ONTOLOGIES MAPPING

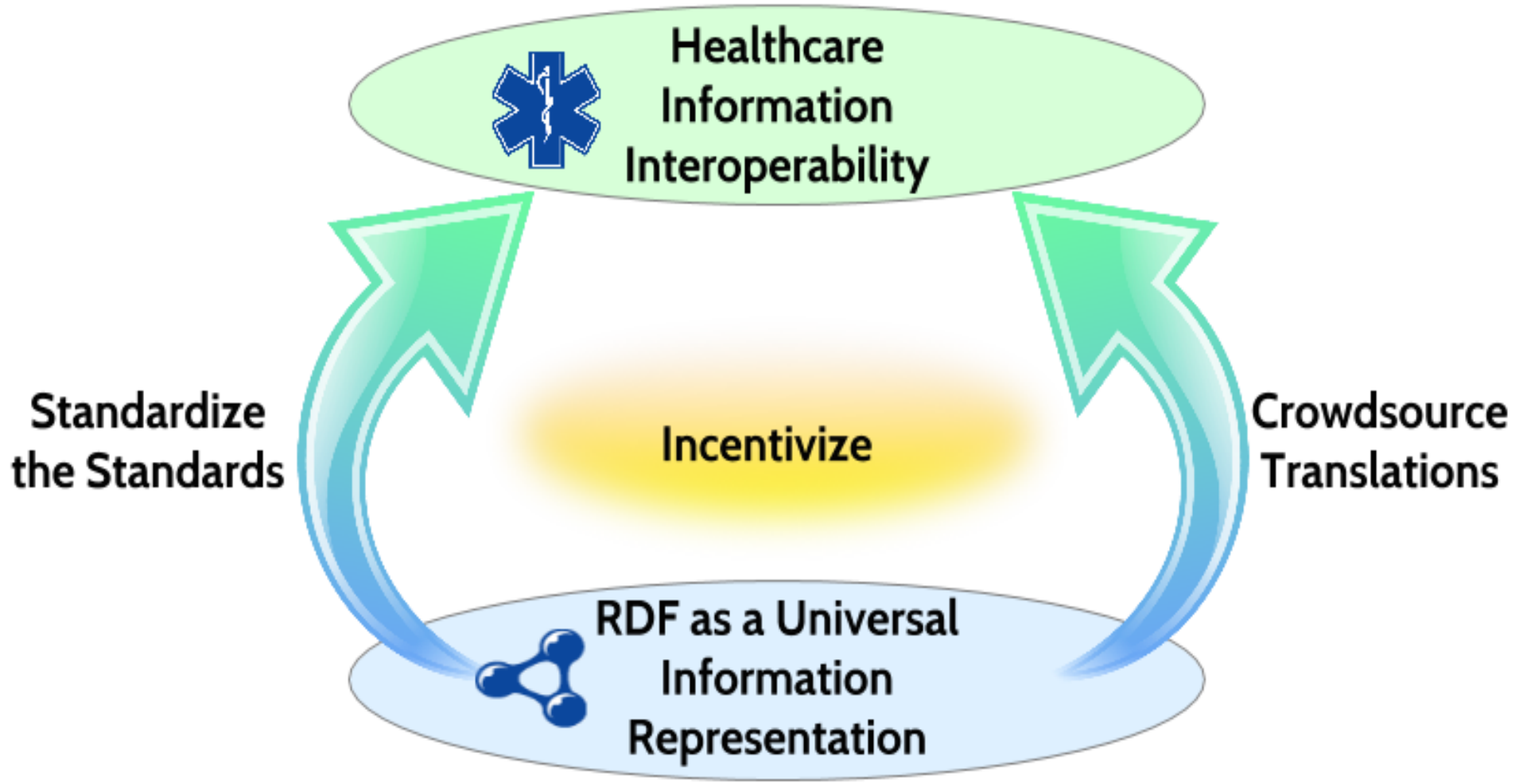
Pistoia Alliance

The Pistoia Alliance Ontologies Mapping Project

12th November 2015

Dr Ian D Harrow for the Project Team

Interoperability Roadmap



http://YosemiteProject.org/

Outline

1. Introduction
2. The Pistoia Alliance
3. Ontologies Mapping Project

ONTOLOGIES MAPPING

Pistoia Alliance

1. Introduction: Ian D Harrow

- Owns and directs Ian Harrow Consulting which provides project management and expert services in bioinformatics and text mining for the life science industry

- Strong involvement in consortia such as Open PHACTS, GA4GH, BioExcel and Pistoia Alliance – Project Manager

- Over 30 years experience in the Pharmaceutical and Life Sciences Industry.

- Post-doctoral studies at Columbia University, New York.
- PhD from the University Of Cambridge, UK.
- BSc. Hons in Zoology from University of Nottingham, UK.

© Pistoia Alliance

2. The Pistoia Alliance

- A global, not-for-profit alliance of life science companies, vendors, publishers, and academic groups that work together to lower barriers to innovation in Research & Development.
- An Open Innovation Community:

http://www.pistoiaalliance.org

© Pistoia Alliance

Membership Levels

- Three tiers of membership are available, allowing all types of companies to get involved.

