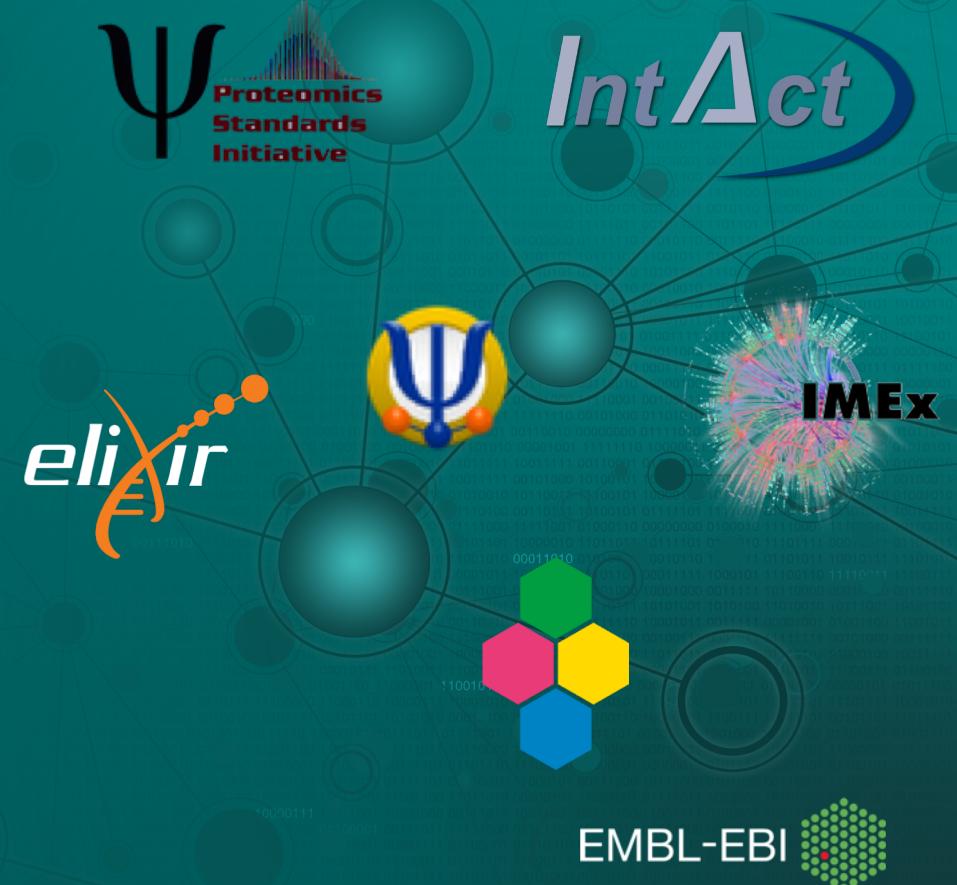


Prototyping the new PSICQUIC 2.0

PSICQUIC: <http://psicquic.github.io/>

BioHackathon project: <https://git.io/fpeGd>

Noemí del Toro Ayllón
Software Engineer
Molecular Interaction Team
EMBL-EBI
Hinxton, Cambridge, UK

