The Spanish National Bioinformatics Institute (INB/ELIXIR-ES) Renovating its portfolio in its 20th anniversary
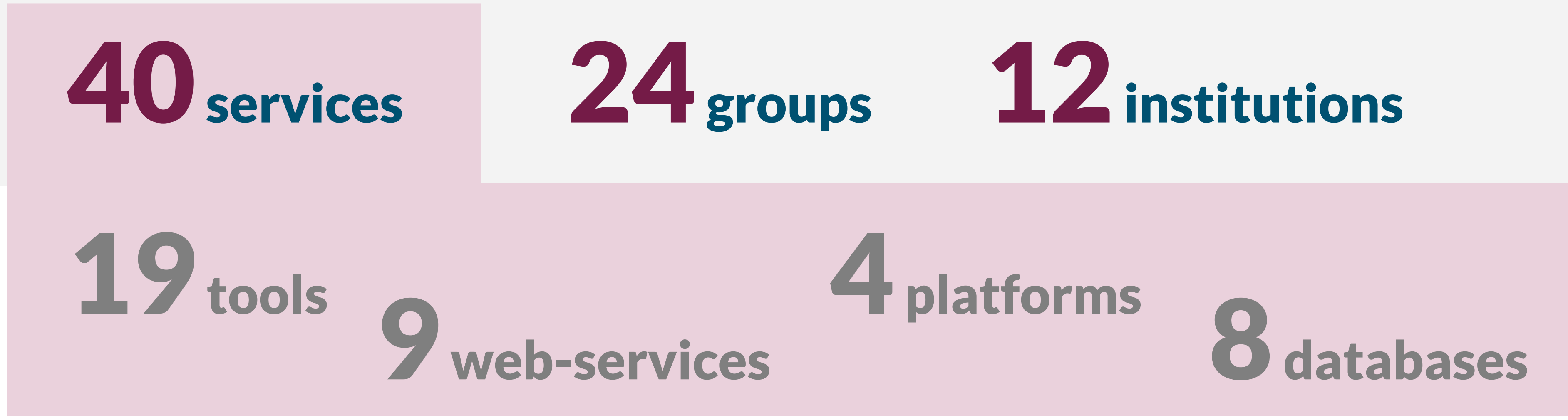
Eva Alloza[1,2], INB/ELIXIR-ES members, Alfonso Valencia[1,2,3] & Salvador Capella-Gutierrez[1,2]

[1] Spanish National Bioinformatics Institute (INB/ELIXIR-ES)
[2] Barcelona Supercomputing Center (BSC)
[3] Catalan Institution for Research and Advanced Studies (ICREA)

elixir SPAIN

The **Spanish National Bioinformatics Institute (INB)** is the **ELIXIR Node in Spain (ELIXIR-ES)**.

The **INB was founded in 2003** as a distributed network of nodes with a central coordination hub. At present, INB counts with 26 research groups distributed across 14 institutions in Spain. Currently, the two main overarching objectives are: **deepen its alignment with ELIXIR and its communities with a strong focus on human data related activities; and increasing its translational impact within the Spanish National Health System** regarding the secondary use of health-related data for research. At the national level, the INB/ELIXIR-ES leads the Translational Bioinformatics Network (TransBioNet) and coordinates IMPaCT-Data, the Data Science pillar of the Spanish National Infrastructure for Precision Medicine.

Renewed Service Delivery Plan

40 services **24** groups **12** institutions

19 tools **9** web-services **4** platforms **8** databases

INB services in purple and TransBioNet members in blue, as of June 2023

Alignment and involvement in ELIXIR

➔ Co-leadership of the **ELIXIR Tools** and **Training** platforms; the **Federated Human Data**, **Rare Diseases** and **Biodiversity Communities**; and the **Cancer Data Focus Group**.

➔ Secured controlled access to human data through the **European Genome-phenome Archive (EGA)**, jointly managed by EMBL-EBI and CRG with the technological support of BSC. EGA is part of the **ELIXIR Core Data Resources** (CDR) and the **ELIXIR Recommended Deposition Databases** (EDD).

➔ Establishment of the **Spanish Federated EGA node** (FEGA-ES).

➔ Development of the ELIXIR platform for supporting technical monitoring and community-driven benchmarking efforts **OpenEBench**.

➔ Designation of ELIXIR-ES resources 3D-Bionotes and FAIRtracks (in collaboration with EMBL-EBI and ELIXIR-NO) as part of the ELIXIR **Recommended Interoperability Resources** (RIR).