C	P	M
Core for life sciences companies	**Participating** for technology, content & services providers	**Contributing** for individual and small start-ups (≤ 7 FTE)
Eligible to stand for election to board		*Join conference and webinar organising groups*
Attend members-only events		*Participate in project teams and working groups*
Help prioritise, select and fund Pistoia Alliance projects		*Networking across the membership*

http://www.pistoiaalliance.org/about/resources

© Pistoia Alliance

ONTOLOGIES MAPPING

Current Members

Pistoia Alliance

ACD / Labs (M)	Boehringer Ingelheim (P)	Dotmatics (P)
Amgen (C)	Bristol-Myers Squibb (C)	Hewlett Packard (P)
AstraZeneca (C)	Cambridgene (M)	Eagle Genomics (M)
Bayer (C)	CCDC (M)	EMBL-EBI (M)
Binocular Vision (M)	Certara (M)	EPAM (M)
Biochemfusion (M)	ChemAxon (P)	EPCC (M)
BioIT (M)	Connected Discovery (M)	Fulcrum (M)
BioReference Laboratories (P)	Databiology (M)	Genestack (M)
BioVariance (M)	Deloitte (P)	GlaxoSmithKline (C)
BIOVIA (P)	DeltaSoft (M)	Grit Systems (M)

Ian Harrow Consulting (M)	Merck & Co. (C)	Promeditec (M)	Swiss Institute of Bioinformatics (M)
InChiTrust (M)	Merck KGaA (P)	QFAB (M)	Syapse (M)
Industrial Lab Automation (M)	Molecular Connections (M)	quattro research (P)	The Edge (M)
Instem (P)	Novartis (C)	Roche (C)	The Jackson Laboratory (P)
IO Informatics (M)	Omixon (M)	Royal Society of Chemistry (M)	The Hyve (M)
Ipsen (M)	Oracle (P)	Sanofi (C)	Thomson Reuters (P)
Jeeva (M)	Osthus (P)	Schrodinger (M)	Titian (M)
Johnson & Johnson (C)	Patcore (P)	SciencePoint (M)	UCB (C)
KNIME (M)	PerkinElmer (P)	Scitegrity (M)	UCL (M)
Lundbeck (C)	Pfizer (C)	Semtific Software (M)	University of Reading (M)

KEY:

Core Members (C) → 13

Participating Members (P) → 16

Contributing Members (M) → 41

70 organisations + individuals

© Pistoia Alliance

Examples of Success

HELM

Hierarchical Editing for Large Molecules

HELM
Hierarchical Editing Language for Macromolecules

Rapidly becoming established as the global standard notation for large molecules.

BEST PRACTICES AWARDS ★ Bio-IT World 2014

www.openhelm.org

CSCS

Controlled Substance Compliance Service

CSCS

Gaining traction across the industry as pharma companies and vendors alike seek to standardize their approach to compliance.

www.cscs-experts.org

tranSMART

tranSMART FOUNDATION

A widely adopted technology platform for managing and analysing translational science information, now managed by the tranSMART Foundation.

