



ELIXIR and Industry

Jerome Wojcik, CEO, Quartz Bio

18 December 2013

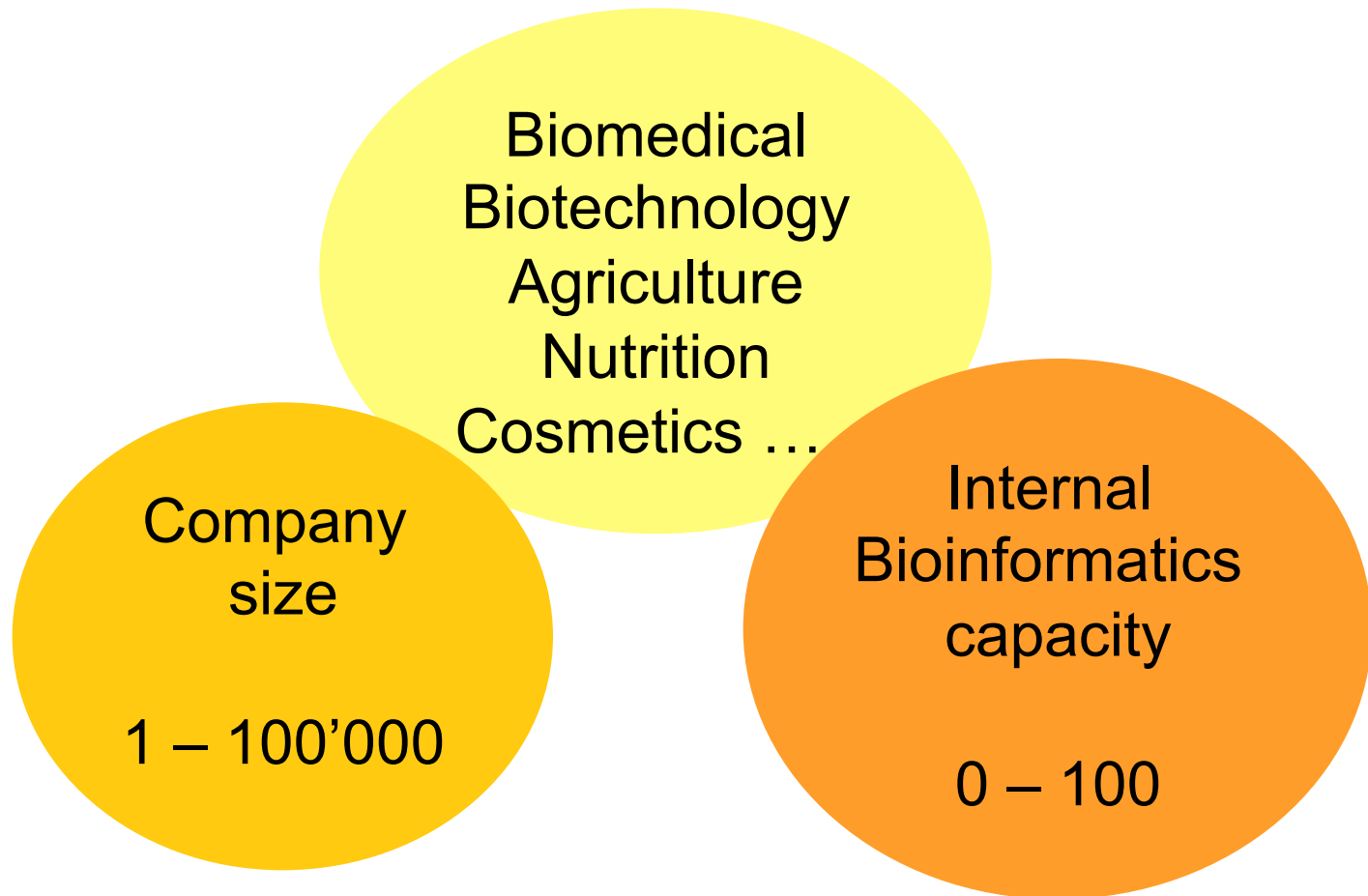


European Life Sciences Infrastructure for Biological Information
www.elixir-europe.org

Outline

- Why ELIXIR is important for the Industry?
- How the Industry can contribute to ELIXIR?

