Instruct Data Management Strategy and the "conection" with Swiss OpenEM project

José-Maria Carazo

Biocomputing Unit, CNB-CSIC
INSTRUCT-ES. Instruct Image Processing Center, Madrid
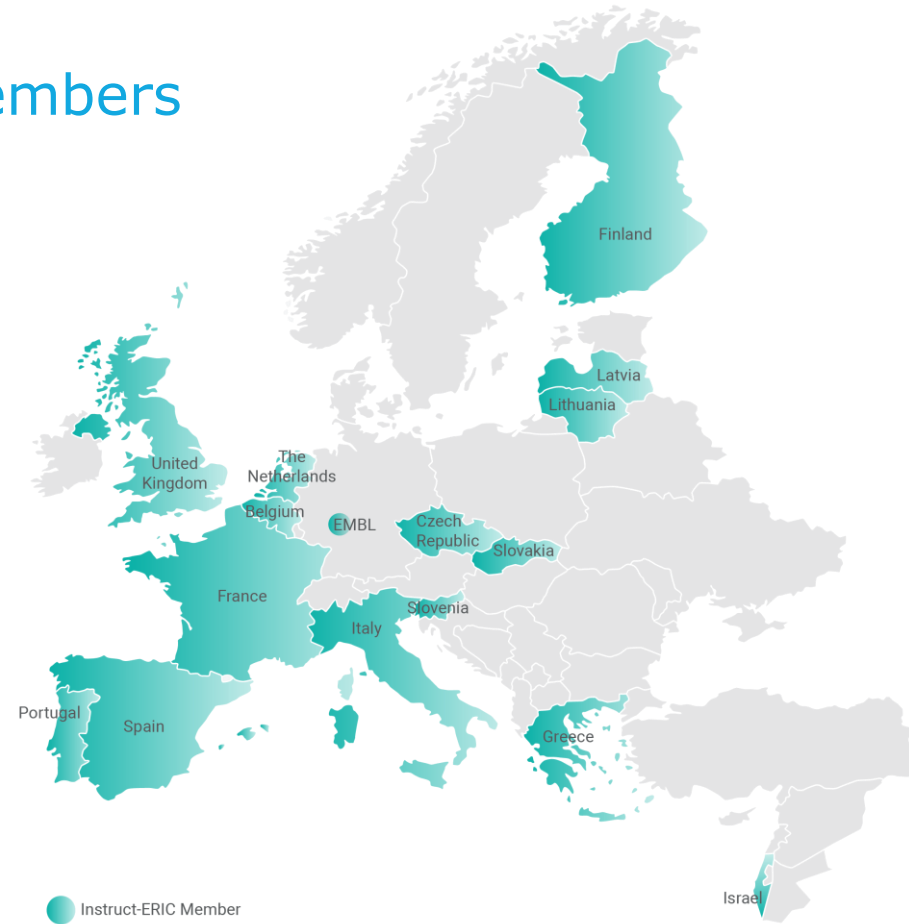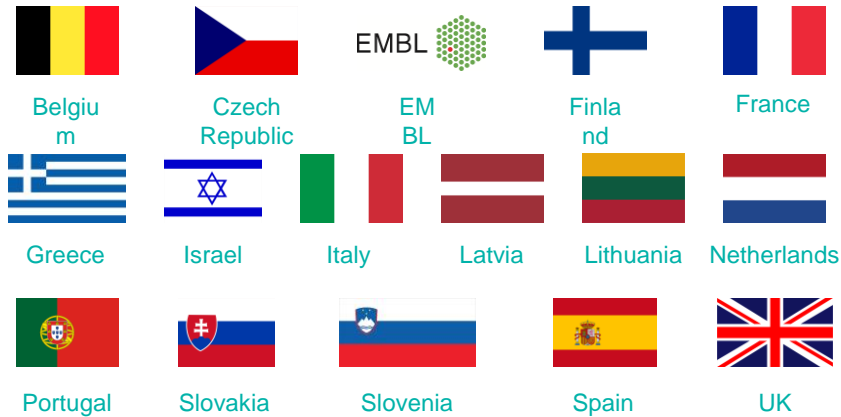
CNB Christmas workshop 2023

What is Instruct-ERIC?   Members

Instruct-ERIC is the single point of access to technology and expertise for structural biology research.

Instruct has 16 Members that each pay an annual subscription to allow their scientists to access the range of services that are available through Instruct.

Belgium	Czech Republic	EMBL	Finland	France
Greece	Israel	Italy	Latvia	Lithuania	Netherlands
Portugal	Slovakia	Slovenia	Spain	UK

Instruct-ERIC Member

CSIC

instruct Integrating Biology

Connection between Instruct and SwissOpen

- **The two projects have the general goal of making cryoEM Facility data FAIR (within constraints)**

- **They also some "budget" and a similar time scale (end 2025)**

- **All started with Spencer calling JM on the phone**

- **Several TC's since then, basically to incorporate interested parties and to confirm goals**

Instruct Data Management: Current state

- **Instruct has a Data Management Policy since 2015. Essentially:**
  - **Data belongs to the user (or user institution)**

- **Instruct Facilities are not required to store users data**
  - **A problema making the data discoverable and accessible**

- **Data have a maximum embargo period of 3/5 years**
  - **Too long**

- **Instruct access users are encouraged to make publicaly available the original raw data under a CC-By (or like) license**
  - **"Encouragement" does not work**

Instruct Data Management: Our goal (FAIR Data)

- **We need to make the data adquired during an Instruct Access visit discoverable in association with the Access project**
  - Make the project management system (ARIA) aware of data…. and Instruct Access projects discoverable in themselves

- **We need to ensure that data are indeed made accesible**
  - Keep a copy of the data and make them available after embargo

- **Streamline deposition in EMPIAR and EMDB of suitable data**

Instruct Data Management: Current resources

- **On going InfraTech project on Fragment Based Screening with a specific WorkPackage for Data Management (in FBS). Until March 26**

- **ISIDORE JRA on Data Management. Until early 25**

- **Note: We care about ALL Instruct provided data, and not only cryoEM, although collaboration with SwissOpen EM will impact cryoEM only**

- **So, END OF 25, which is SwissOpen EM final date, fits perfectly within our Schedule!... For that to work on a synergy between the two projects make sense: We should end in about the same time, and we have resources (surely more on the Swiss side....)**

Instruct Data Management: Schedule

- **Financial resources are alocated until about end of 2025**

- **The standardization work should be done by end 2024/early 2025**

- **The first implementation work at (at least some) of Instruct center and the start of the "production phase" must happen during 2025**

- **We are not thinking in "pilots", but on concrete practical procedures for end of 2025**

    - **Note that this approach also means that we have to be effective and "humble", since we cannot aim for the "perfect" solution, but for a good solution that can evolve still better**

Instruct Data Management: Relationship with non cryo-EM

- **In Instruct we must care for ALL Structural Biology techniques, not only cryoEM**

- **Indeed, there are Instruct financial resources for many SB techniques Data Management (although focused on Fragmet Based Screening)**

- **The collaboration with SwissOpen EM will impact (1) cryoEM, of course, but also (2) the way the Instruct Project Management System (ARIA) interacts with data**

  - **At Instruct we must work always thinking that the data management solutions should be extendable to other techniques**

Reiterating the focus of this meeting

- ## Are we all "meaning" the same?

- ### _Standarization of cryoEM data adquisition models (Facility oriented)_

  How to Access data will come later, and we need it, but we first start with standards!.

  ON standards…. Whatever we can re-use, lets reuse it