



# The Pistoia Alliance

*And its Role in the  
Virtualization of Big Pharma*

Michael Braxenthaler

17 October 2011

PRISME SIG, Cambridge, MA

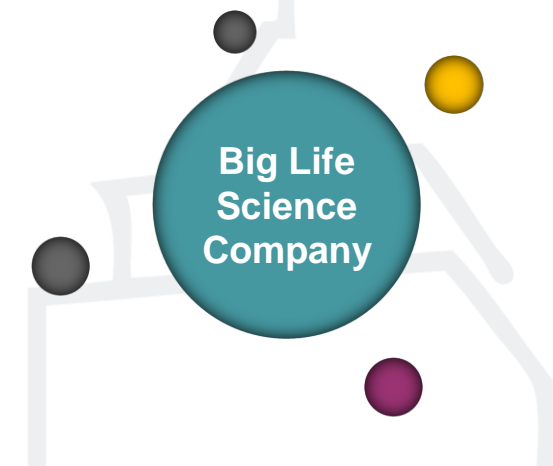
<http://pistoiaalliance.org>



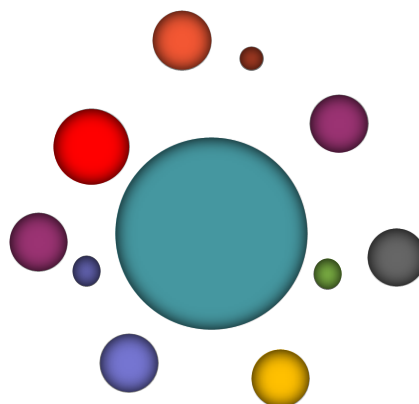
# Life Science Information Landscape

## A rapidly evolving ecosystem

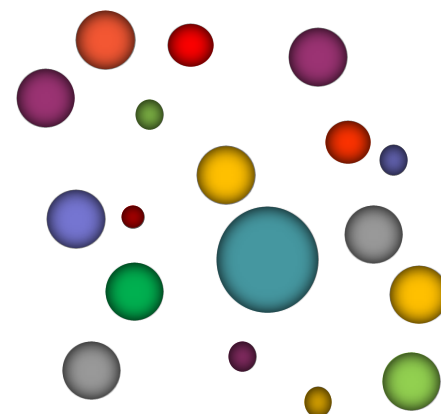
Yesterday



Today



Tomorrow



	Yesterday	Today	Tomorrow
Innovation Model	Innovation inside	Searching for Innovation	Heterogeneity of collaborations. Part of the wider ecosystem
IT	Internal apps & data	Struggling with change Security and Trust	Cloud/Services
Data	Mostly inside	In and Out	Distributed
Portfolio	Internally driven and owned	Partially shared	Shared portfolio