Life Sciences Industry Landscape



A Career Path of a Industry Bioinformatician ...

- 1999: Ph.D. in Biology
- 2000-2005: Director of IT and Bioinformatics in a **Biotech** company
- 2005-2012: Director of Bioinformatics in a **Pharma** company
- 2012-: CEO and Founder of a clinical bioinformatics **SME**



... and its Dependency on Public Infrastructure

HYBRiGENiCS

Proteomics

- Definition of a protein-protein interaction **standards**



Genomics

- Use of public **databases**
- Collaboration with **HPC** centres



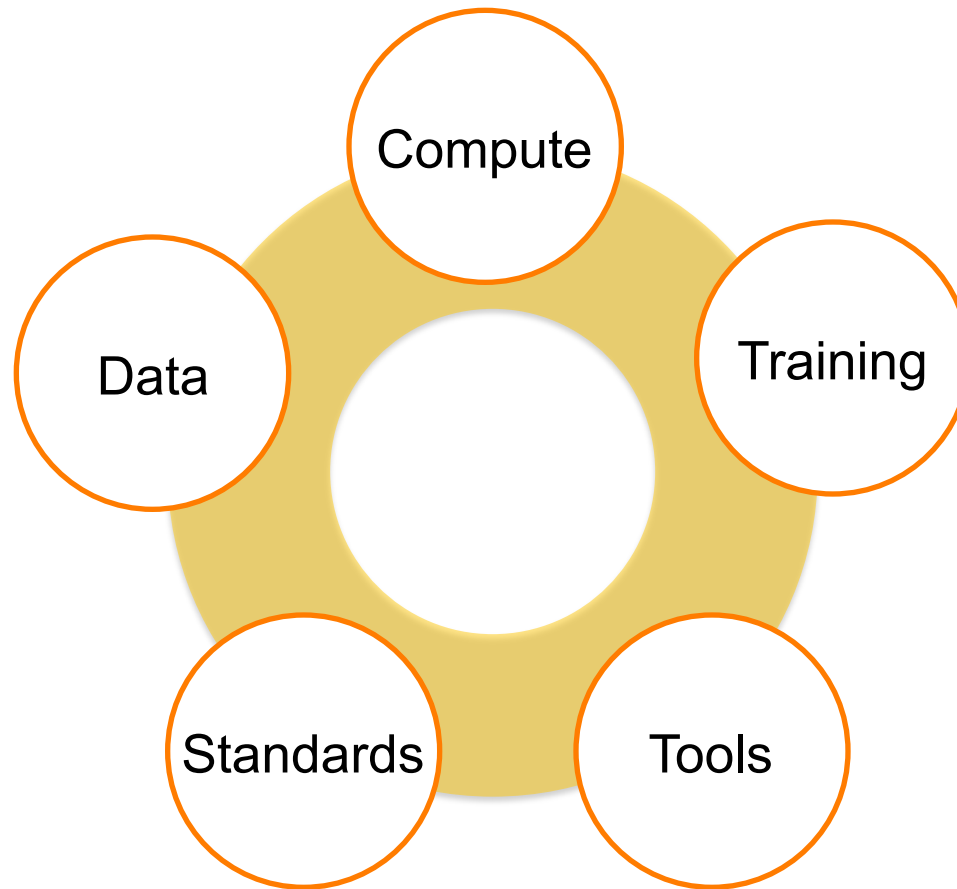
Quartz Bio

Precision Medicine

- Contributor to open-source **tools**
- Actively **recruiting** young talents



Industry as a Consumer of Bioinformatics



Why ELIXIR matters for Industry Consumers?

- Life science companies usually have limited internal R&D bioinformatics resources. They have to rely on external expert resources.

ELIXIR is a network of expertise

Quality

- External dependencies are a risk for industry if they are not stable. Maintenance over time of a high-quality relationship is key.

