusable across the lifecycle.

Goals and Intended Deliverables

- Expanded, modular ontology with capability-based equipment mapping.
- CAR-T manufacturing.

erability
ELN, MES
MS

AI-ready and contextualized data

CMC-PO Version 1.0 Released to Members May 2026

(Public Release scheduled for Nov 2026)

Time Savings

Machine-readable recipe exchange cuts rework, errors, and timelines

Faster Innovation

Standardized, contextualized data are ready for analytics and AI

Improved Compliance

Traceability across steps, materials, equipment, QA

Increased Quality

Unambiguous vocabulary from R&D, manufacturing, and quality to regulatory

J&J Lilly Roche AMGEN abbvie GSK ZS CROWN POINT TECHNOLOGIES MERCK AstraZeneca The Hyve Takeda

FAIR Community of Experts (COE)

Our mission is to enable FAIR data principles implementation to accelerate trustworthy, interoperable, and reusable data practices across the life sciences with best practices, enabling methods and tools.

PISTOIA ALLIANCE

Industry Challenges

1. FAIR transformation demands industry-wide engagement. Single company cannot drive this alone.
2. Lack of common standards and practices hinder interoperability across teams, service providers and partners
3. Legacy systems, processes, culture, skill sets. FAIR data transformation journey requires organizational changes

Goals and Intended Deliverables

framework and calculator
ools and infrastructure
es, webinars and tutorials

FAIR Maturity Matrix v1.2
Released April 2026

FAIR Business Value Framework *alpha* Release
June 2026
(Release to Members scheduled for Sept 2026)

dustry
llective
shapes
delivers
utions.

Knowledge Exchange across companies, among experts and peers

Community Facilitator

Increasing industry-wide quality reduces errors and rework

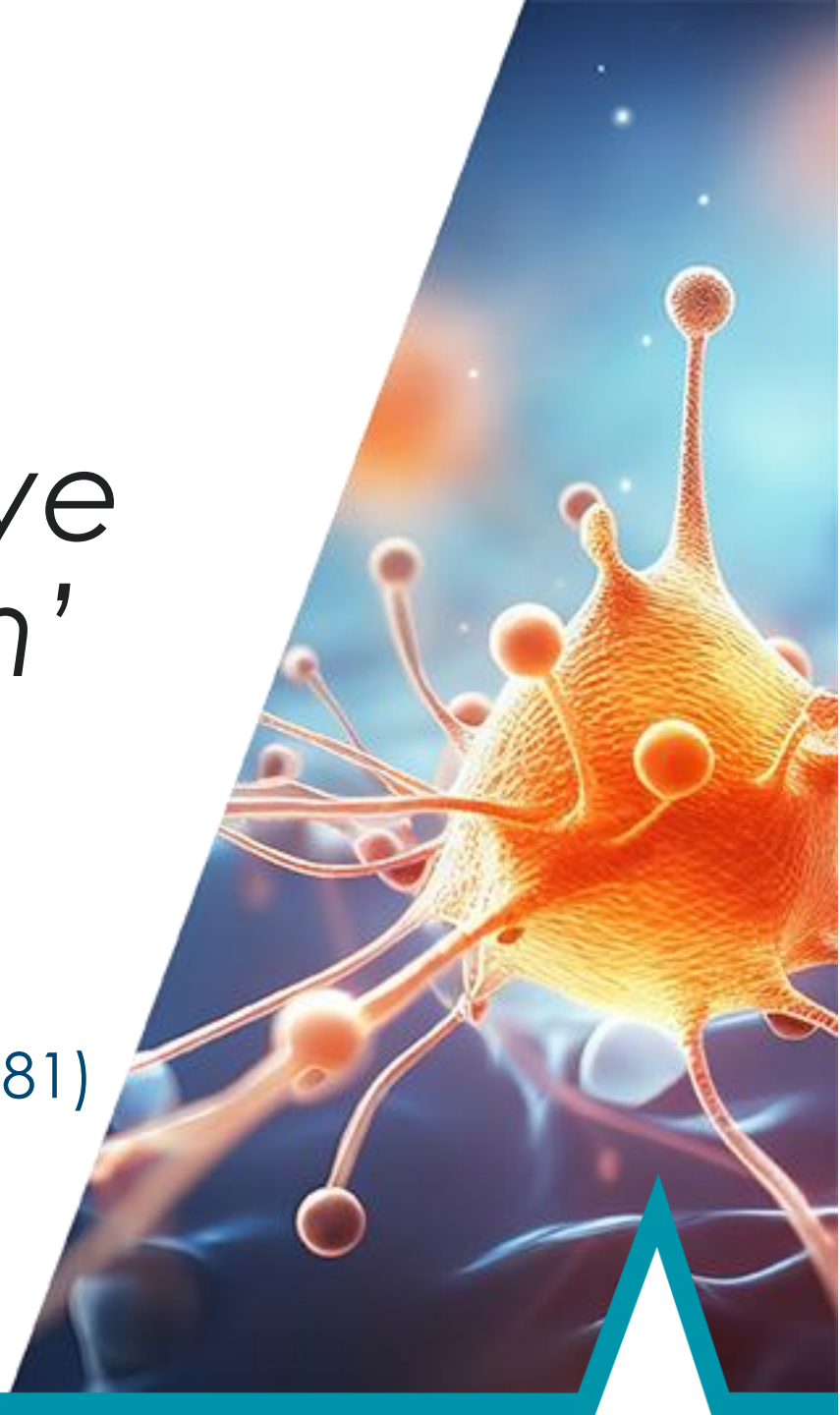
Dates & Budget

Jan. Dec. 2026- Min $50k - Ticket: $10k for a steering group seat (large pharma)