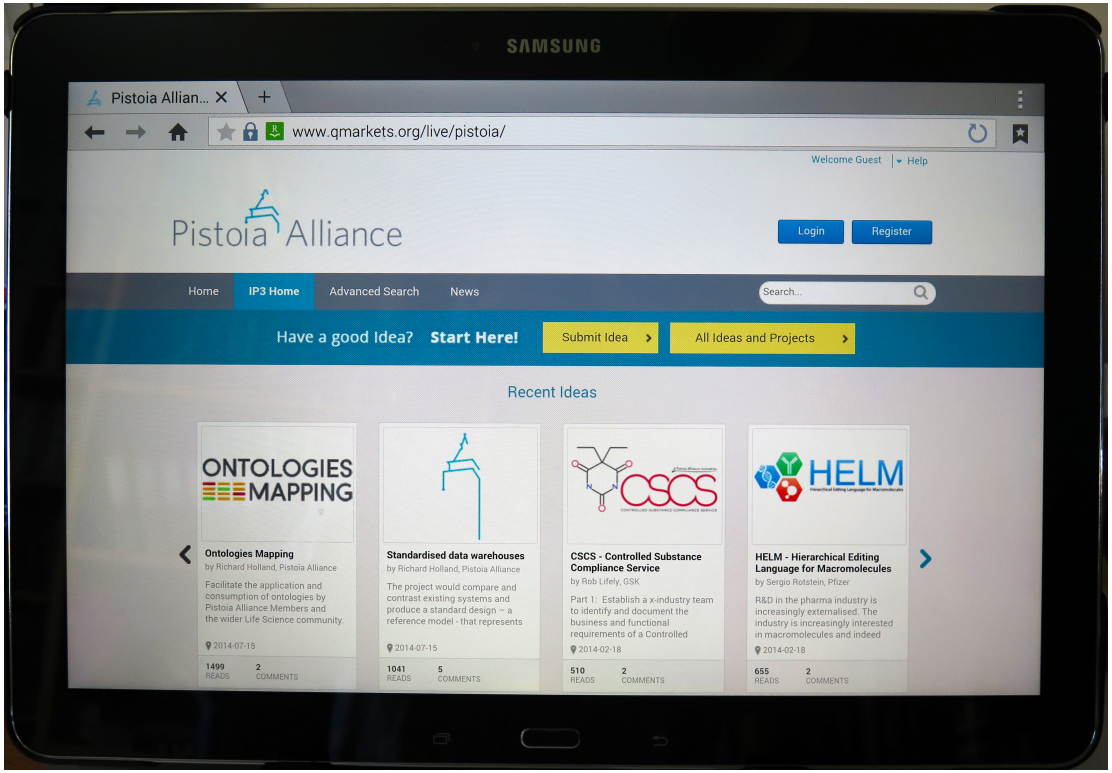
www.transmartfoundation.org

Pistoia-Alliance-Overview.pdf

© Pistoia Alliance

Interactive Project Portfolio Platform

- News ideas are submitted and discussed using the Interactive Project Portfolio Platform (IP3).
- Registration is free and is required should you wish to join in the conversation, add new ideas, or join an interest community:-



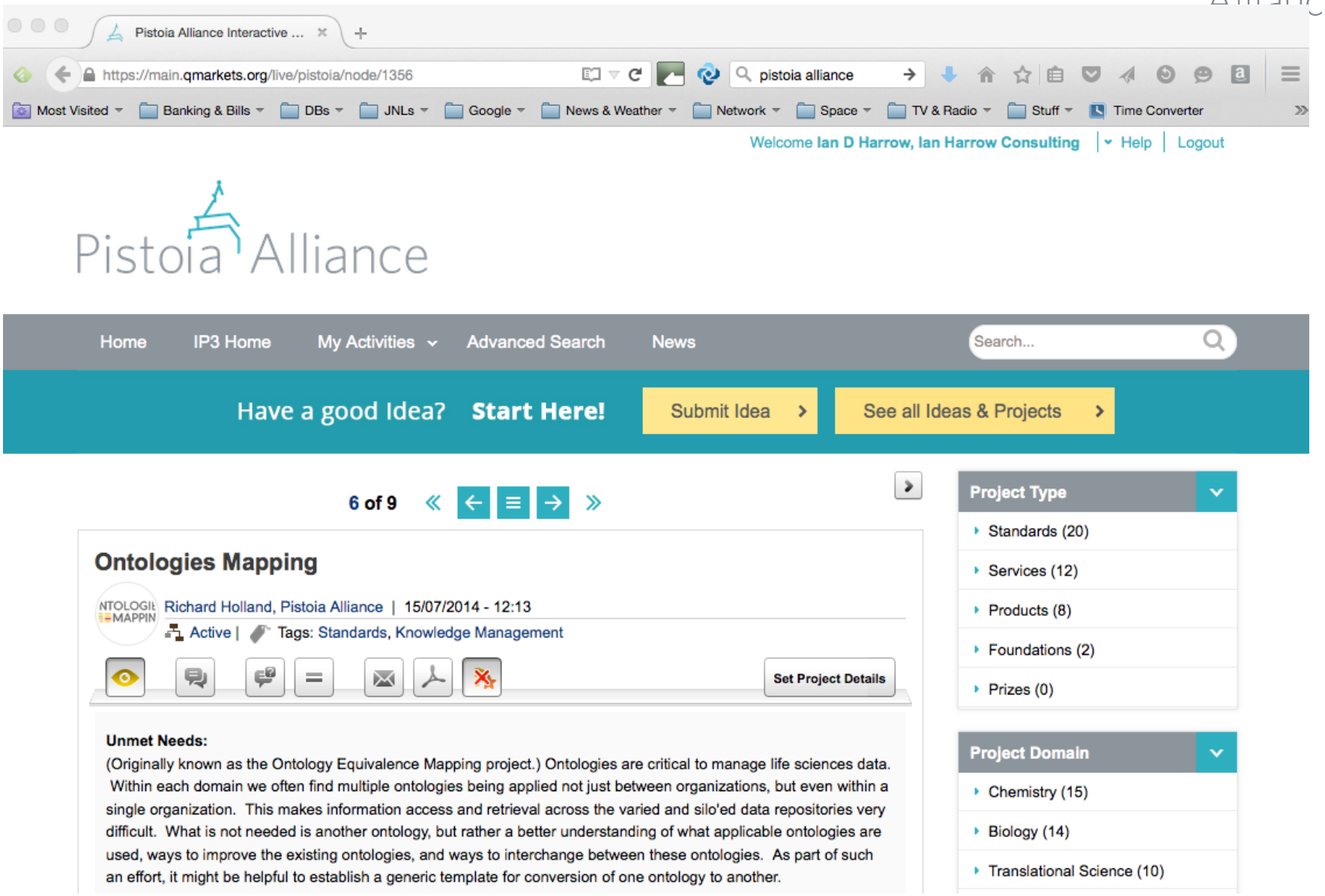
https://www.qmarkets.org/live/pistoia

© Pistoia Alliance

3. ONTOLOGIES MAPPING PROJECT

© Pistoia Alliance

Idea for Ontologies Mapping



© Pistoia Alliance

Business Case: Problem statement

Ontologies are critical to manage data in the life sciences.

