LifeWatch ERIC Blockchain platform in support of Biodiversity & Ecosystems Communities of Practice

Dr. Juan Miguel González-Aranda

Director Common Facilities LifeWatch ERIC in Spain

Acting Chief Technology Officer

LifeWatch ERIC Scientific Community Meeting

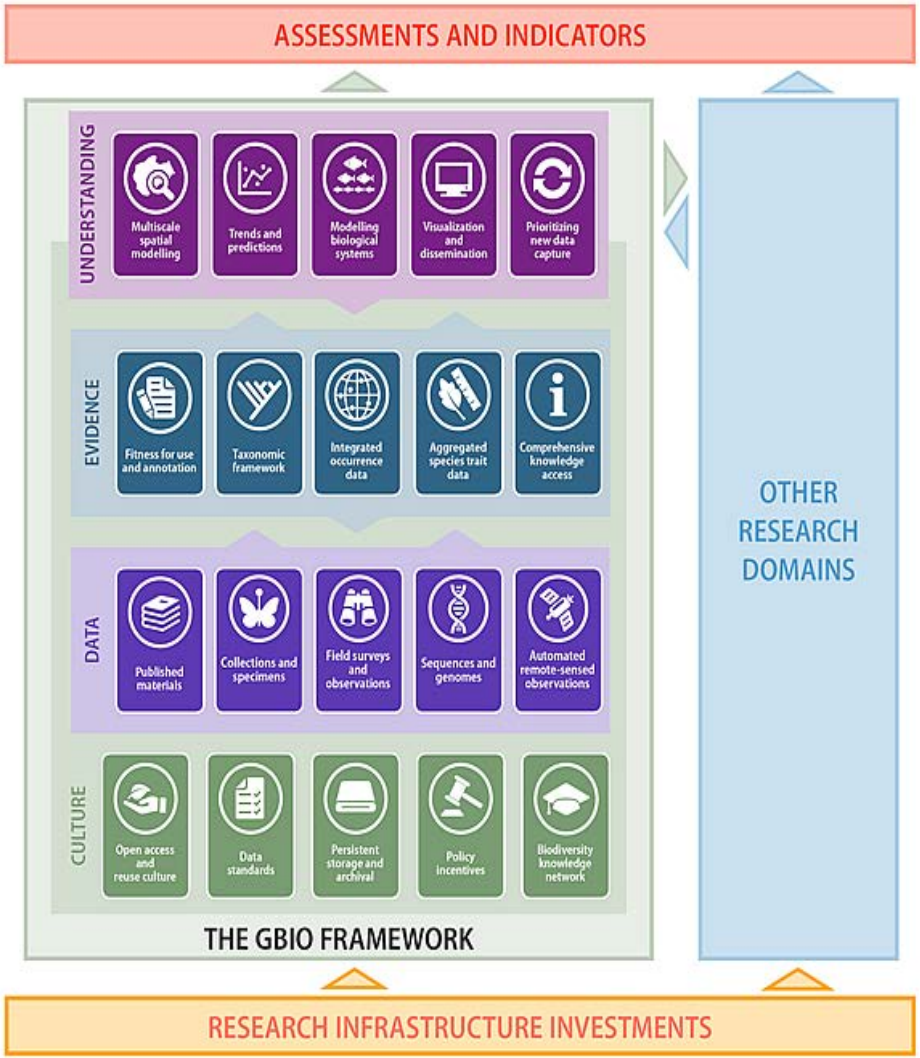
CNR Headquarters, Rome-Italy

May 27th-29th, 2019

Table of Contents

PART I

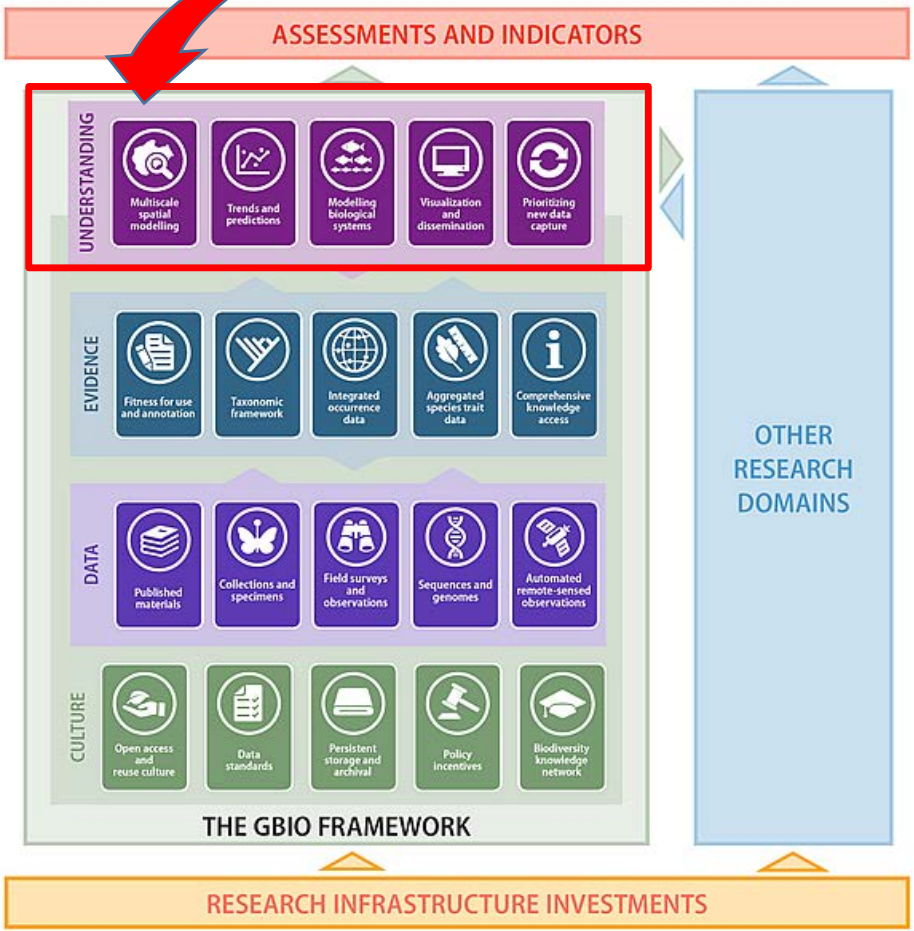
LifeWatch ERIC in the international (distributed) Research Infrastructures context



The Global Biodiversity Informatics Outlook (GBIO) Framework (Hobern et al. 2012) identifies 20 components as essential elements of biodiversity informatics and organized as four layers: Culture, Data, Evidence and Understanding.

The engagement of the (distributed) ICT component of LifeWatch ERIC with other pan-European Research Infrastructures focusing on Biodiversity and Ecosystem Research (GBIF, DiSSCo, e-LTER and AnaEE, among others), is being achieved in a number of ways. This approach is facilitating meeting GBIF's overall objective stated as: "Connecting data and expertise: a new alliance for biodiversity knowledge" (Hobern et al, 2019).

LifeWatch ERIC



UNDERSTANDING: Building modeled representations of biodiversity patterns and properties, based on any possible **EVIDENCE**, based on the following components:

(a) Multiscale species modelling;

(b) Trends and predictions;

(c) Modelling biological systems;