abbvie BAYER ONTOFORCE Pfizer Roche

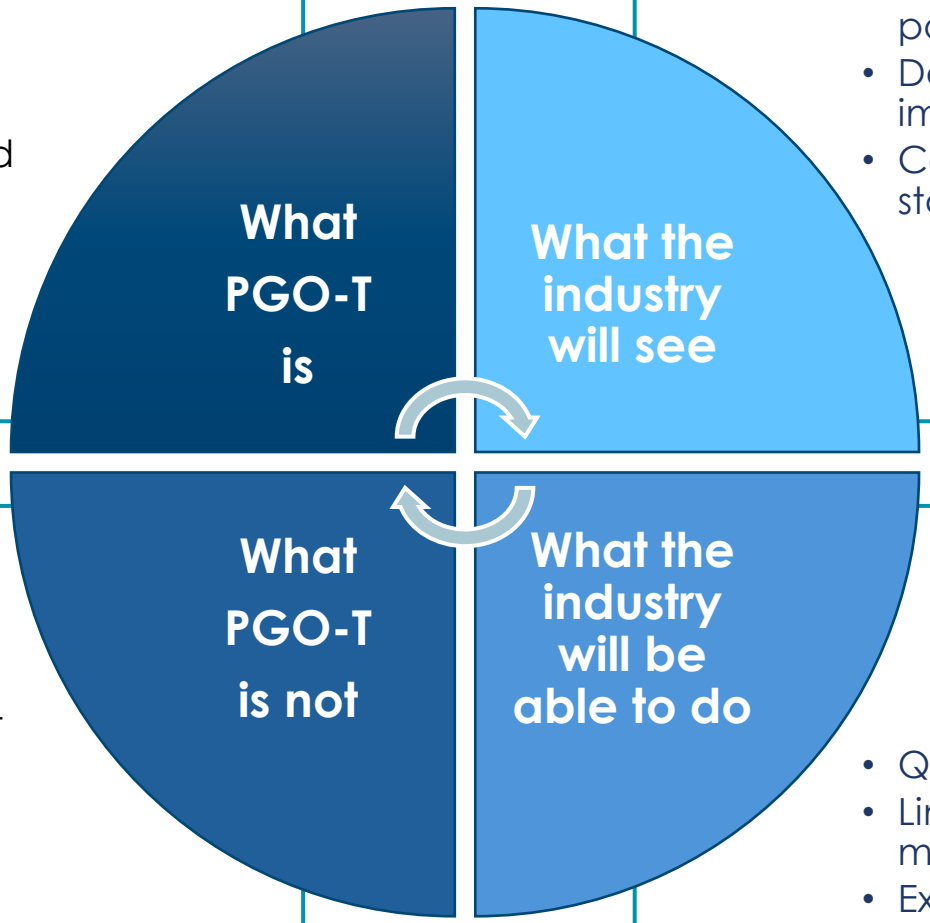
‘The nice thing about standards is that you have so many to choose from’

Andrew S. Tanenbaum, Computer Networks (1981)



Pharma General Ontology (PGO)-Terminology

- A community-governed reference set of core pharmaceutical concepts
- A "lingua franca" — declares clearly which terms are used and why they were chosen
- Open access



- Machine-readable Core-terminology pointing to public definitions
- Declaration / endorsement: who is implementing PGO-T internally
- Community resource implementing W3C standards

- Not an ontology or replacement for MedDRA, SNOMED, CHEBI…
- Not a top-down mandate: community alignment
- Not domain-specific: bridges all pharma domains

- Quicker decisions on reference sources
- Linkable & reusable URIs with unifed mapping of core-concepts
- Exchange data across CROs, partners, & regulators while maintaining internal URIs

Covered Concepts (v1.0)

Example

Class: Indication

A sign, signal, circumstance, or symptom which serves to indicate or point out the cause, pathology, treatment, or issue of an attack of disease; the basis for initiation of a treatment for a disease or of a diagnostic test (causal, or symptomatic, or disease-specific indication)
https://evsexplore.semantics.cancer.gov/evsexplore/concept/ncit/C41184

▶ Alternative Descriptions

▼ Aliases

- Therapeutic Indication

▼ Examples

Value	Description
abnormal blood glucose concentration	too much or too little glucose in the blood).
Hypotension	A condition in which the blood pressure in the arteries is abnormally low.

▼ Annotations

property	value
definition_source	https://evsexplore.semantics.cancer.gov/evsexplore/concept/ncit/C41184
definition_source_attribution	NCI-Thesaurus (NCIT)
definition_source_licence	CC-BY-4.0
definition_access_date	2025-05-30
expert_approval_date	2025-06-12
conceptual_clarification	The group acknowledged the complexity of the term 'Indication', which spans both pre-approval (e.g., clinical investigation) and post-approval (e.g., regulatory labeling) contexts. Core challenges in terminology harmonization included--

- The restrictive use of terms such as 'abnormal' or 'pathological', which may unnecessarily narrow the scope.
- Differentiation between 'sign' and 'symptom'.
- The regulatory distinction between investigational and approved uses of a medicinal product.
- The need for a resolvable and publicly licensed URI, a key requirement for incorporation into PGO.

While definitions from IDMP-O and CDISC were considered valuable, the absence of resolvable unique identifiers in some cases limited their immediate applicability. Notably, the CDISC glossary version of 'Indication' appears to reuse the NCIt concept C41184, but with a modified definition—raising concerns about versioning and semantic traceability. |

PGO-T: Long-Term Vision

To Supply Chain and Beyond!