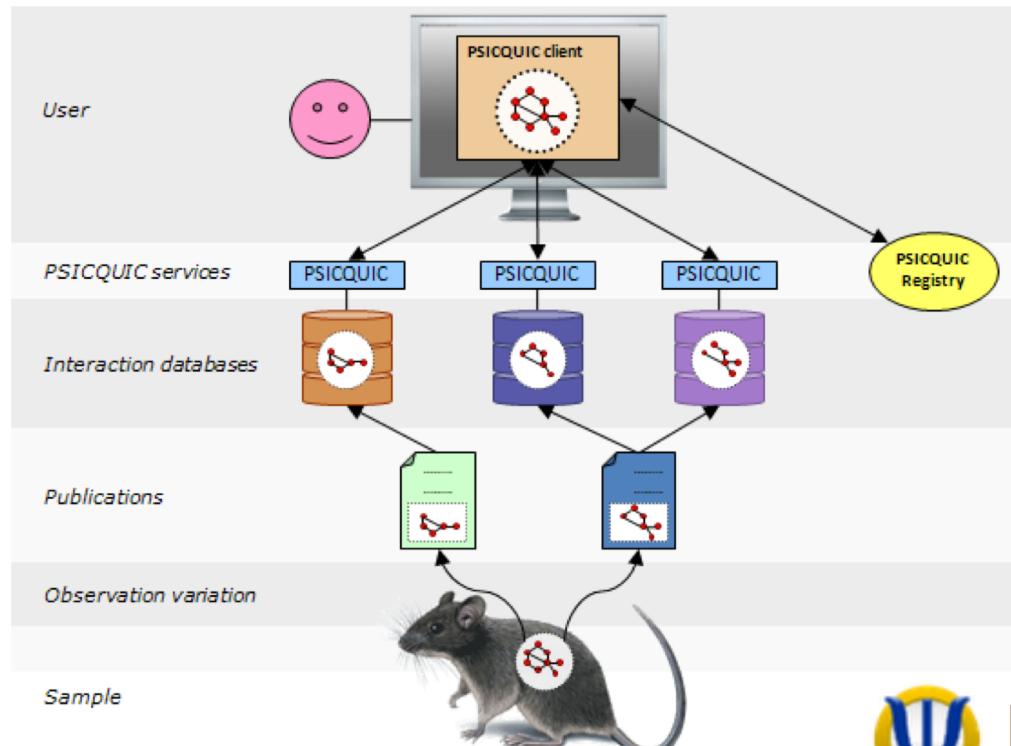


Background information

PSICQUIC

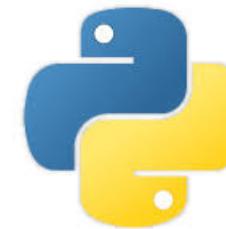
What is PSICQUIC?

- PSICQUIC is an effort from the HUPO Proteomics Standards Initiative (HUPO-PSI) to standardise programmatic access to molecular interaction databases.



What is PSICQUIC?

- PSICQUIC is established as a fundamental tool for the molecular interaction community. Key tools and databases such as Cytoscape, mentha, HiPPIE or GeneMania make use of it for data retrieval.
- External groups have developed R and Python packages to facilitate its use for specific communities which highlights its popularity.