# The Pistoia Mission

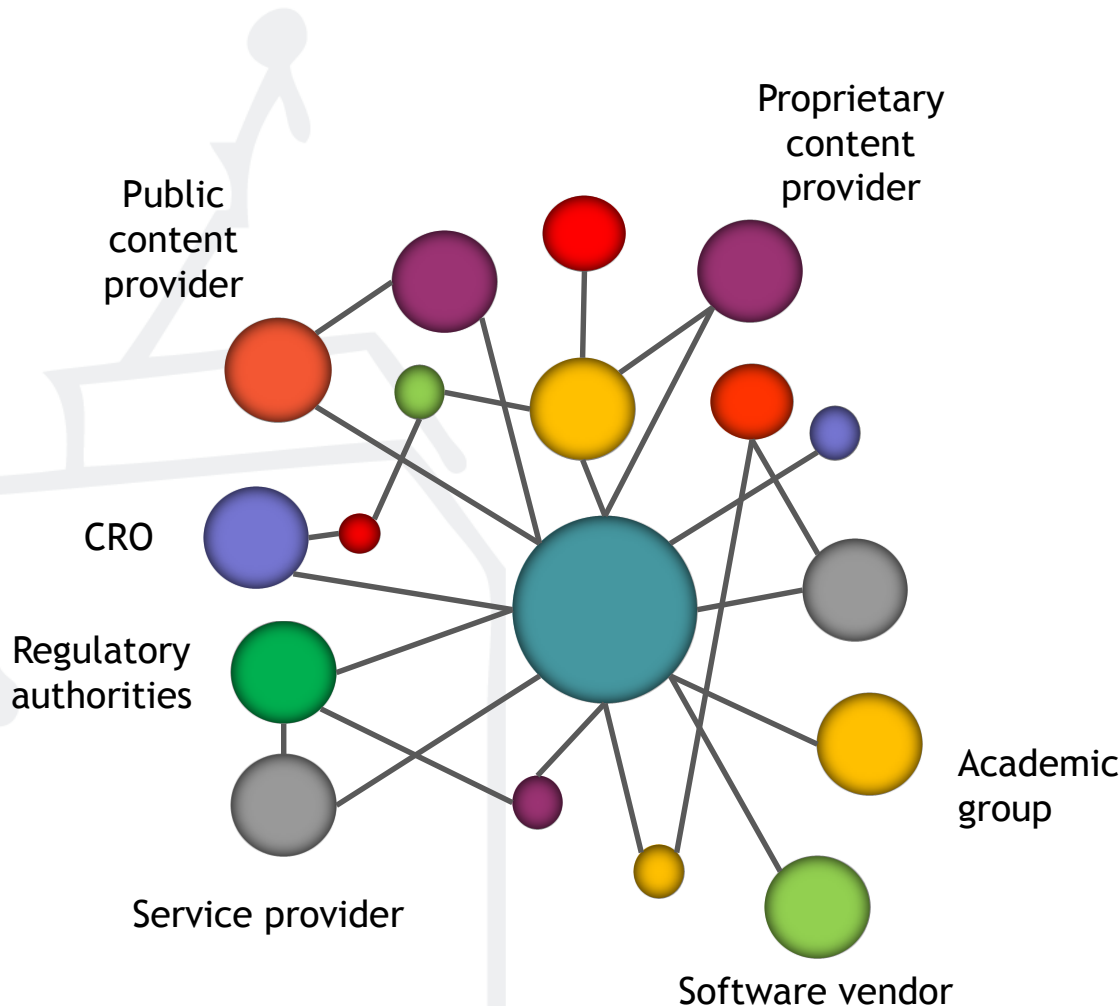


Lowering the barriers to  
innovation

by improving inter-operability of  
R&D business processes

through pre competitive  
collaboration

# The Role of Pistoia: Node or "Glue"?



- Pistoia is not a node in its own right
- Pistoia supports the connections between the nodes
- Pistoia's strength is its diversity of members, representing substantial parts of this ecosystem

# Pistoia Alliance Membership



# Pistoia Governance

- Board of Directors

## Core

Ingrid Akerblom  
Christoph Brockel  
Ramesh Durvasula  
Tom Flores  
Claus Stie Kallesøe  
Bryn Roberts  
Jason Swift

Merck  
Pfizer  
BMS  
GSK  
Lundbeck  
Roche  
AstraZeneca

## Participating

Frank Brown  
Joe Donahue  
Alex Drijver  
Sean Ekins  
Alexander Scheer  
Michael Stapleton

Accelrys  
Thomson Reuters  
ChemAxon  
CDD  
Merck Serono  
Perkin Elmer

- Operational Team

Michael Braxenthaler  
Ashley George  
John Wise  
Nick Lynch  
Deborah Ausman  
Martyn Wilkins  
Vacant

Roche  
GlaxoSmithKline  
Independent  
AstraZeneca  
Independent  
Independent

President  
Treasurer  
Executive Director  
External Liaison  
Communications  
Secretary  
Technical Committee Chair

Supported by Working Group Chairs and Project Teams

# Role of Pistoia

## Risk- and Cost-Sharing

- Standards: Landscape, gaps, sign-posting, agile standards development
  - Example: VSI/Oxford collaboration (S. Sansone, Biocatalog)
- Fostering shared services
  - Example: Sequence Services Project
- Open innovation
  - Example: Sequence Squeeze Competition
- Targeted support of academic grant proposals
- Pre-competitive collaboration on curation of public content
- Developing shared use cases
  - Example: “Disease Explorer”