ELIXIR has the critical mass

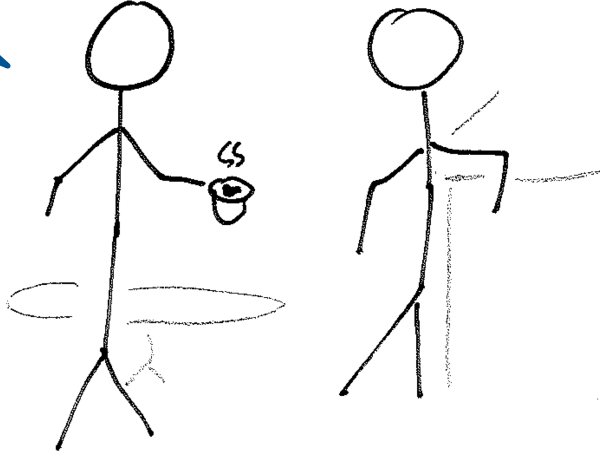
Sustainability

- Industry prefers smooth and light administrative processes.

ELIXIR is the unique European contact point

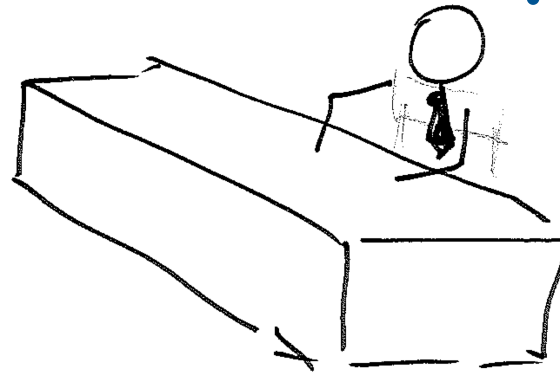
Efficiency

I need
bioinformatics

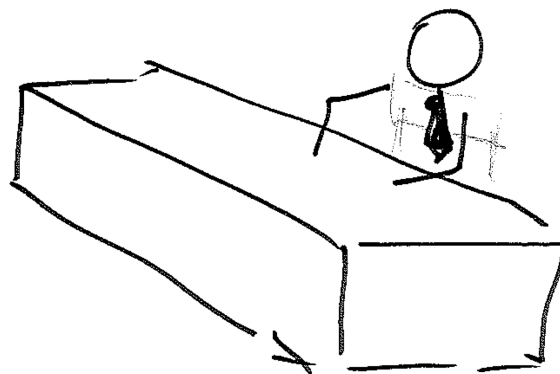


ELIXIR

I plan to pull the mineR package from github to parse the dbOmic database uploaded on our sFTP server by our collaborator (...)



I plan to liaise with
ELIXIR



Ok!

Only a Consumer?

- Industry publishes

A collage of scientific publications from various journals, including Human Heredity, BMC Proceedings, BIOINFORMATICS, Annals of human genetics, nature genetics, nature medicine, and ARD. The publications cover topics such as genome-wide association studies, false discovery rate estimation, gene-based bin analysis, epistasis detection, and genetic interactions.

Human Heredity
A Fast, Unbiased and Efficient Case-Control Association Test

BMC Proceedings
Universal False Discovery Rate Estimation Methodology for Genome-Wide Association Studies

BIOINFORMATICS
Proceedings
Gene-based bin analysis of genome-wide association studies

Annals of human genetics
ExactFDR: exact computation of false discovery rate estimate in case-control association

nature genetics
Detecting Epistasis with Restricted Response Patterns in Pairs of Biallelic Loci

nature medicine
Functional variants in the B-cell gene *BANK1* are associated with systemic lupus erythematosus

ARD
Genetic and physical interaction of the B-cell systemic lupus erythematosus-associated genes *BANK1* and *BLK*

nature medicine
A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer

- Industry releases tools to the public domain and sponsors foundations that maintain these tools



- Development of the Open-source mindset in the Industry
- Up to Open Data?