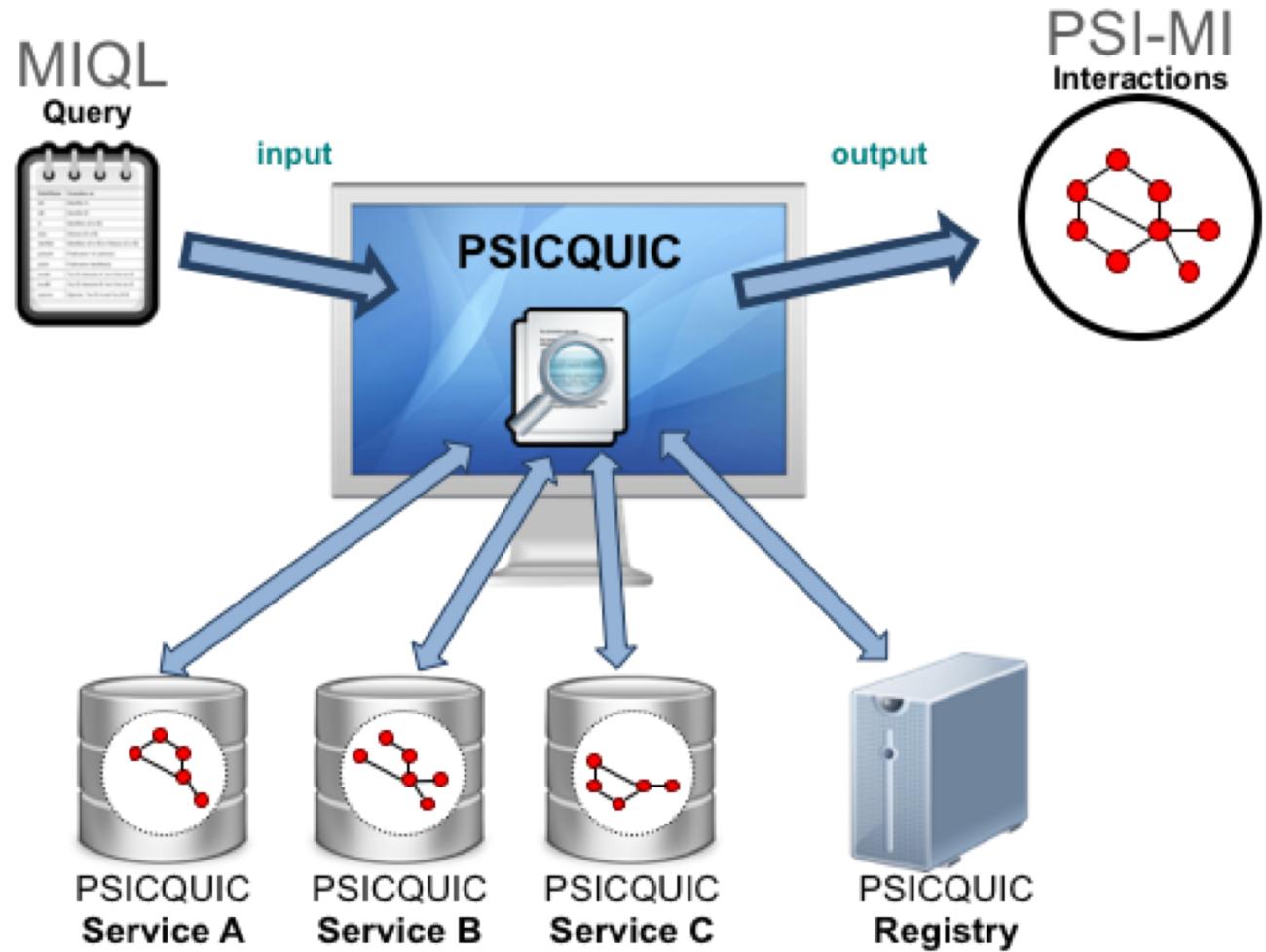


GENEMANIA



PSICQUIC Services

- PSICQUIC Services are Web Services (REST & SOAP)
- Independent databases
 - Adopting the same specification
 - Implementing the same interface
- Same methods to query services
- Same format to provide results