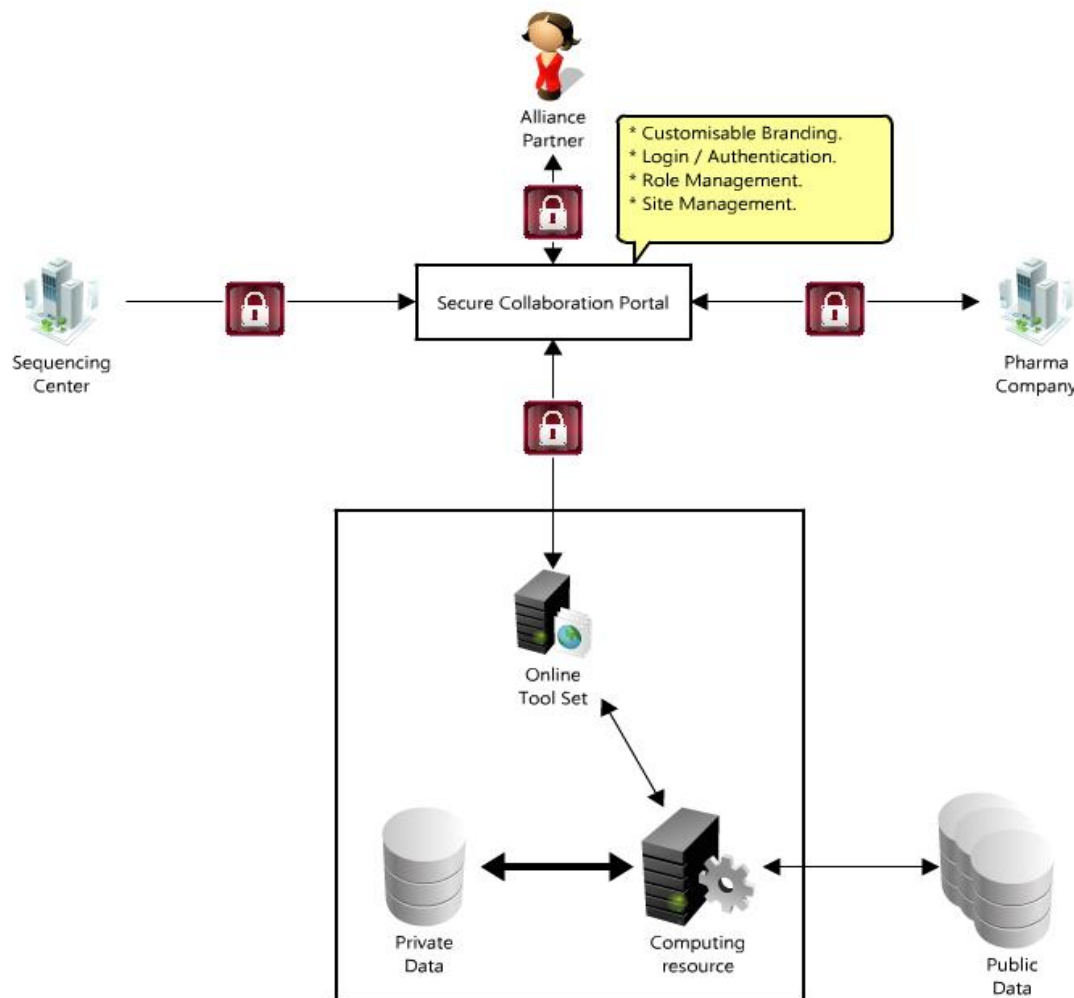
# Role of Pistoia

## Enabling and Fostering Connections

- Enabler of more 1:1 or 1:2 conversations
  - Informal networking
  - Sharing of experience and ideas
    - Just start them - no formal approach needed
    - Potential start of bigger projects
- Pistoia Workshops
  - Start: Life Science Information Ecosystem Workshop
  - Keep growing membership connected and engaged
  - Repeat annually
  - Spin-off of projects and specific workgroups
- Pistoia Clinics
  - Open sessions on a chosen area to share ideas and current challenges
  - Platform to gather pains and ideas



# Shared Services: Sequence Services Phase 2



- NGS-ready Service
- Collaboration
- Security
- Integrated public & proprietary data
- Public toolset for analysis
- Commercial and proprietary tools

# Sequence Services

Project lead: Simon Thornber GSK

## Status:

- Phase 1 achieved
- Phase 2 commenced

## Objective

- Define standards for the provision of secure access to pre-competitive, shared, hosted, gene sequence databases and software tools.