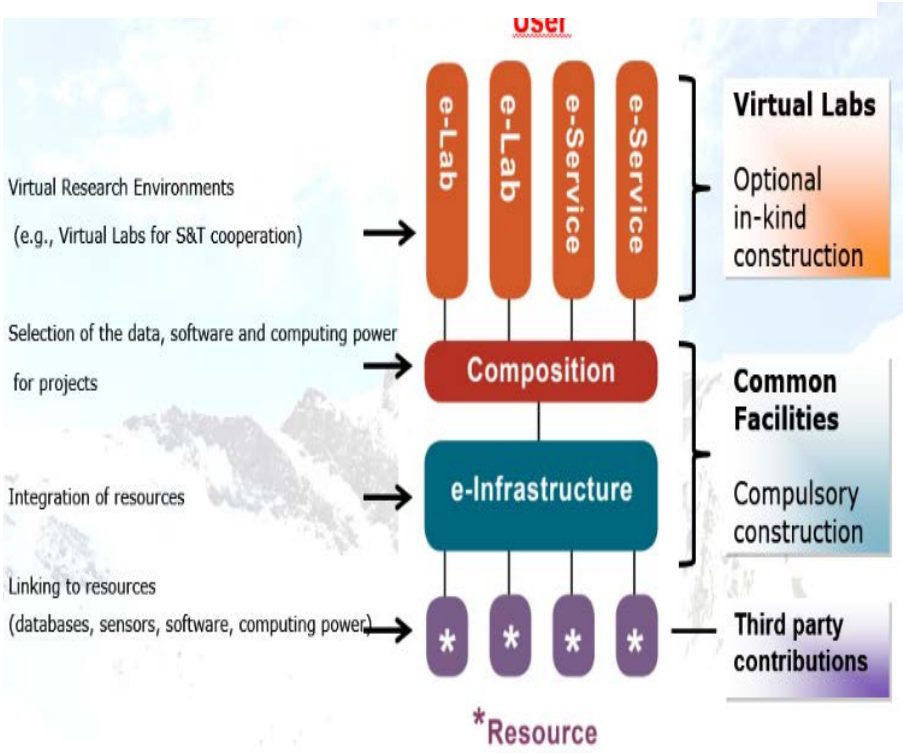
(d) Visualization and dissemination;

(e) Prioritizing new data capture.

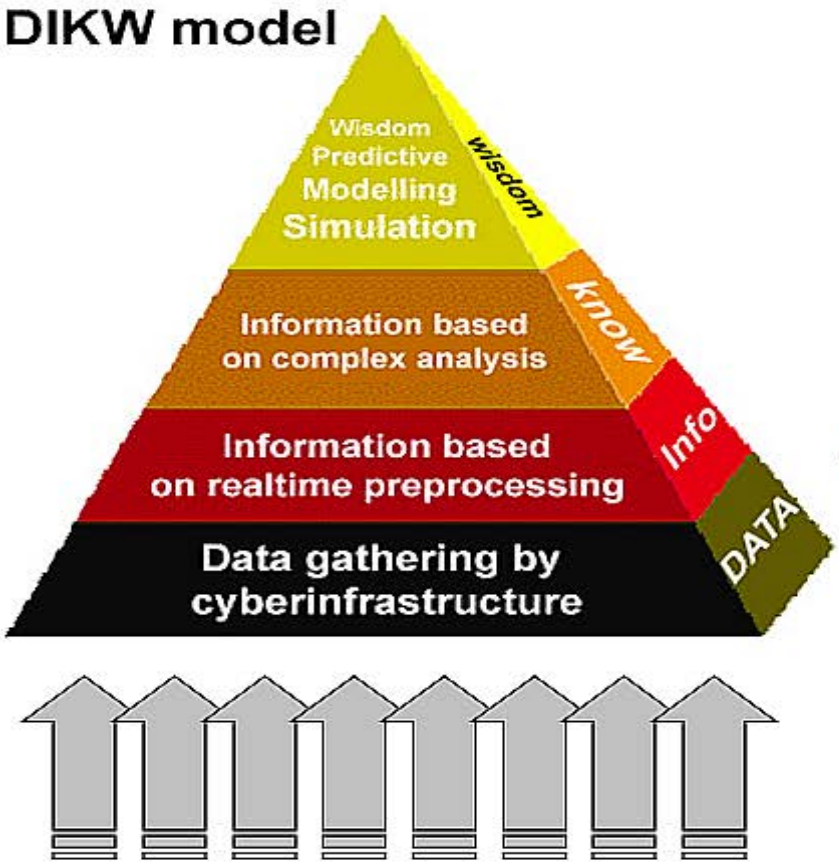
2nd Global Biodiversity Information Conference -GBIC2- held under the auspices of GBIF in Copenhague-Denmark (July 2018): One of the main conclusions was that the **Evidence layer** is the fertile interface to develop sound synergies for collaboration (counting on previous cited RI, among others) in order to support in turn GBIF through the development of concrete activities:

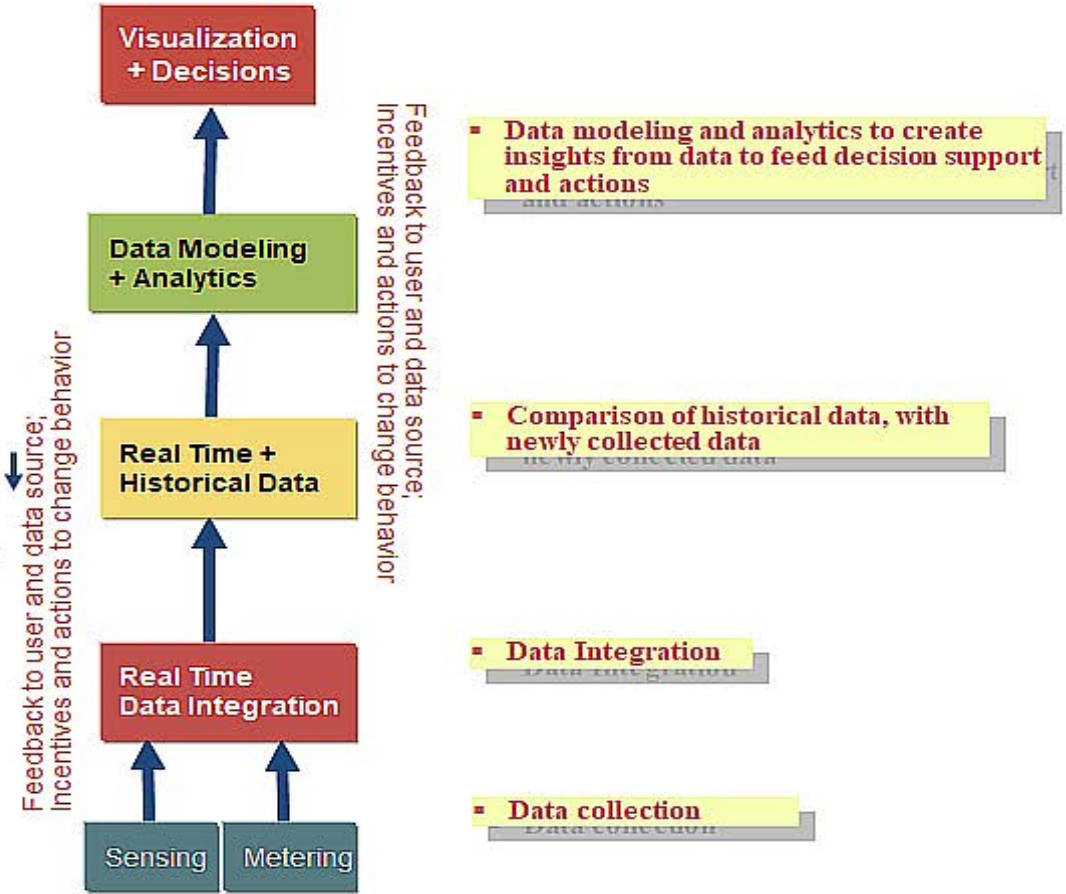
- ✓ **Co-design, development and deployment of a multi-purpose Virtual Research Environment (VRE)** specifically by integrating existing distributed e-Resources and e-Services, and by engaging their associated Communities of Practice;

- ✓ **Co-design and co-implementation of relevant e-Services in LifeBlock (LifeWatch ERIC blockchain-based technology platform);**