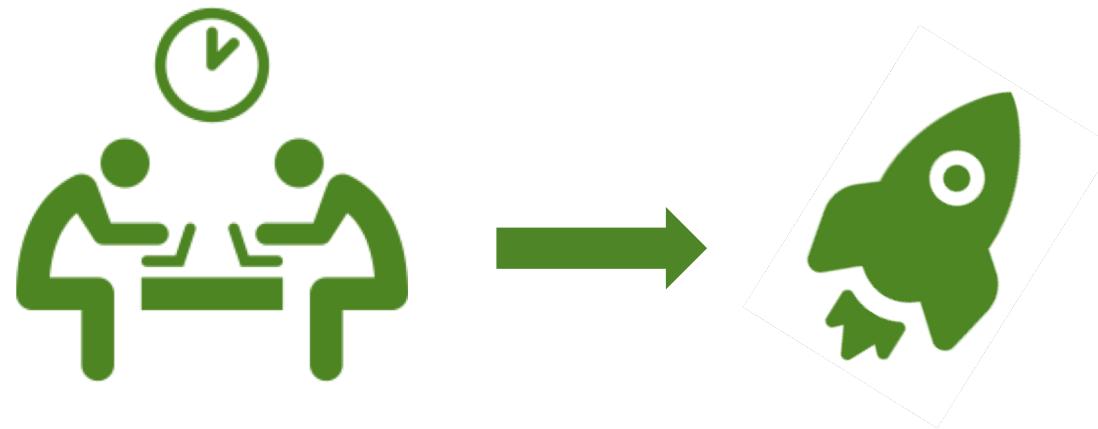


Goals of the hacking project

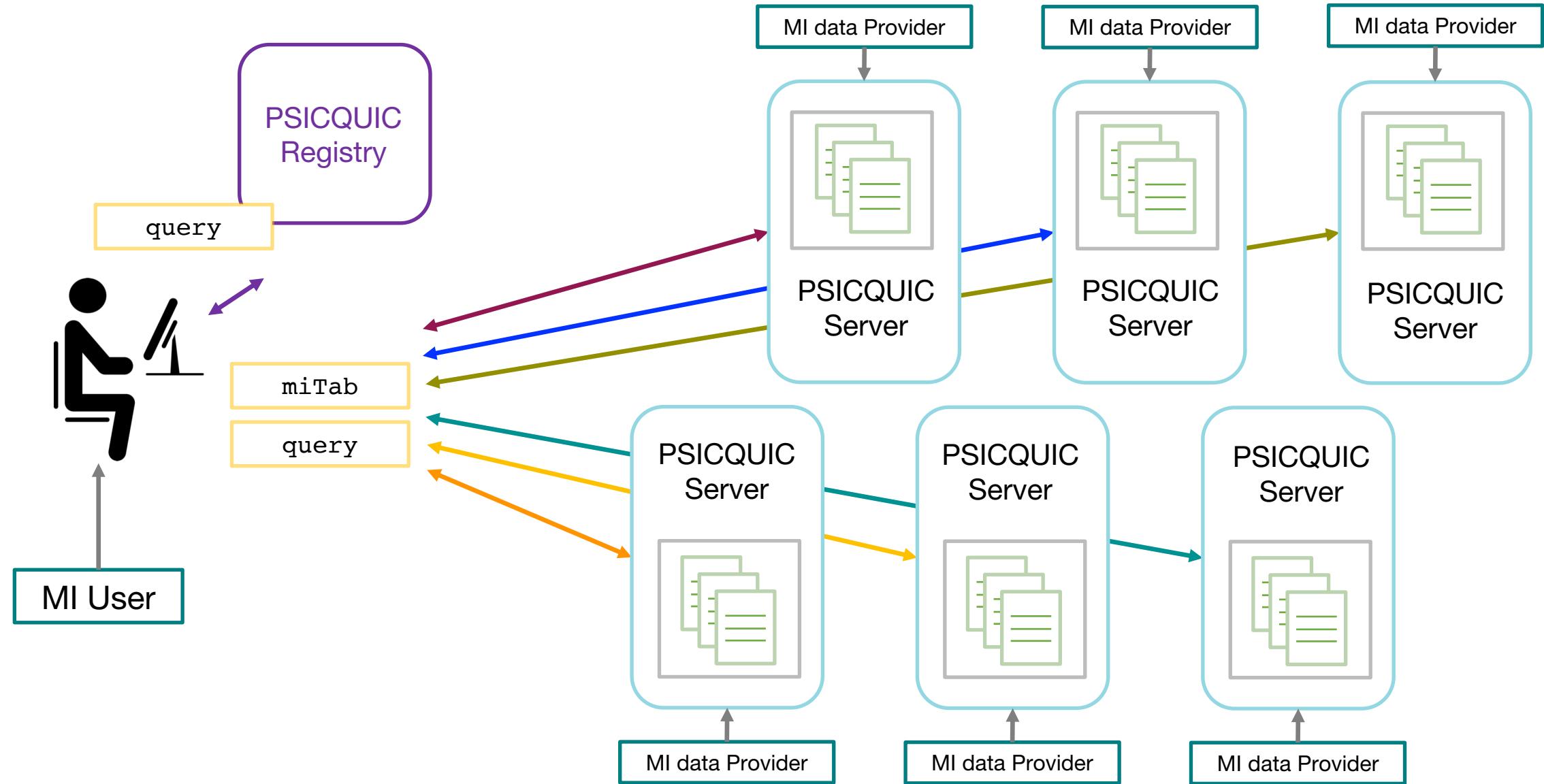
PSICQUIC 2.0

Motivation for the BioHackathon project

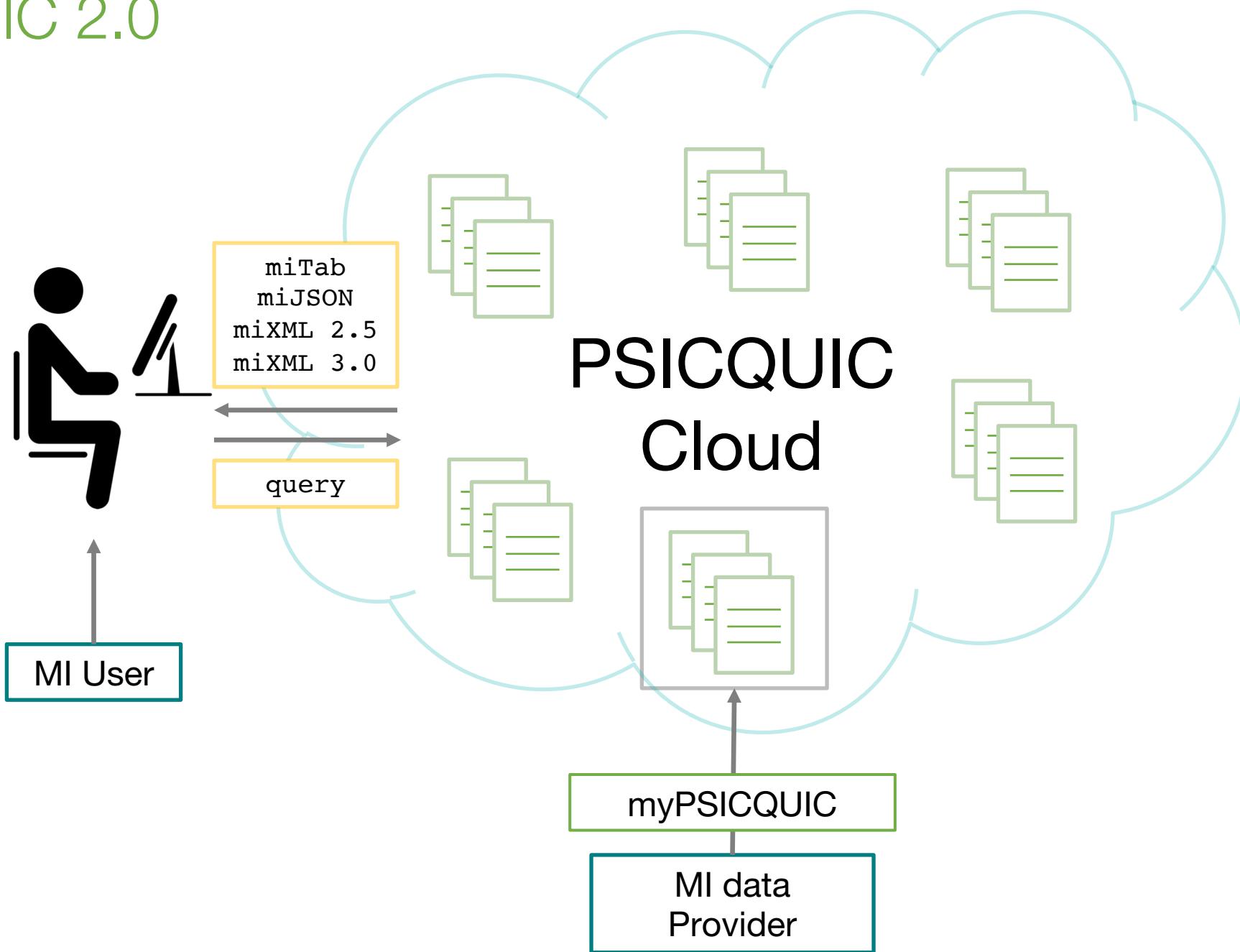
- The PSICQUIC user base has grown considerable since the service was created in 2011 and it has started to face significant challenges from both, data providers and users, among other reasons due to big increases in the data volume.
- Current architecture implies a long term commitment and maintenance from the data providers that they cannot always afford.