Within each domain we often find multiple ontologies being applied not just between organizations, but even within a single organization.

This makes information access and retrieval across the varied and silo'ed data repositories very difficult.

What is NOT needed is another ontology, but rather…

1) a better understanding of what ontologies are available for particular domains.
2) best practise guidelines to improve existing ontologies.
3) a tool to map between existing ontologies to ensure better exploitation and a service for doing this.

© Pistoia Alliance

Business Case: Value Proposition

Facilitate the application and consumption of ontologies by Pistoia Alliance Members and the wider Life Science community.

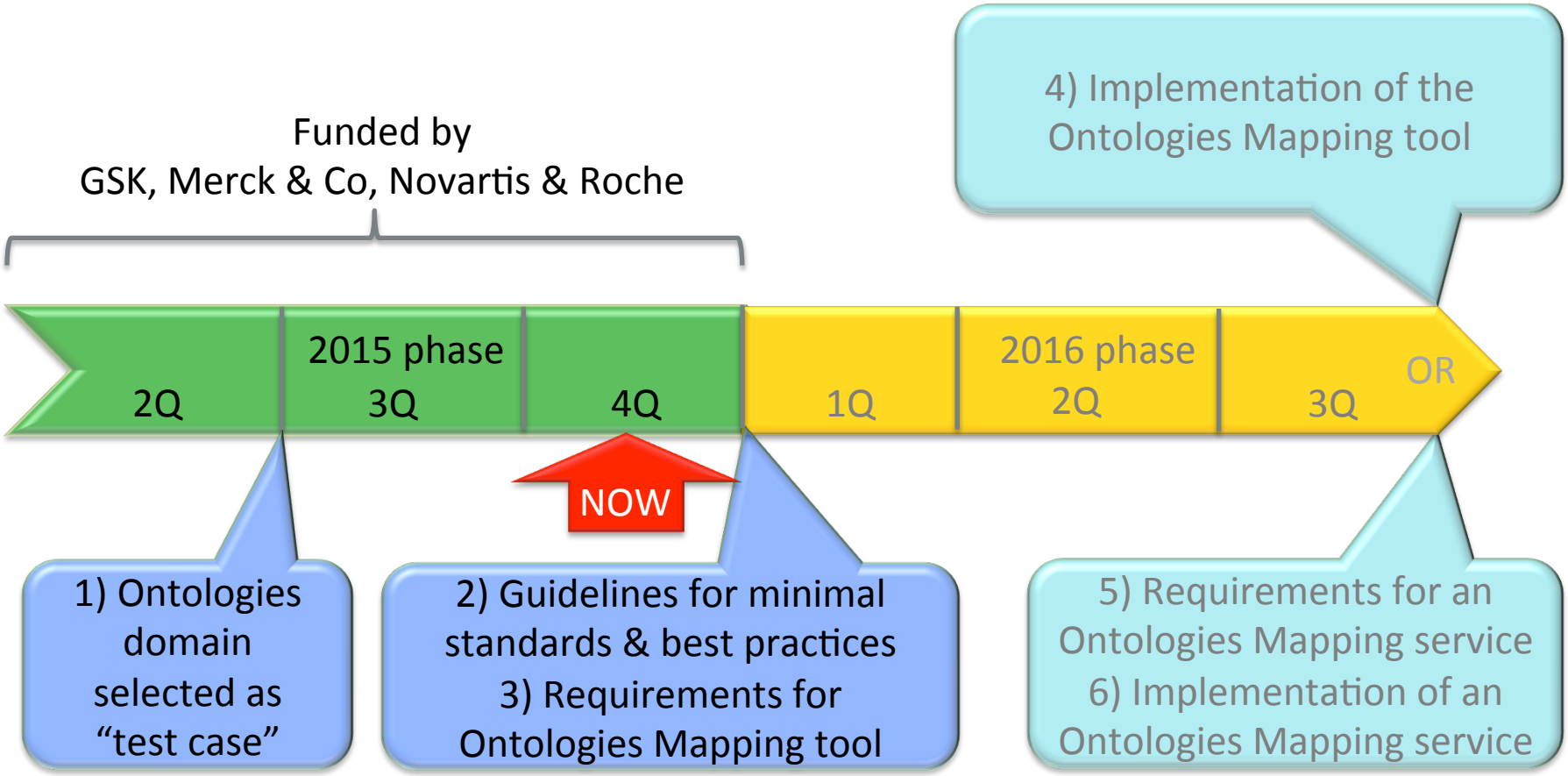
Distill a consensus-based opinion from Pistoia Alliance Members to influence and support providers of ontologies in commercial and academic organisations.

Establish a Project Team and Community of Interest to:-

- Select ontological domain(s) as "test case" based on member needs.
- Define guidelines for best practise to support application and mapping.
- Develop the requirements and implement an ontology mapping tool.
- Devise and create a sustainable ontology mapping service.