## Timeline & Funding

- £200,000
- end April 2012

## Deliverables

- Working implementations demonstrated at the Pistoia Alliance User Conference April 2012

## Next steps

- Phase 2 launched.
- RFP published with enhanced user requirements including addressing NGS data.
- 10 RFP responses received and presented last week
- Phase 2 PoC demonstrations at PA User Conference - April 2012

# Sequence Services RFP respondents

- Happiest Minds / Bionivid
- Cognizant / Accelrys
- Constellation / Microsoft
- Eagle Genomics / Cycle Computing
- GenomeQuest
- Genestack
- HP/ Zurich Genome Center
- Infosys
- IDBS
- TheHyve

# Open Innovation: Sequence Squeeze

Project lead: Simon Thornber GSK

Status:

- On track

## Objective

- Find a new algorithm/technique for compressing next-generation sequencing data for use in the Sequence Services project

## Timeline & Funding

- £12,000 + VAT @ 20% to manage the prize
- \$15,000 prize money
- April 2011 (winning solution ready for announcement at the Pistoia Alliance Users Conference)

## Deliverables

- Competition website: <http://www.sequencesqueeze.org/>
- Established a judging panel
- State-of-the-art gene sequence data compression tool

## Next steps

- Project launch - start Q4 2011
- Entries close end Q1 2012
- Winners to be announced at the Pistoia Alliance Annual Conference in April 2012.

# Use Case “Disease Explorer”

There is a need to be able to “peel away” layers of information around a disease. For example:

- For a given indication, what are the higher-level pathophysiological processes?
- For each process, what are the cells, tissues and events within?
- To what elements of the disease does each process contribute (initiation, progression, exacerbation etc.)
- How does this then break down into molecular pathways?
- Where are current therapies targetted?
- Are we all targeting the same few pathways in the same process?
- Are there processes that are completely untested?
- How does the literature map onto this?
- Where are the new emerging areas?
- Where are the overlaps between diseases? Aan obvious example being the repurposing of viagra due to the enzyme, PDE5 s control of cyclic GMP, and that cGMP is involved in sleeping, diabetes etc.)

# Use Case “Disease Explorer” (cont’d)

## Conclusion:

- We should have information systems able to explore disease through different levels of resolution and map databases accordingly.
- This requires a functioning information eco-system, where the data is mapped to standards that facilitate this exploration and there is a wide body of developers producing elements of this "browser".



# The Pistoia Alliance

## *Information Ecosystem Workshop*

10 October 2011  
Hannover, Germany

<http://pistoiaalliance.org>



# Information Ecosystem Workshop: Participants

## 38 Attendees

- 6 Life Science companies
- Academic groups
- Technology companies
- Scientific software vendors
- Scientific service providers
- Commercial content providers

Focus on work in breakout groups –  
interactive / collaborative  
approach





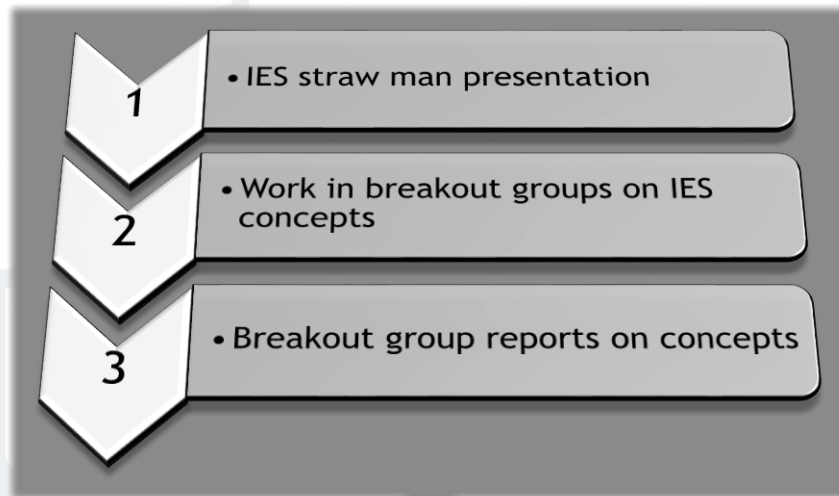
# Pistoia Information Ecosystem Workshop