My 2 most important Current Needs

2. Data management

1. Man power

Data Management: Key and Time-consuming

- Personal experience: 40% of the workload in a fee-for-service project is spent on biomarker data management, *i.e.* transferring, checking, integrating, standardising data
- This step is key as the quality of the input data is the first limiting factor of the quality of the overall analysis
- The use of unanimously accepted **Standards** will dramatically speed-up the data management process



Use Case: CDISC for Clinical Data

- In July 2004, FDA Commissioner L.M. Crawford announced the desire of the FDA to receive data in a **standard** format, the CDISC SDTM



"The importance of a standard for the exchange of clinical trial data cannot be overstated," said Dr. Crawford, "FDA reviewers spend far too much valuable time simply reorganizing large amounts of data submitted in varying formats. Having the data presented in a standard structure will improve FDA's ability to evaluate the data and help speed new discoveries to the public."

- The CDISC format is now widely adopted by the Pharma Industry



ELIXIR can set the Bioinformatics standards that will save time to all analysts



What is the Value of Bioinformatics?

- Industry aims at creating Value
- Where does the value come from in bioinformatics?
 - Structured **Databases, Tools, Standards,** and **Computers** are facilitators and accelerators. They provide Value by saving time and increasing Quality.
 - **People** are from far the most important. Their expertise and experience, and the time they spent exploring data, are instrumental!

Life Sciences need more trained bioinformaticians

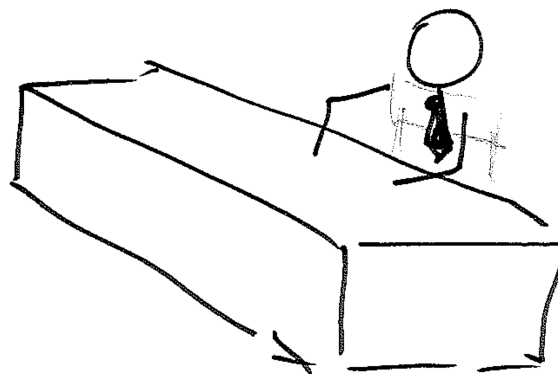
Education is always a Good Investment

- Training bioinformaticians
 - Technical & scientific trainings
 - Cultural trainings:
 - open-source
 - use of standards
 - collaborative developments
- **Educating decision-makers**

I need this \$1M new sequencer and 2 FTEs to analyse the data



We have the \$1M budget. Cannot you take on your internal resources to deal with the data?