© Pistoia Alliance

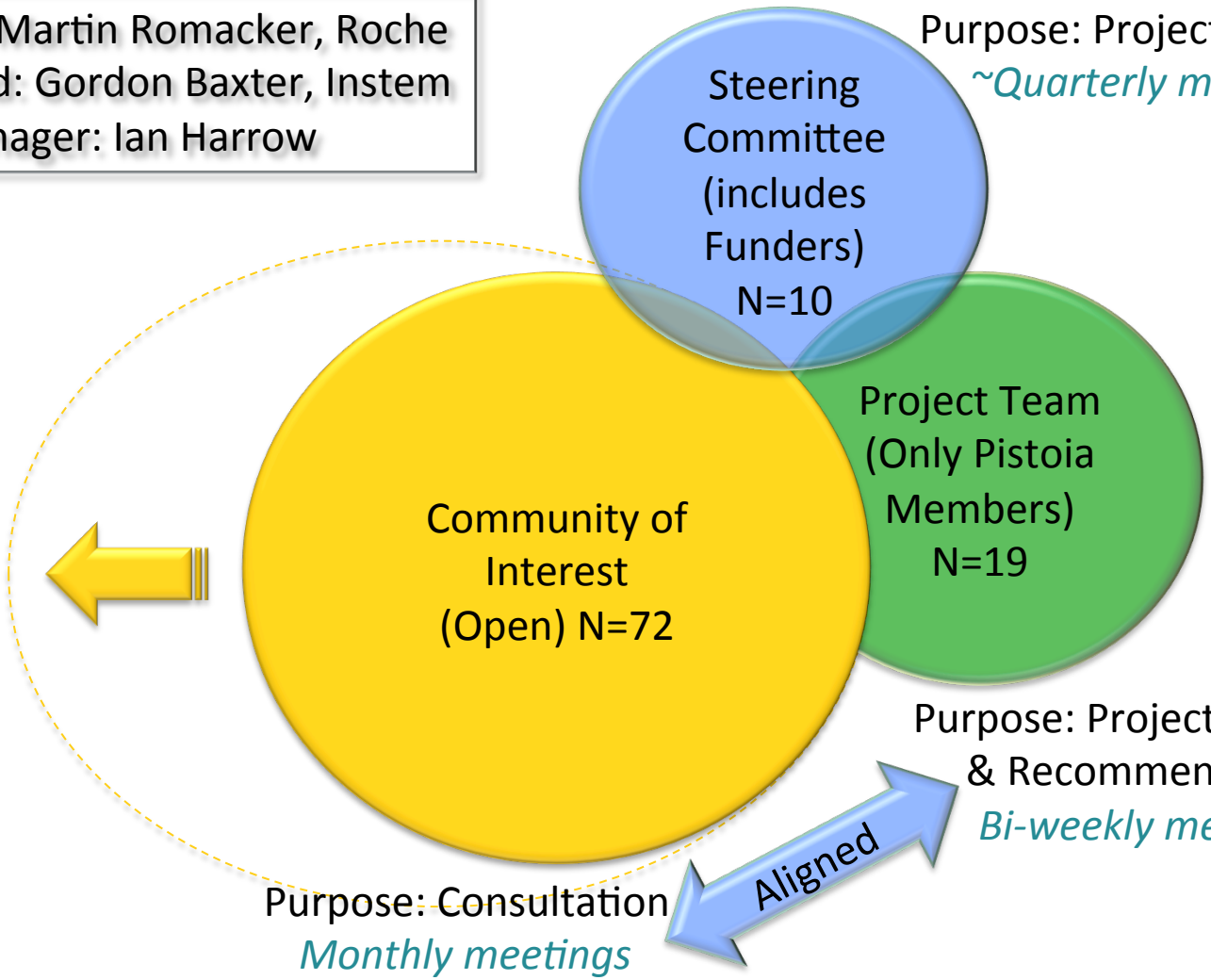
Project Timeline and Deliverables



© Pistoia Alliance

Project Structure and Communication

Champion: Martin Romacker, Roche
Project Lead: Gordon Baxter, Instem
Project Manager: Ian Harrow



© Pistoia Alliance

Progress

1) Survey to select "test cases"

Consultation with the Pistoia membership and Community of Interest through a formal survey.

Recommendations from analysis of the survey results:-

- Disease and Phenotype (single domain)
- Experimental Investigation (multiple domains)
- Recognise ontology types: reference, application and upper