## Agenda and Deliverables Outline

Morning session:

### **What is the IES? What should it be?**

Develop common understanding on what the IES should deliver

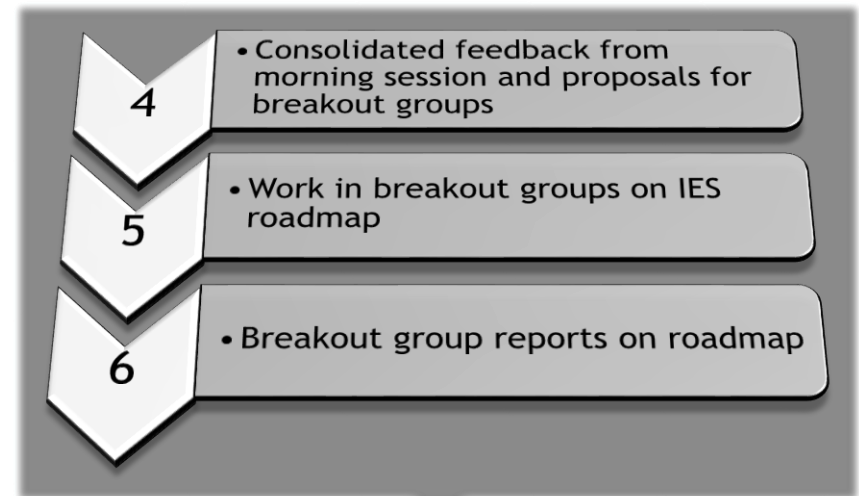


**White Paper:**  
The Key Concepts  
of the Information Ecosystem

Afternoon session:

### **How do we get there?**

Develop a roadmap of projects and activities



**IES Roadmap:**  
Projects and  
Follow-up activities

# Morning Session - Breakout Groups

## Overarching theme:

### Develop a shared understanding of the problem space:

What does 'information ecosystem' mean?

What are the current shortcomings?

What should it look like to help us solve our problems?

- Workgroup 1: The industry/academic collaboration space
- Workgroup 2: Information exchange with CROs - discovery/chemistry
- Workgroup 3: Information exchange with CROs - biology/NCD
- Workgroup 4: Integration of public and proprietary content
- Workgroup 5: The Standards Landscape