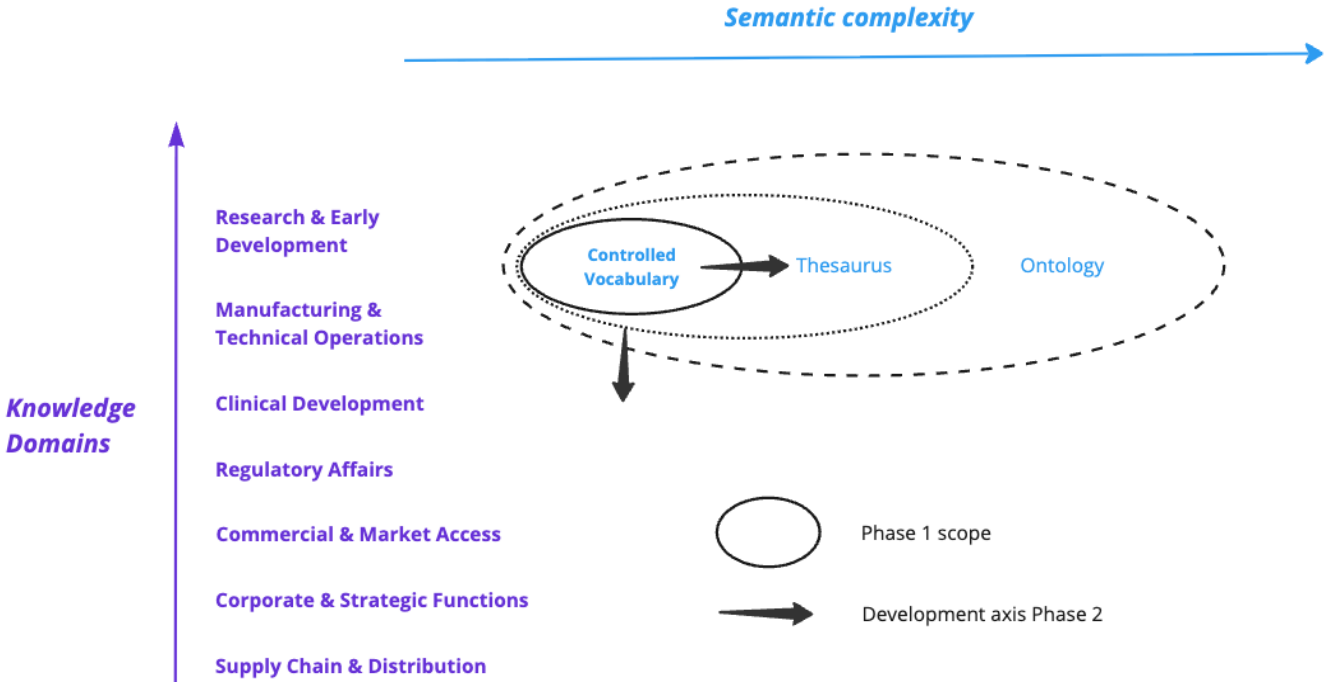
Phase 2 (2025–2026) — In progress

Expanding Coverage:
Additional concepts include Adverse Event, Excipient, Instrument, Medical Condition, Molecular Target, Study Site, Study Subject (> 20 concepts)

Expanding Usability:

- Implementing schemas for machine-readable output
- Generating mappings to external ontologies

Ensuring Sustainability:
PGO as permanent, maintained, industry reference data product



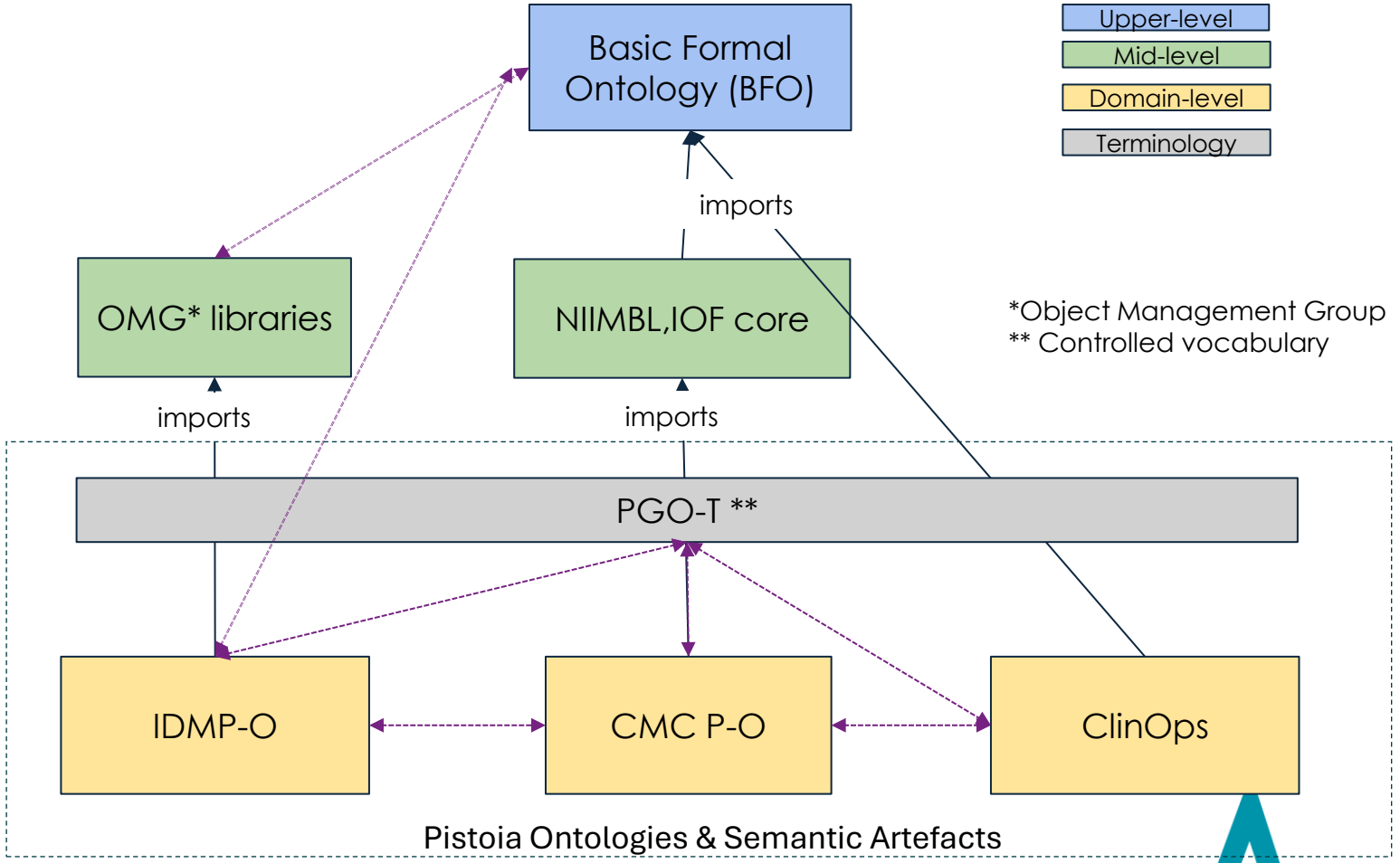
Aligned Semantic Activities at Pistoia Alliance

– Pharma General Ontology (PGO-T)
Shared terminology & semantic reference for cross-organisation interoperability

– IDMP Ontology
Regulatory identification of medicinal products

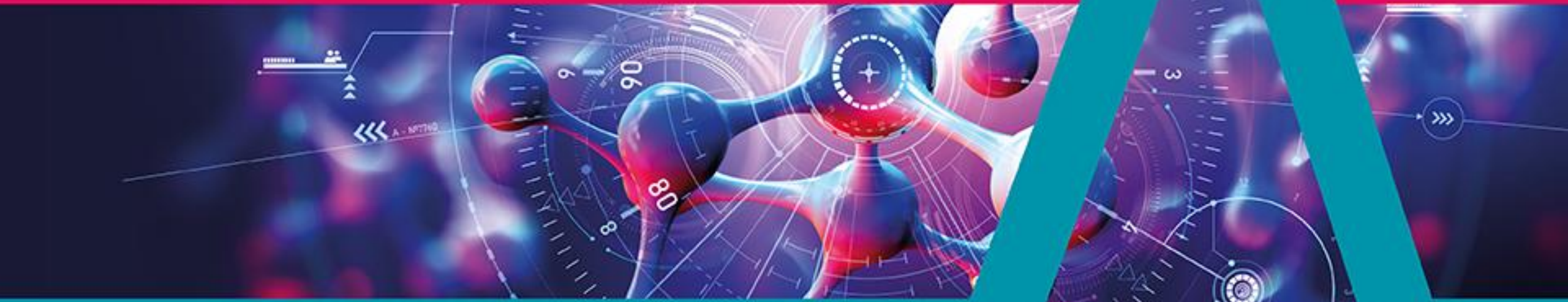
– CMC Process Ontology
Chemistry, Manufacturing & Controls process semantics

– ClinOps Ontology
Clinical operations shared vocabulary



PISTOIA ALLIANCE

Practical Artificial Intelligence



LLM & NLP Use Case Database

A Use Case Database built around real-world pharma applications of LLMs that provides an essential foundation for informed decision-making, faster adoption, and risk mitigation.