PSICQUIC

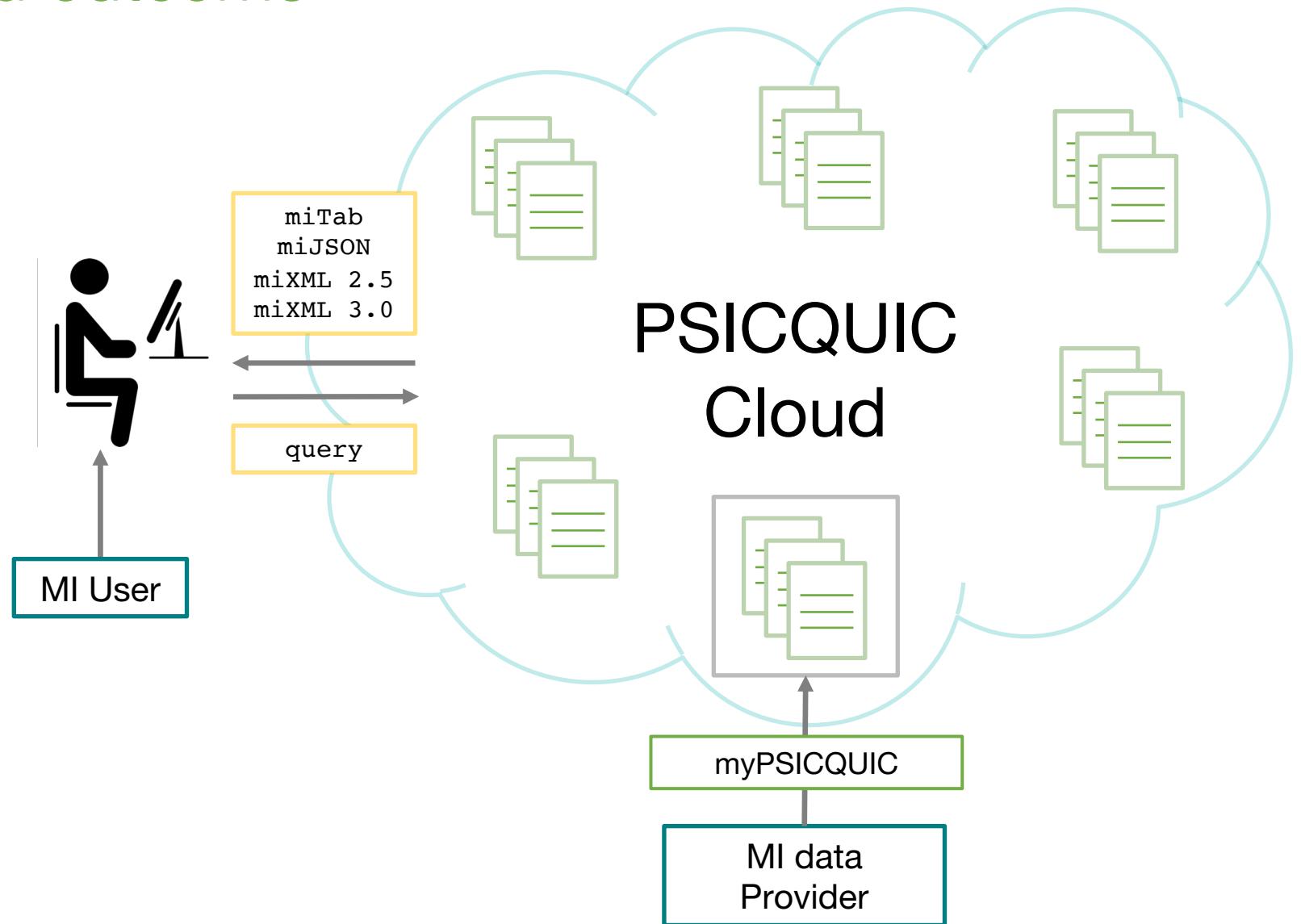


PSICQUIC 2.0



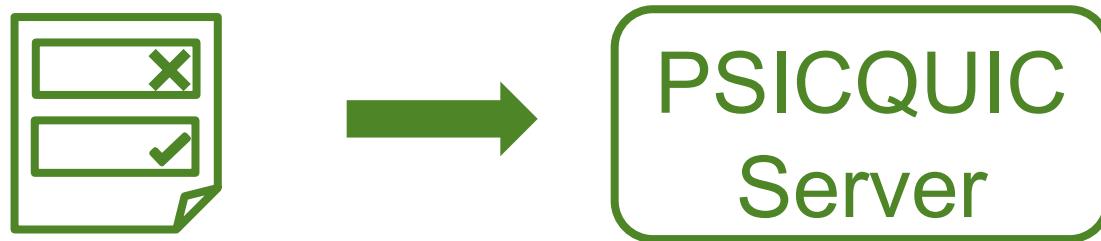
Goal and expected outcome

- The output of the BioHackathon is expected to be in the two main components of PSICQUIC 2.0.
- Both components can be developed in parallel and be integrated at the end.