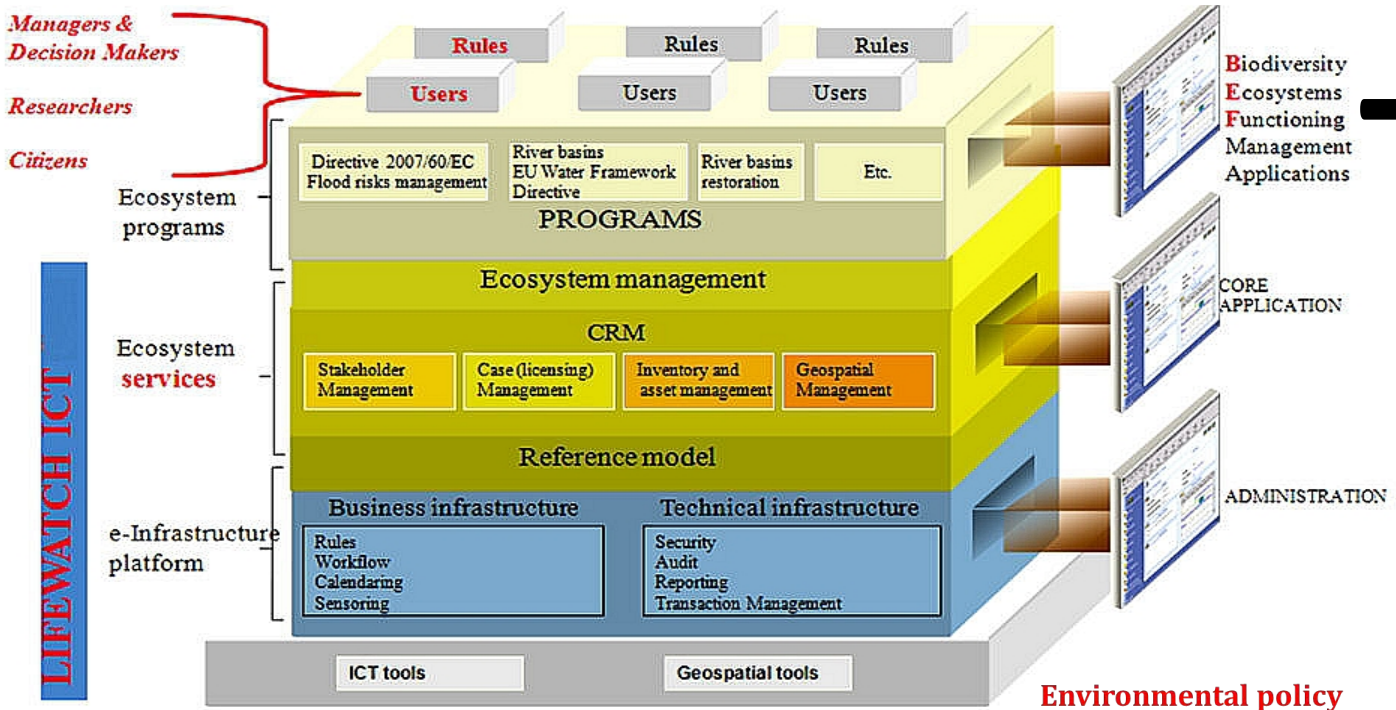


Therefore, LifeWatch ERIC will serve as an ICT "Backbone" to support common & rest of distributed facilities and initiatives and their users': Researchers, Decision Makers & Managers, Citizen Scientists.





All together: RESEARCHERS, Managers & Decision Makers, and Citizens (Science)



e-Science

Environmental policy & management

DATA PROVIDERS & RESEARCH COMMUNITIES ↔ LifeWatch ↔ POLICY & DECISION MAKERS

INTEROPERABILITY

Ecosystem Services (ES)

LifeWatch ERIC will create and deploy pilot use cases relying heavily on the paradigm of ES and aims to increase the reliability of benefit transfer by developing Virtual Research Environments –VRE- engaged to a Blockchain-based platform ("LifeBlock") capable to analyze multiple unstructured inputs from remote and in-field sensing, Remote Sensing-GIS datasets and original studies in order to produce value estimates.

The application of VRE will make it possible to integrate for the first-time environmental parameters and socioeconomic conditions with aspects of biodiversity conservation and its role in the provision of ecosystem services in diverse frameworks, in order to provide new transfer functions that might eventually replace (or at least complement) the meta-regression techniques primarily used at present.

In fact, the active participation on integrating collections data as the main resource needed for the integration of GLOBIS-B developments (EBVs) and results (Kissling et al., 2016) jointly species traits in LifeBlock platform, to feed Ecosystem Services needed to supply Biodiversity Ecosystem Services VRE, to further support thematic areas (in particular, Invasive Alien Species, through the LifeWatch ERIC Common Facility and Distributed Nodes Internal Joint Initiative, a.k.a. IJI).