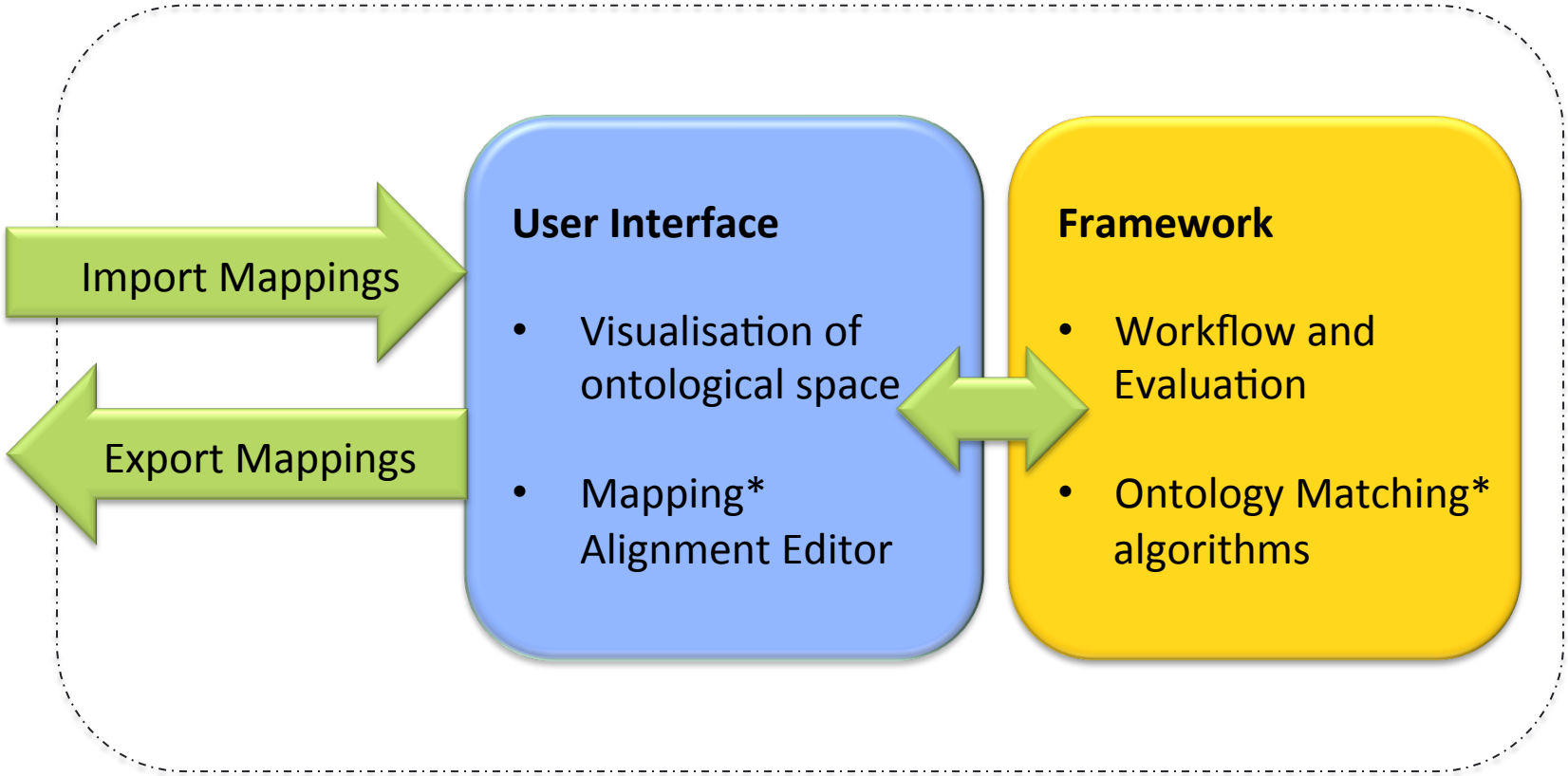
2) Draft Guidelines for Best Practise

- Guidelines for best practise to support application and mapping of ontologies have been drafted on project wiki.
- Serves as a useful checklist for selection of ontologies suitable for in a particular domain or application.