# Afternoon Session - Breakout Groups

## Overarching theme:

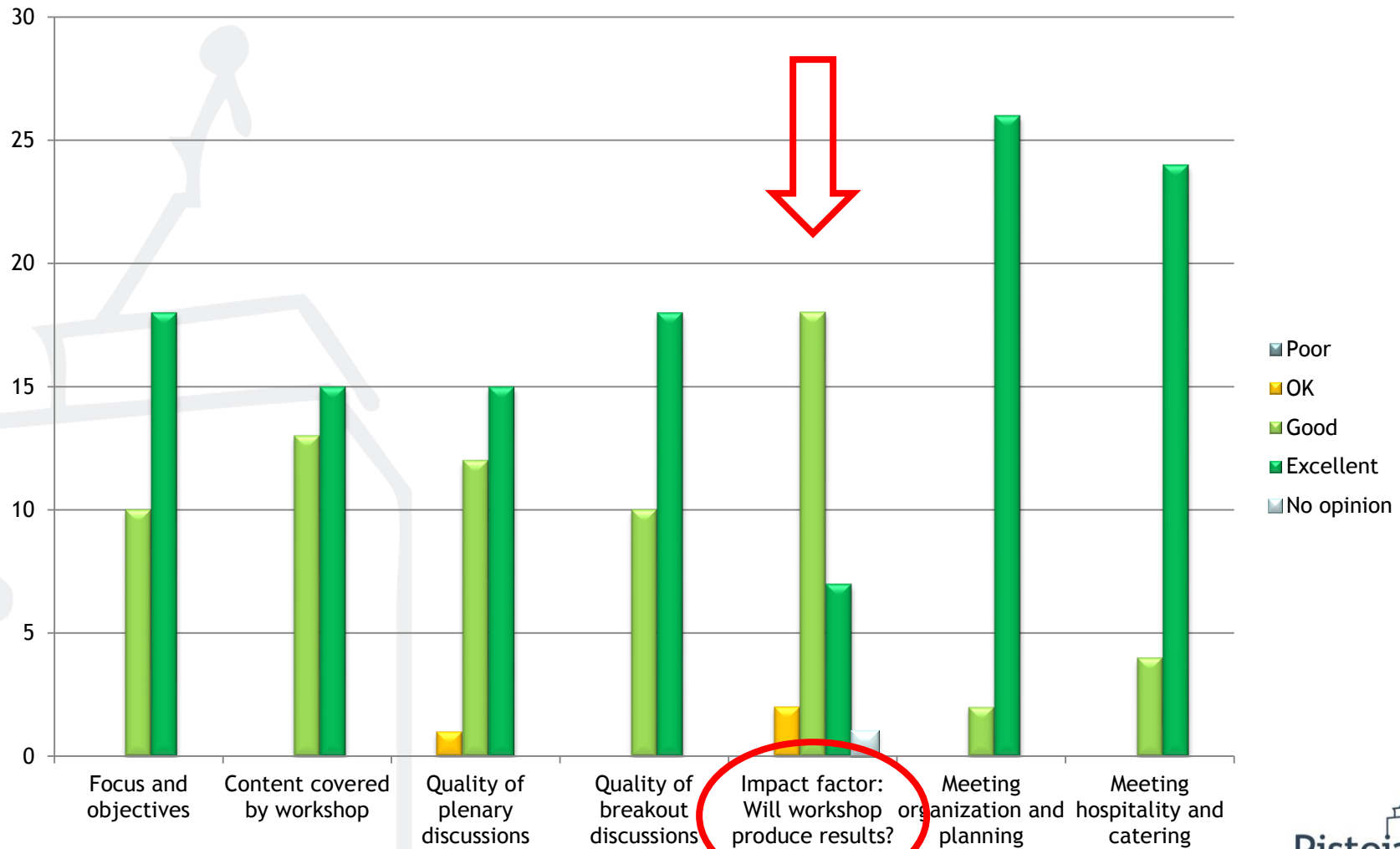
## Develop solution outlines to address the current problems:

Proposals for short term projects and mid term activities

Looking beyond the scientific-technical horizon: Economic feasibility and viable business models

- Workgroups 1 and 2: Low-hanging fruit - What can we tackle right away within a 1-year time frame?
- Workgroup 3: Continuing focus areas and mid-term activities - How do we approach more complex problems?
- Workgroup 4: The Standards Landscape - What needs to be done?
- Workgroup 5: Business Models - Rethinking the game

# Feedback at the workshop...



# Why this workshop?





# Why this workshop?



# Thank you!

Questions





Backup Slides

Pistoia Alliance

# INFORMATION ECOSYSTEM WORKSHOP





Pistoia Alliance  
Information Ecosystem Workshop

# MORNING SESSION

# AM Workgroup 1

## The industry/academic collaboration space

- Increasing number of bi- and multilateral collaborations between life science companies and academic groups.
- Need for collaboration space
  - Not just sharing and exchange of documents and data
  - Secure, 'science-aware' collaboration environment
  - Allows collaborative data analysis and discussion
  - Easy and fast way to spin up and turn off instances

What would such a space ideally look like?

What are the minimal and optimal requirements for such a space?

# AM Workgroup 2

## Information exchange with CROs - discovery/chemistry

- See use case “Research Externalization on a Collaborative Framework”

# Use Case “Research Externalization on a Collaborative Framework”

## Introduction:

Declining R&D productivity, rising costs of commercialization & shorter exclusivity period have driven up the average cost of launching a successful new drug to US \$ 1.7 billion. To facilitate drug development & to lower the cost & risk of launching new drugs on their own, pharmaceutical companies have increasingly turned to alliances for outsourcing.

## Actors:

Pharmaceutical Company, Scientist, Regulatory body, Discovery CRO, Chemistry service providers, Biology service provider, assay development, screening, lead optimization & other vendors.