Goal and expected outcome: myPSICQUIC

- Due to the complexity of the molecular interaction data, the data must be validated before it is uploaded to a PSICQUIC server. Currently, there is an online tool that offers that validation (<https://www.ebi.ac.uk/intact/validator>)



- The participants will need to work on the extension of the current validator by offering the option to upload the data at the end of the validation process
 - Changes in the front-end to incorporate the new functionality
 - Implement data transfer from the validator to the PSICQUIC server



Goal and expected outcome: myPSICQUIC

- PSI-MI Validator extension
 - Front-end tasks
 - Modernize current web application to new Javascript Framework technologies
 - Extend the interface to give the option to upload the data to PSICQUIC
 - Back-end tasks
 - Trigger indexing process to upload the data to PSICQUIC



Goal and expected outcome: PSICQUIC Cluster

- During the hackathon it is expected that the participants will be involved in the migration of Solr from the current standalone version to SolrCloud with the extra capabilities that it provides.
 - Currently, the reference implementation of PSICQUIC is based in a previous version of Apache Solr that gives all the capabilities of indexing and searching molecular interaction data under the hood.
- This will allow the continued use of Solr at the back-end and allow updating PSICQUIC to a distributed architecture.



Goal and expected outcome: PSICQUIC Cluster

- Migration to SolrCloud
 - Back-end tasks (to be developed for the reference implementation)
 - Migration of the current schema.xml to be compatible with the latest version of Apache Solr
 - Update/rewrite current indexing libraries to be compatible with the latest version of Apache Solr
 - Update/rewrite calls from the PSICQUIC web service to Solr



Other possible outcomes

- Due to the nature of the data some participants interested in network visualization could collaborate on the development of reusable web components for interaction network visualization.
- Improving or developing new PSICQUIC clients in other languages like R or Python could be an interesting outcome.
- Extending output formats to support `miJSON` (molecular interaction json format), `sif` (single interactions file) and `mixML` 3.0 as well as extending PSICQUIC to support `mitab` 2.8 (known as causalTAB) could be attractive for hackers familiarized with the domain.



Hack organisation

Organisation of the hacking project

- Duration: 04 hacking days
- Project representation at the BioHackathon
 - Noemi del Toro
- Invited resource people and associated expertise
 - John Zobolas (back-end developer)
 - Elisabet Barrera (front-end developer)
- Expertise from BioHackathon attendees
 - Programmers (front-end and back-end)
 - Previous knowledge in Java, Spring and Solr is recommended but not mandatory.

Steps and tasks

- Define in GitHub project folder: <https://git.io/fpeGd>
- Pool of tasks divided in different topics



Contact and links

- Contact (s)
 - Noemi del Toro (ntoro@ebi.ac.uk)
 - John Zobolas (john.zobolas@ntnu.no)
 - Elisabet Barrera (barrera@ebi.ac.uk)
- Links
 - PSICQUIC: <http://psicquic.github.io/>
 - BioHackathon project: <https://git.io/fpeGd>



HPI
DB_{2.0}



PSICQUIC Documentation

- PSICQUIC Specification
 - <http://psicquic.github.io/PsicquicSpecification.html>
- MIQL reference
 - <http://psicquic.github.io/MiqlDefinition.html>
- PSI-miTAB format
 - <https://github.com/HUPO-PSI/miTab/>
- PSI-miXML format
 - <https://github.com/HUPO-PSI/miXML>
 - Schema documentation
 - <https://rawgit.com/HUPO-PSI/miXML/master/3.0/doc/MIF300.html>
 - Schema
 - <https://github.com/HUPO-PSI/miXML/blob/master/3.0/src/MIF300.xsd>

Prototyping the new PSICQUIC 2.0

PSICQUIC: <http://psicquic.github.io/>

BioHackathon project: <https://git.io/fpeGd>

Noemí del Toro Ayllón

ntoro@ebi.ac.uk

Software Engineer

Molecular Interaction Team

EMBL-EBI

Hinxton, Cambridge, UK