© Pistoia Alliance

Progress

3) Tool Requirements



Ontologies Mapping Tool

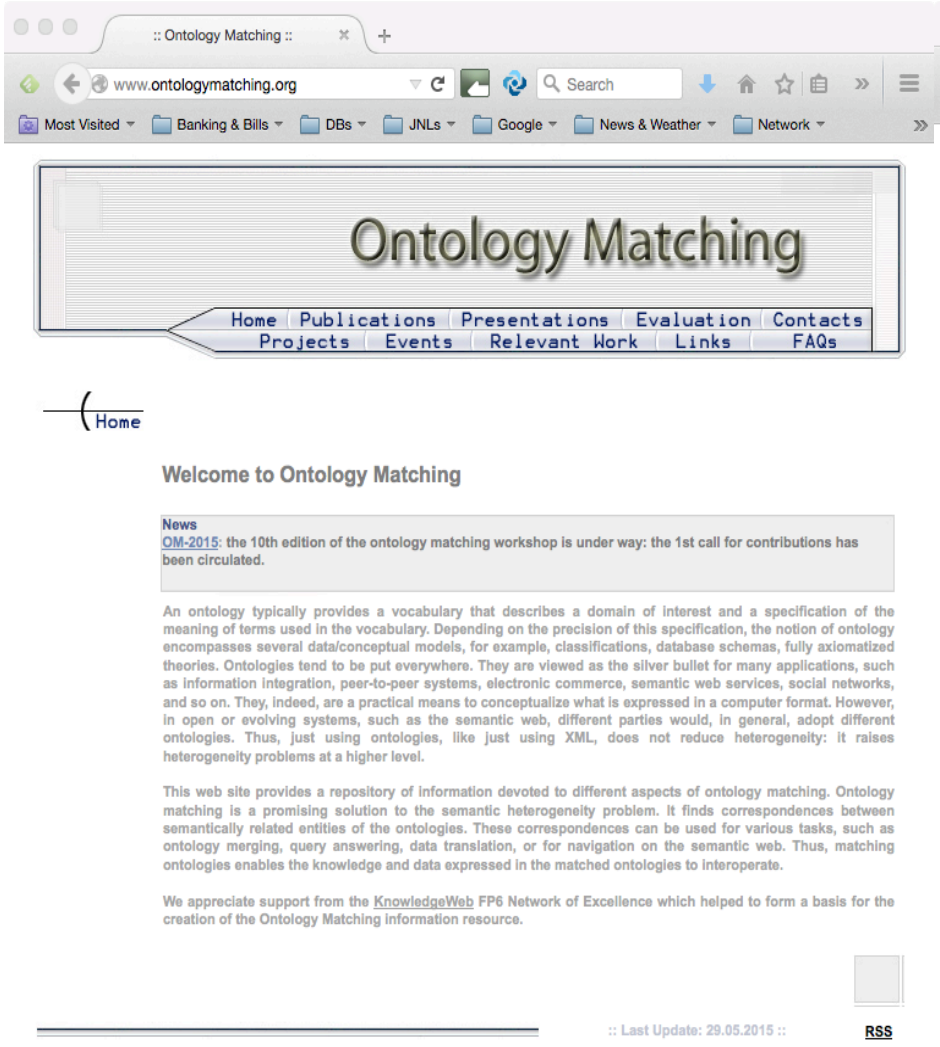
*Mapping = Matching

Requirements for the Tool are being drafted in detail.

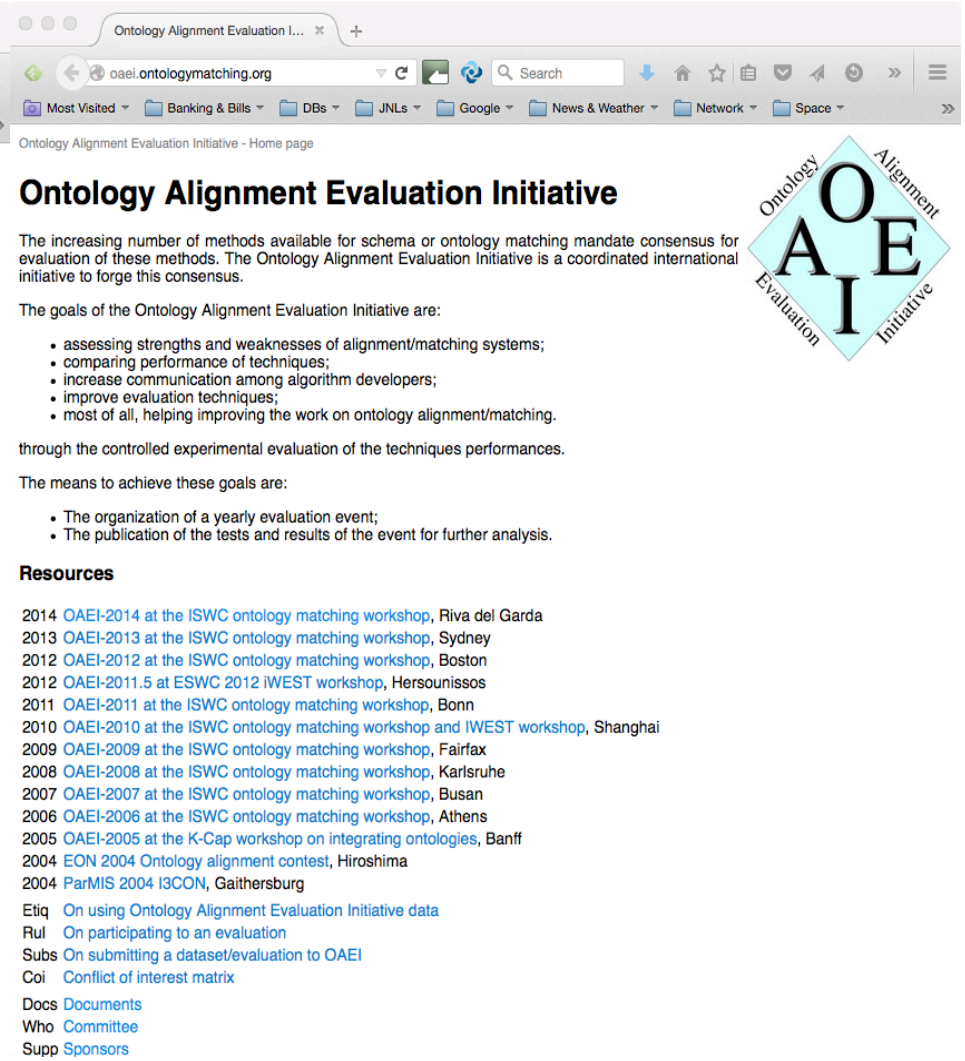
© Pistoia Alliance

Related External Activities

http://www.ontologymatching.org



http://oaei.ontologymatching.org



October 11th or 12th, 2015: Bethlehem, PA US http://www.om2015.ontologymatching.org