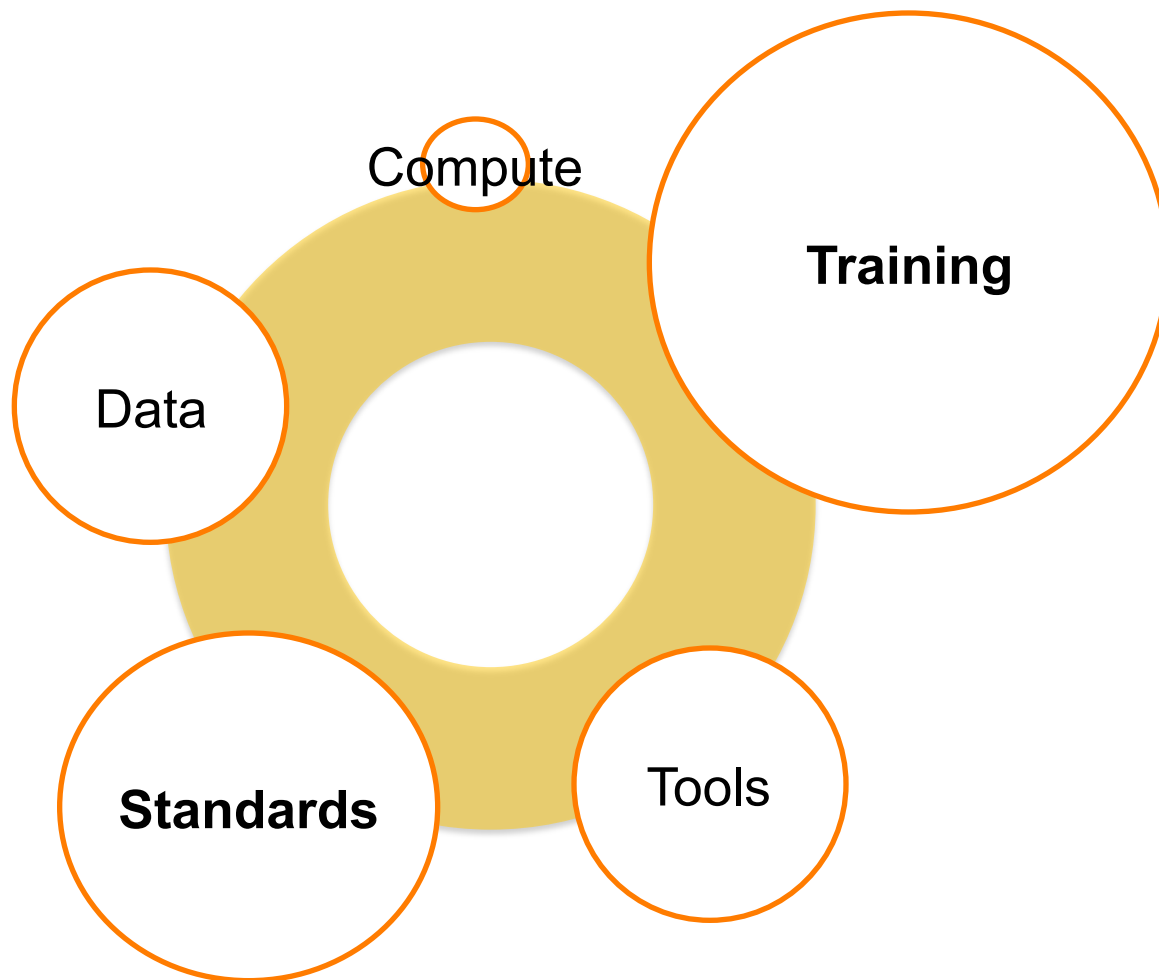


ELIXIR, a Modern Framework for Research

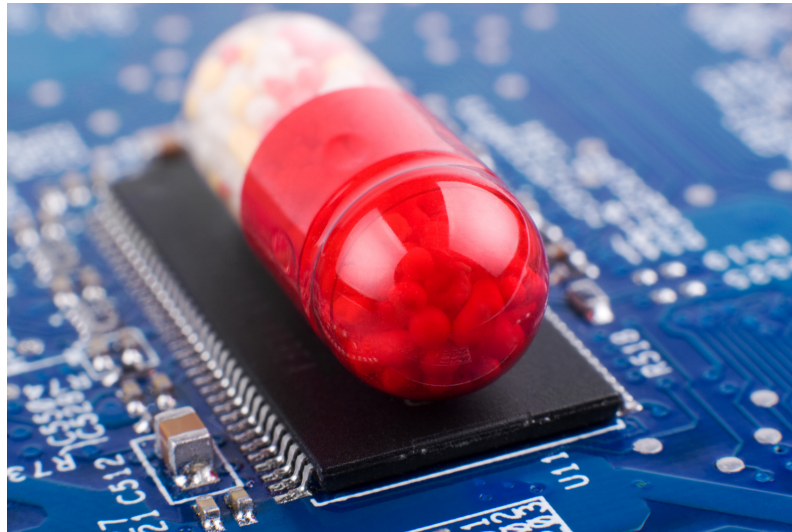
- Many large Life Sciences programs rely on **Big Data**
 - « OMICS data »: genetics, transcriptomics, proteomics, metabolomics...
- Examples are the EU-funded programs, *e.g.* FP7, IMI, H2020
- The success of such programs is driven by 2 key factors:
 - Good collaboration between Academia and Industry
 - Capacity to draw the opportunity from generated big data

ELIXIR offers collaborative bioinformatics framework that meets both requirements

ELIXIR and the Industry: Perspectives



Thank you for your attention



jerome.wojcik@quartzbio.com

www.quartzbio.com