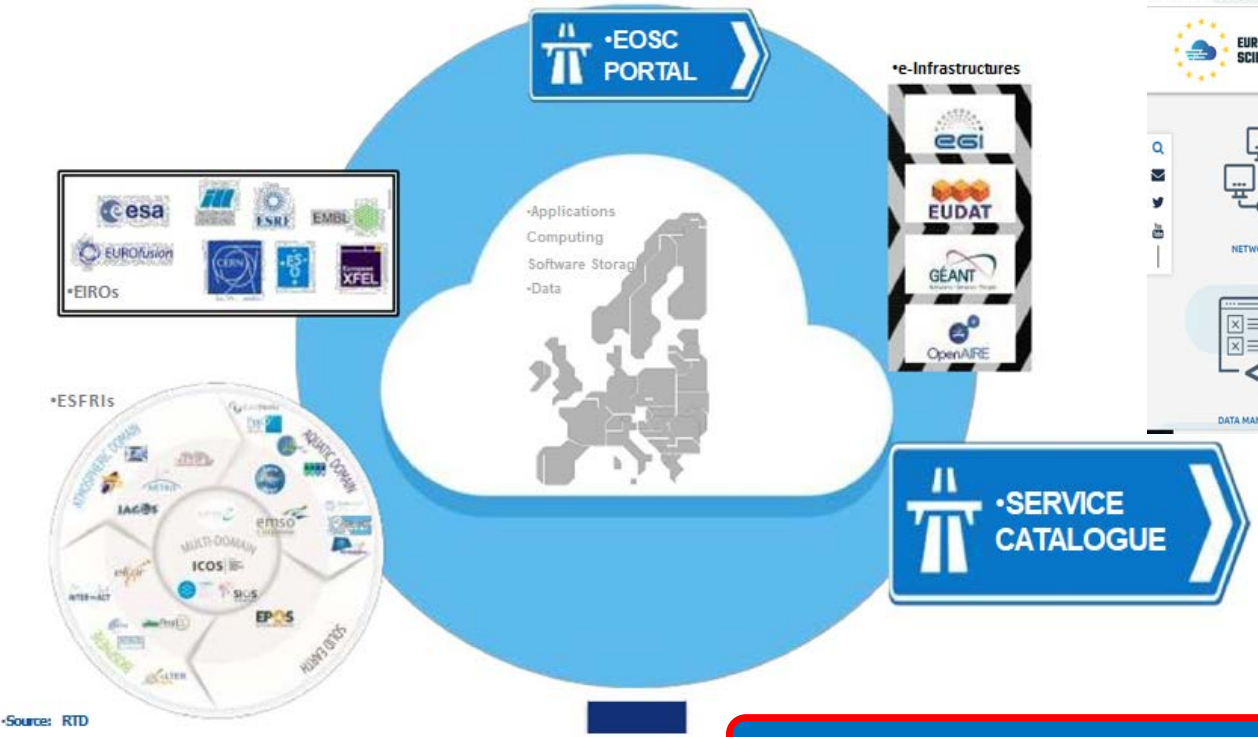
Facing e-Biodiversity challenges together: GBIO framework-based synergies between DiSSCo and LifeWatch ERIC

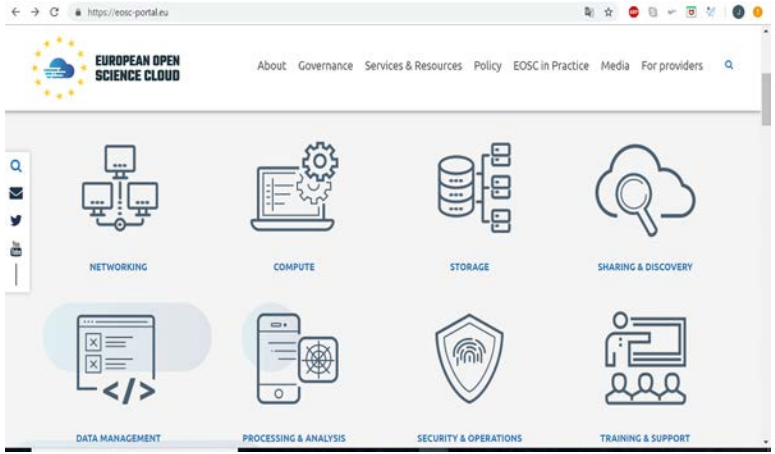
González-Aranda, Juan Miguel[1]; Koureas, Dimitris[2]; Addink, Wouter[2]; Hirsch, Tim[3]; Arvanitidis, Christos[1]; Sáenz-Albanés, Antonio-José[1]; Schalk, Peter[2].