➔ Leadership of the **GA4GH Beacon and ELIXIR Beacon Network**, including recent extensions of the Beacon V2 standard, e.g. clinical, medical imaging.
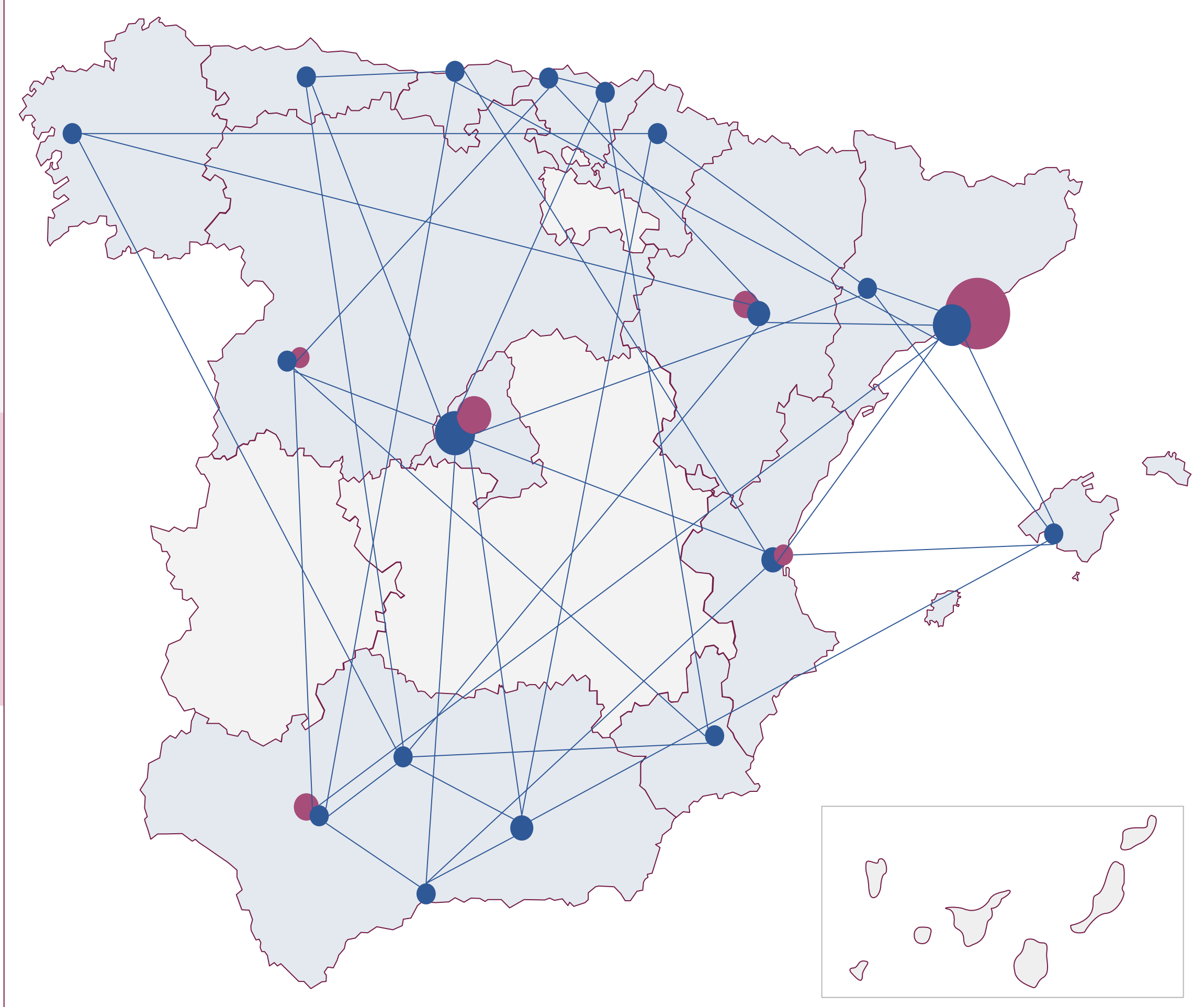
Moreover, groups from INB/ELIXIR-ES have intense activity across the ELIXIR structure:

➔ **5 Platforms**: Tools, Data, Compute, Training, Interoperability.