PISTOIA ALLIANCE

Industry Challenge

1. LLMs are currently the wild west of technology. Many opportunities but lots of unknown.
2. Most companies are working on variations of the same things, to different degrees of success.
3. Many projects succeed in theory but fail to scale and be adopted.

Goals and Intended Deliverables

- A curated catalogue of real-world pharma applications of LLMs for NLP
- Documentation of both successful deployments and failures with causes
- Analysis of trends to determine the underlying factors driving outcomes
- Exclusive live resource for Pistoia Alliance members

Value Proposition

- Avoid common pitfalls & increase successful adoption
- Enable strategic planning through a broad survey of the landscape
- Find underlying factors determining success & failure
- Know what works and what scales for informed decision making

Project Benefits

Cost Savings
Avoid wasted expense by learning from the failures of others

Increased Quality
Avoid pitfalls that work but ultimately lead to failure

Faster Innovation
Stand on the shoulders of giants to push boundrys

Time Savings
Understand what works and does not before starting

Roche

Boehringer Ingelheim

Benchmarks for NL Data Mining with LLMs

PISTOIA ALLIANCE

Industry Challenge

1. Systematic methods to assess LLM-based natural language (NL) data mining tools used in R&D are missing.
2. Lacking performance benchmarks, users risk blind spots in the quality of results.
3. The absence of appropriate test sets complicates the development of NL data mining tools.

Pistoia Alliance Value Proposition

- Novel member legal framework expedites collaboration
- Success as a neutral forum for standard, policy building
- Critical mass of members leads to influence and adoption
- Access to third party vendors leads to fit-for-purpose tools/IT

Goals and Intended Deliverables

- Deliver benchmarks to assess the true performance of LLM natural language query assistants in the context of Pharma R&D Data.
- Measure and ensure the quality of identified or constructed benchmarks.
- Communicate the findings and lessons learned.

Steering Committee Benefits

Faster Innovation
Accelerate adoption of AI tools

Community and Networking
Share benchmarks & lessons learned

Increased Quality
Drive creation & use of better tools

Pharma and Life Sciences AI/ML Training

Academics and Industry, Live in 2025, Now available on demand

Original Syllabus

1 - Introduction to AI for drug discovery
2 - Generative AI for Drug Development
3 - Computer vision
4 - Multimodal Deep Learning
5 - LLMs in early discovery
6 - LLMs as Agents in DD
7 - AI in Computational Medicinal Chemistry
8 - Knowledge graphs
9 - Digital pharmaceutical manufacturing

2026 Update (Sept-Nov 2026)

10 - AI Fundamentals and AI Ready Data
11 - Knowledge Management
12 - AI Tools and Workflows
13 - Governance and Responsible AI
14 - UX and Change Management for AI Adoption
15 - AI ROI Measurement and Value Realization
16 - Building an Enterprise AI Strategy for Pharma
17 - Autonomous Pharma R&D and the Future of AI

Confirmed Trainers from

- AbbVie
- AstraZeneca
- Campana & Schott
- GSK
- Ipsen
- Novo Nordisk
- Rancho BioSciences

PISTOIA ALLIANCE

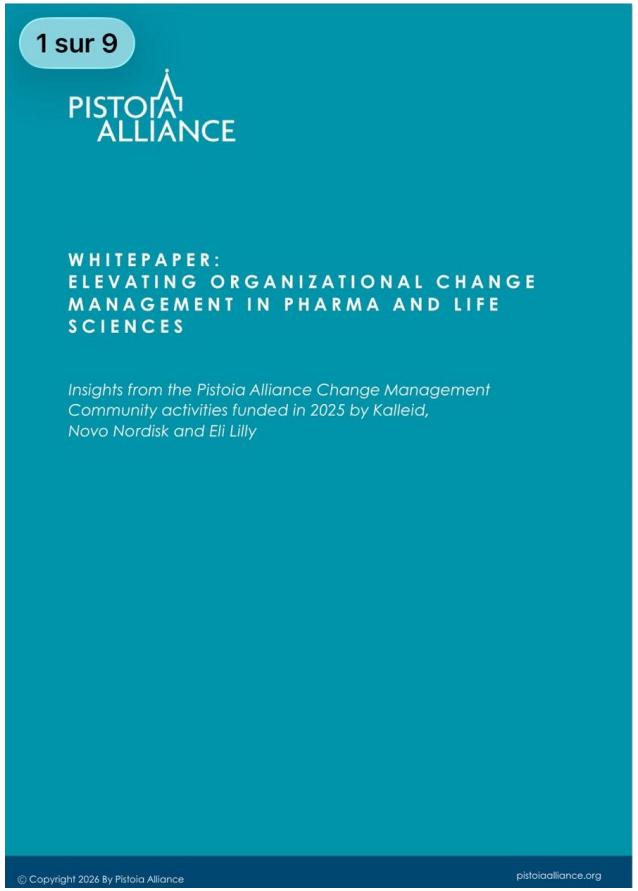
Change, Adoption & Digital Transformation

Change Management and Digital Transformation

The Technology is Not the Problem

- Dynamic community to enhance change management practices to for digital transformation across the industry
- Change management does not get the support that it deserves
- White paper released in Jan 26

  ‘Elevating Organizational Change Management in Pharma and Life Sciences’



At the Core of Digital Transformation

- Successful multi-community workshop run in Spring 2026
  - Combining UXLS, Change management and Future Lab Evolution
  - Accelerating Digital Transformation in Pharma Labs
- Building on the output
  - A series of 6 online sessions from August to November
  - A workshop at the Boston conference
  - Produce guidelines for “Balancing short-term ROI with long-term transformation”

User Experience in Life Sciences (UXLS) Community

A community bringing together life sciences professionals to recognise and embed UX as a key strategic enabler, adapting to AI-driven change and industry trends.