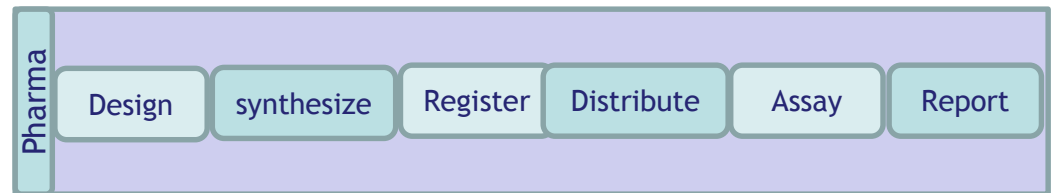
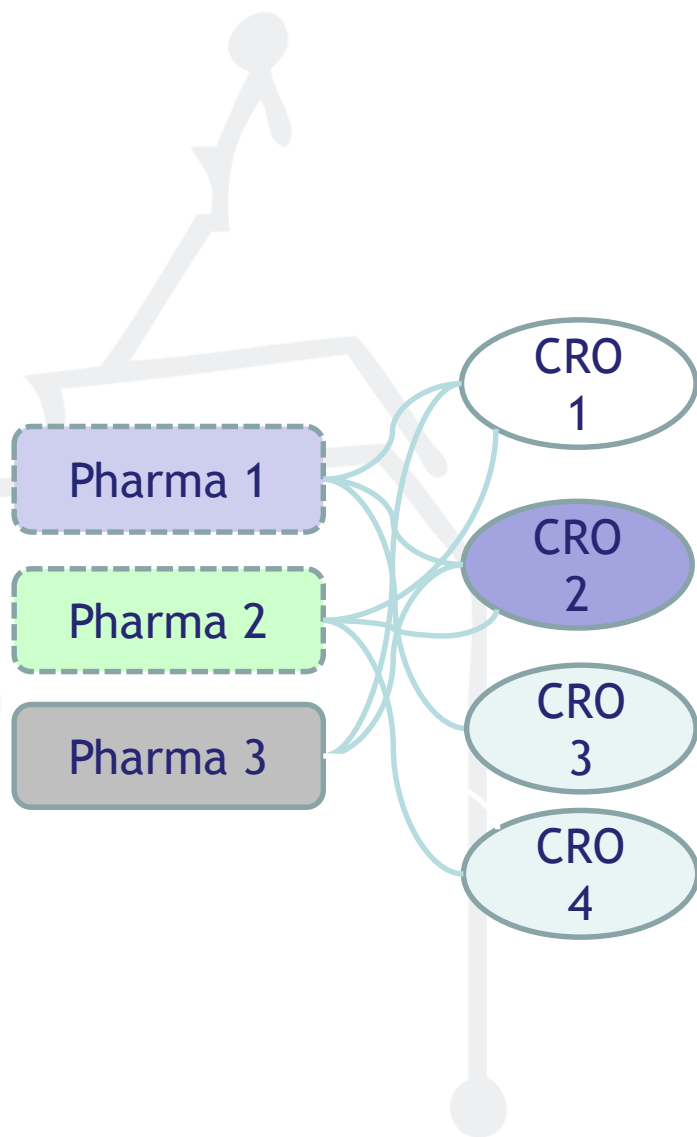
# Use Case “Research Externalization on a Collaborative Framework” (cont’d - 1)