Community of Interest*

Organisation	Organisation
AstraZeneca	Ontology Matching
Bayer	OBO Foundry
EMBL-EBI	PRO Consortium
Jackson Labs	Biosharing/ISA/Uni. Oxford
NCBI	Linguamatics
Connected Discovery & SciBite	Friesland Campina
Thompson Reuters	Allotrope Foundation
RSC	Aridhia Informatics
Molecular Connections	Abbvie
Takeda	ELIXIR Interoperable Services
Teradata	+ more….

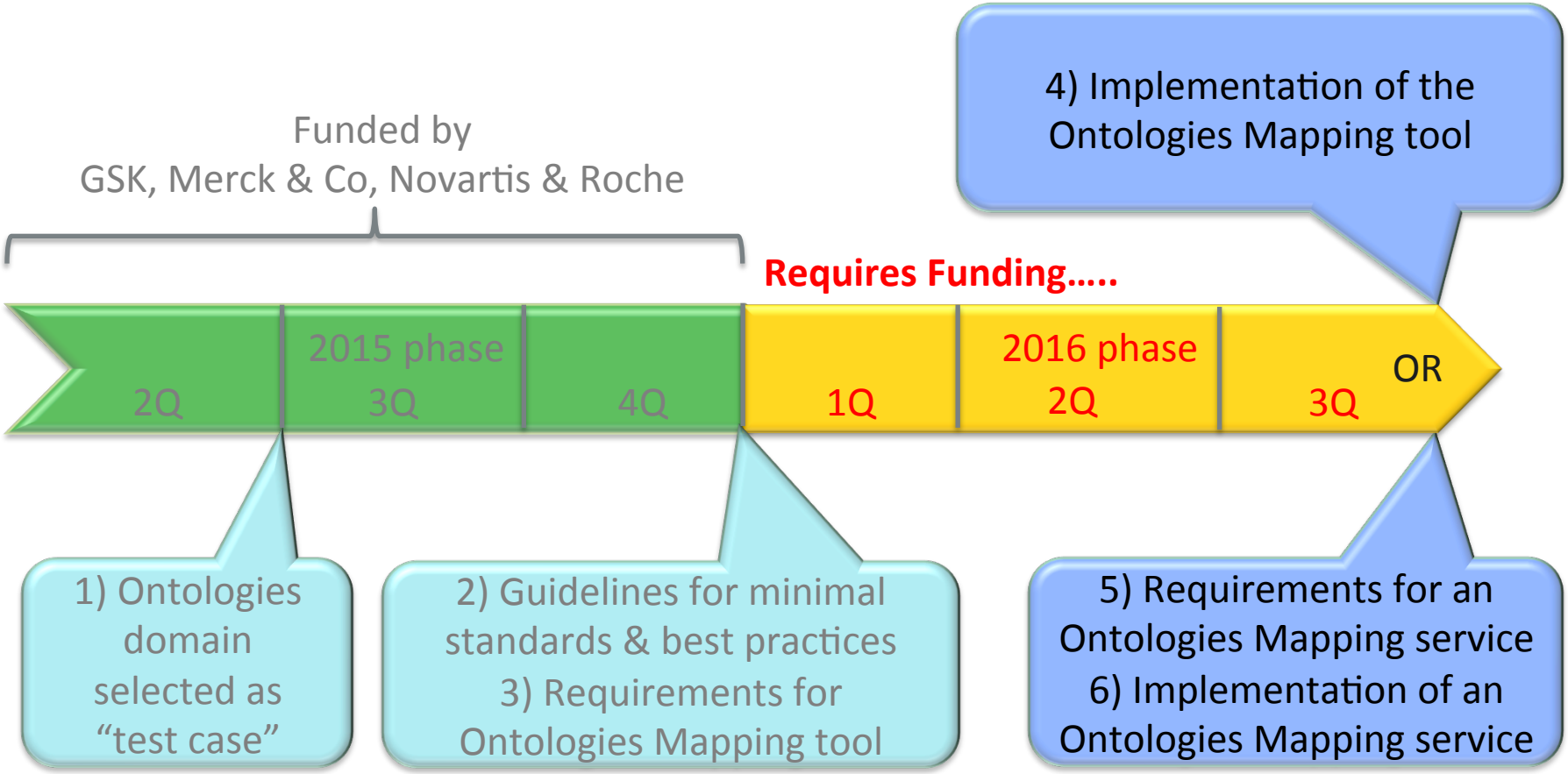
*Membership is open to any organisation with relevant skills and experience

Call for joining the Community of Interest

© Pistoia Alliance

Future Plans

Timeline: Next Phase



© Pistoia Alliance

Thank you for your Attention
Are there any questions?

ian.harrow@pistoiaalliance.org

@pistoiaalliance

www.pistoiaalliance.org