PISTOIA ALLIANCE

Industry Cha

1. UX remains
as mission-
the Life Scie

2. UX processe
be adapted
regulated, c
developmer
of the Life S

3. AI-driven
restructurin
budget pres
demand that UX
prove its strategic
value.

ion experience
typing and testing
in Q2
mittee & community

e
s

Focus on creating and testing truly effective UX for Life Sciences

Access new strategies; avoid known pitfalls, build on best practices

Learn from experts about tools, processes, trends, and practices

Empower UX teams with structures, skills, and roles for UX in LS

Increased Quality
Foster improve usability of LS tools; eliminate friction in value chain

NOVARTIS GSK MERCK

Controlled Substance Compliance & Shipping Expert Community

PM: Birthe Nielsen **Start**: 1-Jan-2026 **End**: 31-Dec-2026 **Next Phase**: 1-Jan-2027

The Cho

Legislation relati
R&D continually
each change a
and potentially
controls. Mainta
compliance is c
consequences f
can be severe.

Shipping of phar
product includin
cold chains, dar
and infectious m
legislation is particularly
complex.

ry-wide Positions on nce

ing solutions,
each other to stay
controlled substance

rch)
olutions

NOVARTIS Johnson&Johnson Innovative Medicine MERCK

CERTARA Scitegrity

FAIR and Data Governance Training

Live from Q3 2025-Q1 2026, Now available on demand

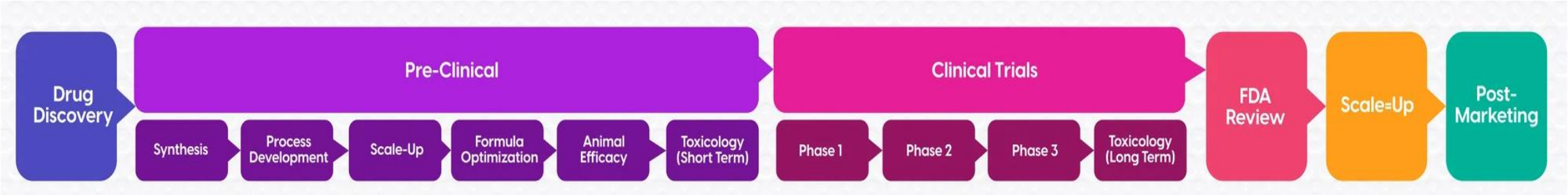
1: Introduction to **FAIR Principles**
2: Making Data **Findable**
3: Making Data **Accessible**
4: Making Data **Interoperable**
5: Making Data **Reusable**
6: **FAIR Implementation Frameworks**
7 – **FAIR Maturity Indicators**
8: **Technical Infrastructure for FAIR**
9: FAIR in Practice (**Case Studies**)
10: FAIR roles in cultural **Change Management**
11: **FAIR in Life Science** from Research to Clinical
12: **Business Value of FAIR**
13: **FAIR for AI, ML & Digital Transformation**
14: **Data Governance**
15: **FAIR Data Strategy & Architecture**

112 Participants so Far

PISTOIA ALLIANCE

Data Exchange

DATA STANDARDIZATION & AI READINESS INITIATIVES



- PPP with FDA

In Vitro Pharmacology

In Vitro NAM Data Harmonization

In Vivo Laboratory efficacy data standards

Minimal Meta data Set (MNMS) project

Data Standardization & AI Readiness Initiatives

Driving consistency, efficiency, and innovation in life sciences through data standardization and AI readiness

	IVP Result Standardization	In Vitro NAM Data Harmonization	In Vivo Efficacy Data Standards
CHALLENGE	• No harmonized IVP result formats • Inconsistent regulatory submissions	• Non-standardized metadata • Siloed data sources	• Fragmented *in vivo* efficacy datasets • Challenging to integrate & analyze
DELIVERABLES	• Standard IVP submission template • Open-access assay repository	• Harmonized metadata standards • Standard result formats	• CDISC SEND–aligned data model • Pilot implementation
BENEFITS	• Reduced duplication & cost • Improved predictive safety modeling	• Human-relevant methods identified • Increased regulatory confidence	• FAIR datasets enabling AI-driven analysis • Faster regulatory submissions

Stakeholders: Researchers • Pharmacologists • Toxicologists • Academics • Data Managers • Statisticians • CROs • Regulators

Some Highlights

- **In Vitro Pharmacology (IVP) :**
  - PPP with the FDA
  - Release of v1.0 IVP assay repository
  - A standardized IVP data template for IND/NDA submissions (screening and dose response)
  - An ontology-aligned framework
  - Development of an IVP Glossary

- **In Vitro Novel Alternative Models (NAM) :**
  - PPP with the FDA
  - Standardised Template for Endpoint Result Data Collection and Assay metadata in Cardiac Safety and Hepatotoxicity In Vitro NAM Assays
  - Controlled Vocabulary and Glossary for In Vitro NAMs
  - Manuscript under preparation

- **Minimal Metadata Set (MNMS) :**
  - Validation of the minimal metadata set by demonstrating ≥95% of MNMS template completion when data from 5 companies are tested
  - Manuscript under preparation

- **In Vivo Laboratory Efficacy Data Standards :**
  - Mapping of oncology data with SEND IG 4.0 and Animal Rule IG 1.0
  - Development of Standardized SDTM-Compatible Domain(s) for Efficacy Studies (In Progress)

Key Take away messages

- With those 4 projects, we contribute to help our members :
  - to be aligned with 3R rules (Reduce, Reuse, Recycle)
  - to leverage the use of AI
  - to speed up regulatory submissions
  - enabling integrated, interoperable, and FAIR-aligned preclinical data exchange

PISTOIA ALLIANCE

Late-Stage R&D

Accelerating Innovation in Late-Stage R&D

Late-stage R&D generates enormous volumes of clinical, regulatory, safety, and real-world data. Success increasingly depends on:

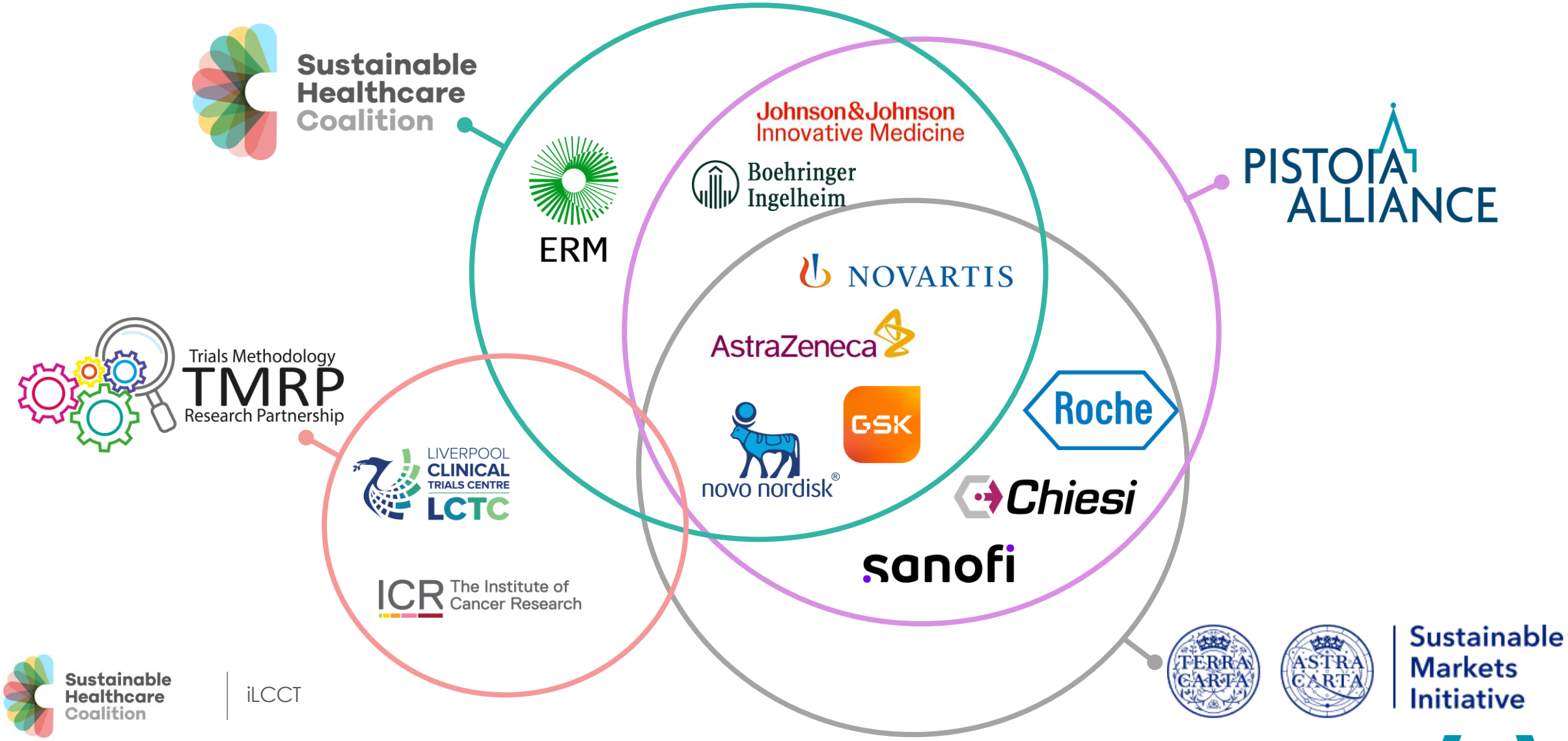
- Cross-company collaboration
- Data interoperability
- AI-ready standards
- Faster evidence generation
- Regulatory trust
- Sustainable digital transformation

All the Pistoia Aliance fundational work and models serve to help late-stage R&D processes

Focus in late stage R&D portfolio

- Carbon footprint clinical trials
- Social Media Listening for RWD
- PV
- AI and Clinical Trials

The industry Low Carbon Clinical Trials consortium (iLCCT)



The iLCCT Clinical Trials Carbon Calculator

Who it's for

The iLCCT Clinical Trials Carbon Calculator is for anyone involved in any stage of clinical trials design and management

What

The iLCCT Clinical Trials Carbon Calculator is free to use and is a step-wise, activity-based process to support design of low-carbon clinical trials and their reporting, including scenarios and Docusaurus functionality for ease of use

Why

To ensure the impact at scale needed to mitigate climate change, we need a consistent approach to reduce the carbon emissions from clinical trials

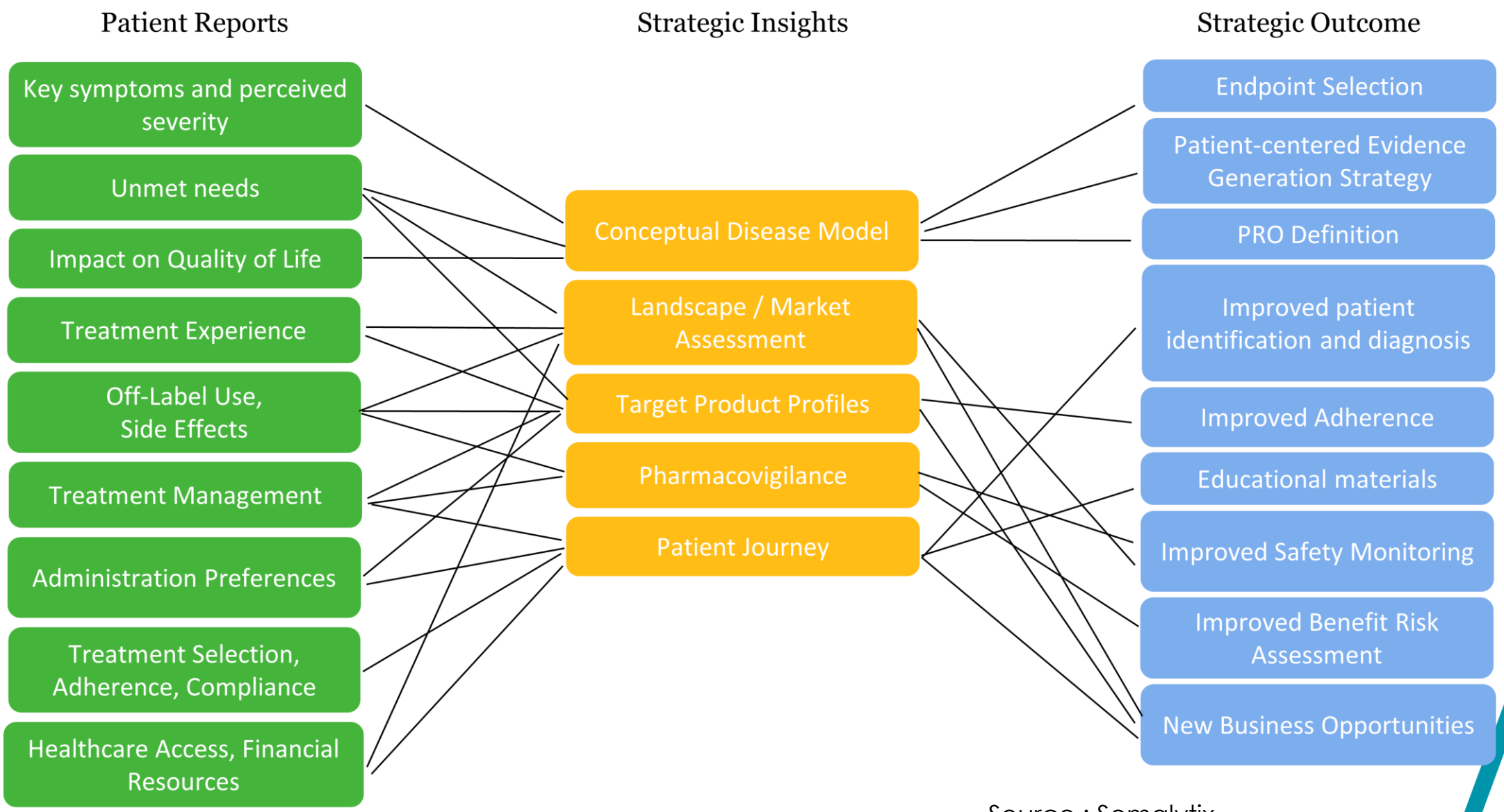
For entire supply chain adoption of low carbon processes, coordination is also essential