➔ **13 Communities**: Federated Human Data, Human Copy Number Variation, Rare Diseases, Galaxy, Proteomics, 3D-BioInfo, Intrinsically Disordered Proteins, Toxicology, Plants Sciences, Single Cell Omics, Systems Biology and Biodiversity, and the emerging Research Data Management community.

➔ **6 Focus Groups**: Machine Learning, Cancer Data, Health Data, EOSC, Impact, Innovation and Industry.

➔ **11 Implementation studies**, including strategic ones like **Federated EGA** framework; **ELIXIR Beacon**; ELIXIR **Rare Disease services**; and Making **container services, sustainable and widely adopted**.

INB/ELIXIR-ES are also active in relevant activities for ELIXIR and the whole community:

➔ Promotion of **FAIR data principles** and scientific software development best practices, and training.

➔ Organised the **ECCB2022**, including an **ELIXIR outreach event with Latin American** bioinformatics communities, and the co-located **ELIXIR Innovation and SME Forum** on Data-driven innovation in Healthcare Diagnostics.

➔ Organised the **Federated EGA Celebration Symposium** and presentation of the inaugural FEGA nodes in the framework of the 10th GA4GH plenary in Barcelona.

➔ **Gender equity** in Bioinformatics: ELIXIR has launched the ELIXIR LEadership And Diversity mentoring programme (**ELEAD Pilot**), which builds from previous experiences of the **Bioinfo4Women programme**.

BSC leads the Spanish participation in **ELIXIR-CONVERGE**, with other members of the node as beneficiaries. In this context, the UPF is responsible for the *FAIR encoding and access to Toxicology data* use-case while the IRB Barcelona leads the *FAIR organisation of biomolecular simulation information* one, which triggered particular efforts on organising biomolecular simulation data through dedicated infrastructures. The CRG co-leads WP7 related to the Federated EGA for transnational access of COVID-19 host data, which served to the development and initial testing of the **Federated EGA Maturity Model**. This maturity model is nowadays the reference for the onboarding of new nodes into the federation. It has also served as reference for the establishment of the maturity model for the Federated One Health Surveillance Platforms. Beyond ELIXIR-CONVERGE, spanish institutions have an active role in other ELIXIR-led projects like **EOSC-Life, B1MG**, **GDI**, and **BY-COVID**.

Translational Bioinformatics

TRANSBIONET

➔ Creation and coordination of **TransBioNet**, Translational Bioinformatics Network, an open experts' forum bringing together 38 bioinformatics groups at Health Research Institutions across Spain.

➔ Share best-practices and experiences, early access to standards, protocols and best practices by ELIXIR and GA4GH, benefit from access to the computational infrastructures provided by the INB and ELIXIR, and training activities, altogether have a **direct impact on researchers working closely to clinicians**.

Precision Medicine

IMPaCT|Data

➔ **IMPaCT-Data** is the Data Science pillar of the Spanish National Infrastructure for Precision Medicine associated with Science and Technology (IMPaCT), created by the Institute of Health Carlos III in 2021.

➔ INB/ELIXIR-ES nodes and TransBioNet members are integrated into IMPaCT-Data, structured around 60 research groups of 47 institutions across Spain, and coordinated by BSC.

➔ IMPaCT-Data aims to develop the **first iteration of a Biomedical Cloud** to establish the mechanisms for **working with genomics, clinical and medical imaging data for research purposes in a distributed manner**.

➔ The project has delivered 1) **two demonstrators** on how to build a federated ecosystem enabling distributed analysis via Galaxy and using GA4GH Passports and Visas technologies, 2) recommendations and guidelines regarding the standardization of biomedical data to facilitate its interoperability, access, and governance.

➔ The **reference implementation of the Biomedical Cloud** under development will include several components as Beacon and Beacon network, Local EGA, Galaxy, openVRE, AAI, CVMFS, OMOP, OHDSI, cBioPortal, and other central services.

➔ IMPaCT-Data aligns with many European projects and initiatives and implements standards pushed forward by community-driven initiatives, e.g. ELIXIR, GA4GH, 1+MG/B1MG, EUCAIM.

find out more in Poster #74

Contact:

INB/ELIXIR-ES Coordination Hub

Barcelona Supercomputing Center (BSC)
C/ Plaça Eusebi Güell, 1-3, 08034-Barcelona, Spain

https://inb-elixir.es ✉ inb.hub@bsc.es 🐦 @INB_Official
Spanish National Bioinformatics Institute (INB)