## Scenario:

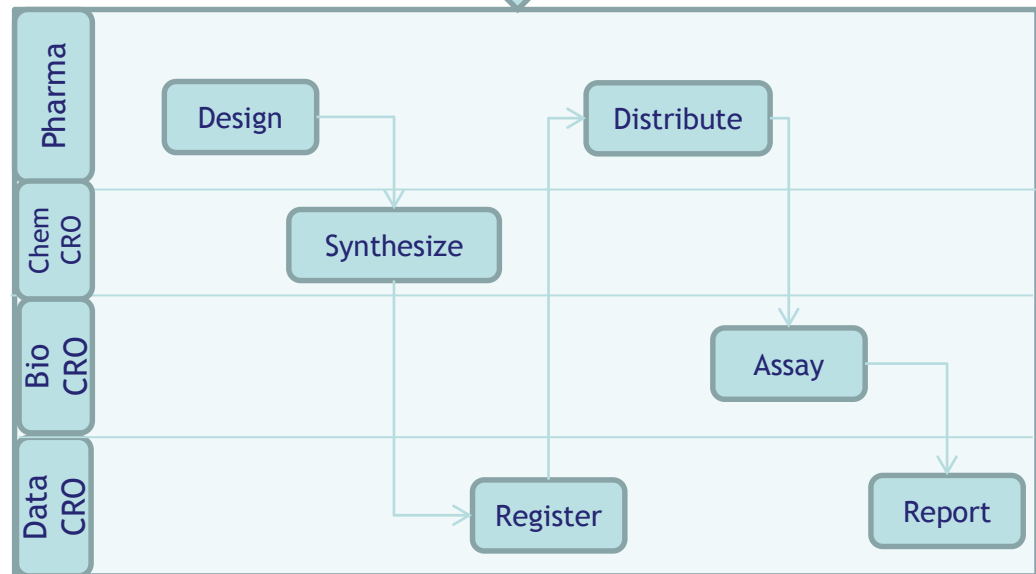
Target molecule has been discovered & validated by pharmaceutical company’s internal R&D department. Lead identification requires expertise, specialized skills, and heavy investments on resources/technology. There already exist vendors in the market with resources/technology/expertise in dedicated areas. Hence the pharmaceutical company decides to externalize the different components of lead identification process to multiple capable vendors, thereby resulting in reduced cost, time and overcoming lack of in-house expertise/resources.

The diagram below depicts a scenario indicating a transition from a siloed/fully internal approach to a collaborative model which enables multiple pharma’s to externalize their research activities to multiple organizations in a secured manner.

# Use Case “Research Externalization on a Collaborative Framework” (cont’d - 2)



Fully Internal Model



Selectively Integrated Model

# Use Case “Research Externalization on a Collaborative Framework” (cont’d - 3)

**The current industry challenges to execute the above scenario on a collaborative framework are:**

- Secured real time information exchange
- IP Protection & infringement
- Process & Data harmonization
- Lack of standard data formats
- Lack of regulatory guidelines around collaboration & networking

## **Summary:**

This collaborative model would not only drive seamless research externalization, but also enable access to information/experts, real time learning's, reuse/sharing of internal assets, increase research productivity, focused approach, effective process tracking, shared workspace for secured communication and risk mitigation.

# AM Workgroup 3

## Information exchange with CROs - biology/NCD

- Currently, data exchange between life science companies and CROs requires substantial manual involvement
  - data formatting issues
  - meta data and context for interpretation
- Many areas in the R&D process affected
  - Genotyping, NGS, Proteomics, Metabolomics, Preclinical Safety - Animal Studies

What would the ideal exchange environment look like?

- Are data and meta data standards sufficient?
- Or would this environment benefit from a collaboration space allowing collaborative data access for QC and interpretation?



# AM Workgroup 4

## Integration of public and proprietary content

- Increasingly large volumes of relevant data for life science R&D are available in the public domain.
- Tendency increasing in volume, complexity, and geographic/organizational distribution
- Conventional approach to bring all relevant data in-house for integration and analysis will not be feasible in the future

What are sustainable models for environments which enable integrated analysis and interpretation of distributed complex data?

See also use case “Disease Explorer”

# AM Workgroup 5

## The Standards Landscape

- Many standards exist or are being developed in the life science information space
- At the same time there are many calls for standards to enable information exchange/sharing/integration

What does the current standards landscape look like?

What are the key issues with the current standard landscape?

Too many? (where? - sign-posting)

Missing? (where?)

Problems with agreement/adoption?



Pistoia Alliance  
Information Ecosystem Workshop

# AFTERNOON SESSION

# Afternoon Session

- Summary of the outcome from the morning session
- More on the intended outcome of the workshop
- Afternoon breakout groups
- Feedback from breakout groups
- Summary and concluding remarks
- ‘Thank you’ reception

# Key Messages from Morning Session

- Identified many opportunities
  - Short term
  - Mid term
- Question will be: How best to source activities?

# Key Messages from Morning Session

- Possible additional roles of Pistoia
  - Certification (of standards?, compliance with standards?)
  - Signposting
  - Call for “Gorilla in the Room”
  - Catalyst for IT/legal involvement
  - Guide and support academic funding proposals
  - Pre-comp content curation