AstraZeneca Johnson&Johnson sanofi Roche GSK Trials Methodology TMRP Research Partnership ERM PISTOIA ALLIANCE Chiesi novo nordisk® Boehringer Ingelheim

Some Highlights

- **Carbon footprint clinical trials :**
  - Calculator V3 launched April 2nd
  - Assessment of impact of “use of digital tools”ongoing :
    - Focus is on saving patients travels and monitors travels
  - Community of Practice is expanding : > 400 members WW

Real World Social Data turns patient voices into strategic outcomes



Source : Semalytix

What have we done ? Key Deliverables

- **Defining Best Practices for Data collection, Analysis & validation :**
  - **2 publications : one position paper, one methodology**

frontiers | Frontiers in Medicine

TYPE Policy and Practice Reviews
PUBLISHED 06 March 2024
DOI 10.3389/fmed.2024.1274688

Check for updates

OPEN ACCESS

EDITED BY
Carla Matos Torre,
University of Lisbon, Portugal

REVIEWED BY
Prathamesh Karmalkar,
Merck Group, India
Carlos Alves,
University of Coimbra, Portugal

*CORRESPONDENCE
Philipp Cimiano
cimiano@cit-ec.uni-bielefeld.de

RECEIVED 08 August 2023
ACCEPTED 12 February 2024
PUBLISHED 06 March 2024

CITATION
Cimiano P, Collins B, De Vuono MC, Escudier T, Gottowik J, Hartung M, Leddin M, Neupane B, Rodriguez-Esteban R, Schmidt AL, Starke-Knäusel C, Voorhaar M and Wieckowski K (2024) Patient listening on social media for patient-focused drug development: a synthesis of

Patient listening on social media for patient-focused drug development: a synthesis of considerations from patients, industry and regulators

Philipp Cimiano[1,2]*, Ben Collins[3], Maria Carmela De Vuono[4], Thierry Escudier[5], Jürgen Gottowik[6], Matthias Hartung[1], Mathias Leddin[6], Bikalpa Neupane[7], Raul Rodriguez-Esteban[6], Ana Lucia Schmidt[6], Cornelius Starke-Knäusel[1], Maarten Voorhaar[3] and Krzysztof Wieckowski[8]

[1]Semalytix GmbH, Bielefeld, Germany, [2]CITEC, Bielefeld University, Bielefeld, Germany, [3]Boehringer Ingelheim International GmbH, Ingelheim, Germany, [4]Chiesi Farmaceutici SpA, Parma, Italy, [5]Pistoia Alliance, Wakefield, MA, United States, [6]Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd., Basel, Switzerland, [7]Takeda Pharmaceuticals Co., Ltd., Cambridge, MA, United States, [8]SageFiber, Ustron, Poland

frontiers | Frontiers in Medicine

TYPE Policy and Practice Reviews
PUBLISHED 30 January 2026
DOI 10.3389/fmed.2025.1703923

Check for updates

OPEN ACCESS

EDITED BY
Cristiana Sessa,
Oncology Institute of Southern Switzerland (IOSI), Switzerland

REVIEWED BY
Md Sadique Hussain,
Uttaranchal University, India
Diana M. J. Delnoij,
National Health Care Institute (ZIN), Netherlands
Chantal Britt,
Bern University of Applied Sciences, Switzerland

*CORRESPONDENCE
Philipp Cimiano
cimiano@semalytix.com

RECEIVED 12 September 2025
REVISED 18 November 2025
ACCEPTED 03 December 2025
PUBLISHED 30 January 2026

Best practices for the collection and analysis of patient experience data from social media for patient-focused drug development

Philipp Cimiano[1,2]*, Nicole Brazda[1], Matthias Hartung[3], Cornelius Starke-Knäusel[1], Ana Lucia Schmidt[4], Maria Carmela De Vuono[5], Aditya Tyagi[7], Jürgen Gottowik[4], Raul Rodriguez-Esteban[4], Ben Collins[6], Krzysztof Wieckowski[6] and Thierry Escudier[7]

[1]Semalytix GmbH, Bielefeld, Germany, [2]Faculty of Technology, Bielefeld University, Bielefeld, Germany, [3]Boehringer Ingelheim GmbH, Ingelheim, Germany, [4]Roche Innovation Center Basel, Basel, Switzerland, [5]Chiesi Farmaceutici S.p.A., Parma, Italy, [6]Boehringer Ingelheim International GmbH, Ingelheim, Germany, [7]Pistoia Alliance, Wakefield, MA, United States

POMELO : Protocol for social media listening online

Presented at ISPOR
May 18th 2026

Applying the Best Practices Framework to Enhance Reporting Standards in Patient Listening Studies from Social Media

PCR45

POMELO: PROTOCOL FOR SOCIAL MEDIA LISTENING ONLINE

Philipp Cimiano[1,2], Nicole Brazda[1], Matthias Hartung[3], Cornelius Starke-Knäusel[1], Ana Lucia Schmidt[4], Maria Carmela De Vuono[5], Aditya Tyagi[6], Jürgen Gottowick[4], Raul Rodriguez-Esteban[4], Ben Collins[7], Krzysztof Wieckowski[7], Thierry Escudier[7]

[1]Semalytix GmbH, Bielefeld, Germany, [2]Faculty of Technology, Bielefeld University, Bielefeld, Germany, [3]Boehringer Ingelheim GmbH, Ingelheim, Germany, [4]Roche Innovation Center Basel, Basel, Switzerland, [5]Chiesi Farmaceutici S.p.A., Parma, Italy, [6]Pistoia Alliance, Wakefield, MA, United States, [7]Boehringer Ingelheim International GmbH, Ingelheim, Germany

01 BACKGROUND & OBJECTIVES

- Social media listening (SML) can capture real-world patient experience.
- Reporting standards remain inconsistent across studies.
- POMELO provides structured best-practice guidance.
- Validation of framework using a T2DM case study.

02 RESEARCH METHODS

- The research team **successfully applied the comprehensive POMELO framework** to evaluate a specific T2DM clinical case study.
- We analyzed social media discussions from **32 online forums between 28 Feb 2023 and 28 Feb 2025**.
- The primary research study focused heavily on **symptoms reported by patients**, **daily burdens**, and **various related comorbidities**.

Validated **POMELO** as a **reporting standard** through a review of **peer-reviewed SML studies** in published over the past two years. Assessed reporting quality based on the presence and clarity of:

01 Study setting · 02 Data sources · 03 Algorithmic Approaches · 04 Analytical methods

05 CONCLUSION

- **POMELO** consolidates best and widely used practices for **SML**.
- Integrates existing methodological standards into one structured framework.
- Addresses key gaps and inconsistencies in study reporting.
- Supports improved **study quality** and **interpretability**.
- Enhances the **credibility** and broader **acceptance** of SML evidence.

FULL PUBLICATION

[1]Cimiano P, Brazda N, Hartung M, et al. Best practices for the collection and analysis of patient experience data from social media for patient-focused drug development. Frontiers in Medicine. 2025;12

Funding
www.pistoiaalliance.org

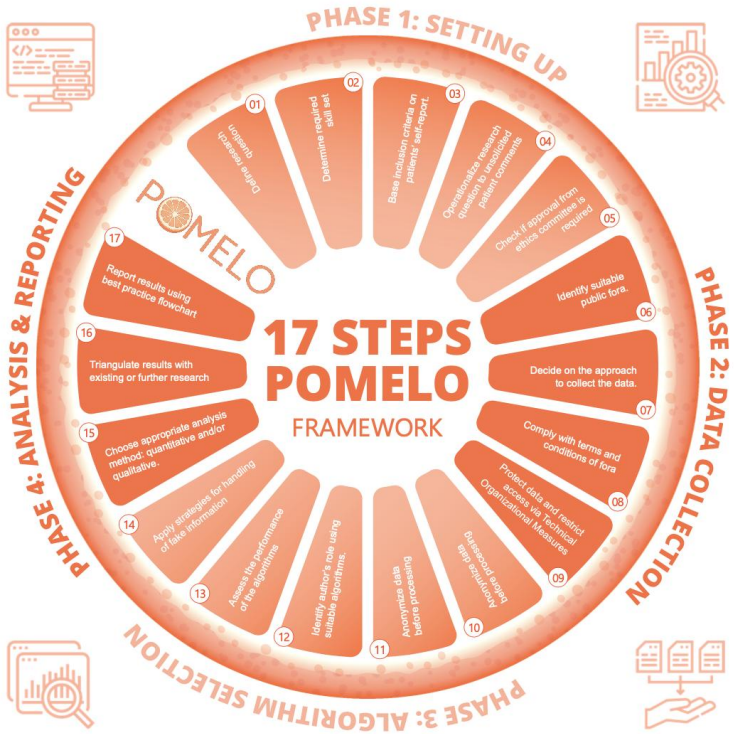
Contact Information
cimiano@semalytix.com

JOIN US!

We are looking for more companies to act as funders and contributors, if you are interested, please contact: aditya.tyagi@PistoiaAlliance.org

TO DOWNLOAD POMELO FRAMEWORK, SCAN THIS QR CODE

@pistoiaalliance
www.pistoiaalliance.org



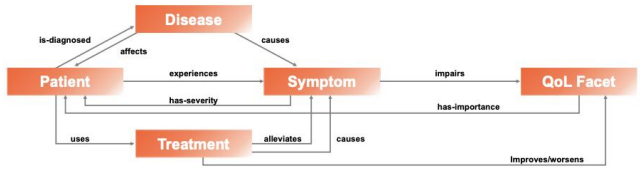


Fig. 1: Conceptual patient experience model (CPEM) to define the relevant information types/variables that can be extracted from patients' posts in the context of SML studies.

03 T2DM CASE STUDY SNAPSHOT

527,268 Document Analysed

3,243 Patients Identified

OBSERVATION PERIOD

February 2023 – February 2025 · 32 Online Forums

FOCUS

Symptoms · Burden · Comorbidities in T2DM

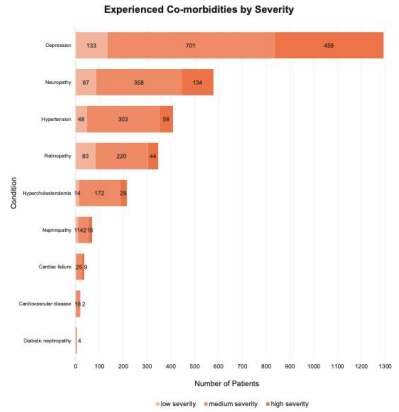


Fig. 2: Bar chart of the experienced comorbidities by severity in the T2DM example case study. Total number of patients reporting a certain symptom, broken down into three severity levels: 1, low; 2, medium; and 3, high. The severity was assessed automatically. From [1].

04 VALIDATION

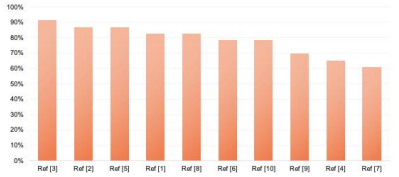


Fig. 3: Bar chart of POMELO concept mapping to published SML studies.

Patient perspective and regulatory interaction

- To investigate and document patient attitudes toward sharing health information on social media and evaluate the potential use of these data in drug development : patient survey conducted in
  - 3 countries
  - 3 therapeutic areas
  - 57 patients interviewed
- Survey is completed, analysis finished, publication under preparation
- Next phase under discussion with Steer Co members

- Interaction with EMA following publication of their "Patient Experience Data position paper"

More in late stage R&D

- **Pharmacovigilance :**
  - case intake project ongoing
  - Forum “Safety and PV AI” ready to start

- **New : AI in Clinical Trials : “ClinSilico”**
  - Collaboration between Pistoia Alliance and CUHP (Cambridge University Health Partners) and CCAIM (Cambridge Center for AI in Medecine) : several use cases planned
    - A range of digital twins
    - Patient recruitment tools
    - Adaptive trial design tools
    - AI clinical trial agents
  - More to be announced in a webinar September 15th

Pre-Conference Workshops (Morning of 10th November)

- From Recipes to Knowledge Graphs: A Hands-On Intro to the CMC Process Ontology
- Defining AI-Ready Data for Life Sciences
- Accelerating Digital Transformation in Pharma Labs

PISTOIA ALLIANCE

Q &A