has been recently approved for oral presentation during the session "SI22-DiSSCo as a model for regional development of collections infrastructure" in the Infrastructure track at the forthcoming joint Biodiversity Next conference in Leiden, The Netherlands, October 20-25, 2019.

EUROPEAN OPEN SCIENCE CLOUD

https://eosc-portal.eu

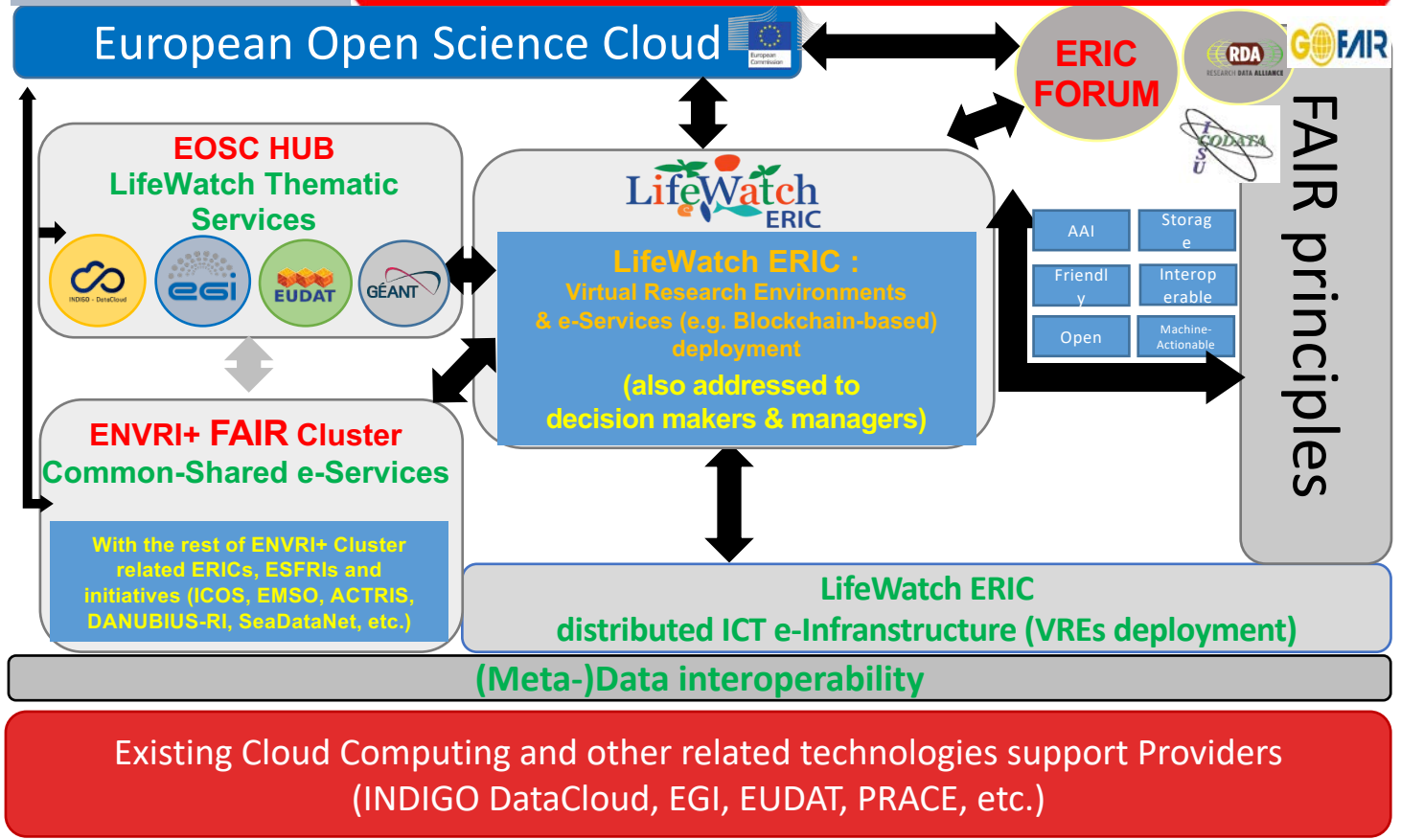




European Open Science Cloud

How to integrate ERIC ? Our approach

- Integration of ERIC development into pan-European e-Infrastructures, particularly in the new European Open Science Cloud-EOSC
- Creating synergies between ERICs and the (Regional) Research and Innovation Strategies for Smart Specialisation-RIS[3]



Existing Cloud Computing and other related technologies support Providers
(INDIGO DataCloud, EGI, EUDAT, PRACE, etc.)

PART II

Why LifeBlock, why using Blockchain technologies ?

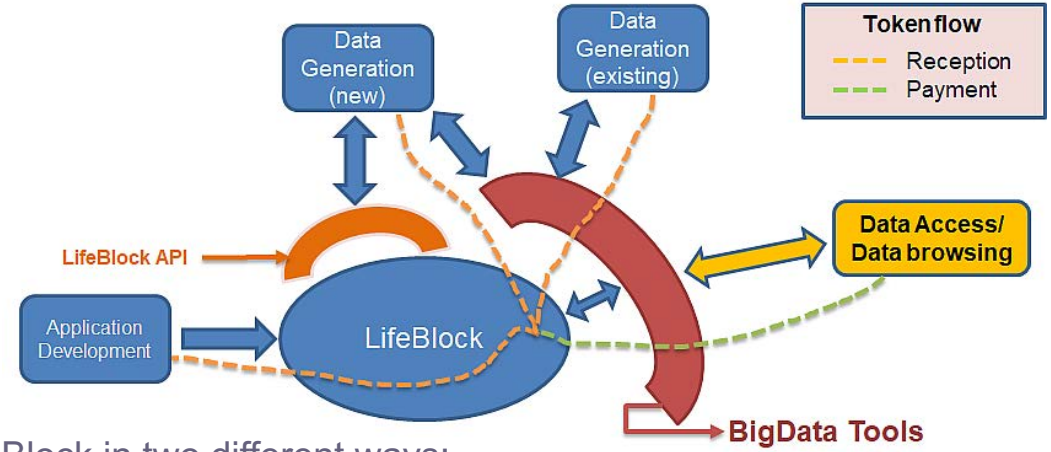
Why Blockchain?

- **Disintermediation.-** The core value of a blockchain is enabling a database to be directly shared across boundaries of trust, without requiring a central administrator.
- **Validation.-** Every blockchain transaction contains their own proof of validity and their own proof of authorization, instead of requiring some centralized application logic to enforce those constraints.
- **Price.-** Blockchains offer a way to replace organizations with a distributed database, locked down by clever cryptography. And once it's been written and debugged, code tends to be an awful lot cheaper.
- **Robustness.-** Blockchain is extreme fault tolerance, which stems from their built-in redundancy. Every node processes every transaction, so no individual node is crucial to the system as a whole. Similarly, nodes connect to each other in a dense peer-to-peer fashion, so many communication links can fail before things grind to a halt.

Paving the way to LifeBlock

- ✓ **The deployment of an "anti-tampering" provenance and traceability e-Services based on Blockchain technologies, the so-called "LifeBlock" (LW ERIC Blockchain) platform will guarantee the availability of FAIR (Findable-Accessible-Interoperable-Reusable) compliant data.**
- ✓ **This new approach represents a change in accessibility of Big Data for biodiversity research and also creates a new type of e-Science European RI facilities. Answering this demand requires the analysis of impacts and managerial actions at different spatial and temporal scales.**
- ✓ **Its achievement entails observation (and monitoring) data from ecosystems, data from manipulative experiments, the relevant appropriate data storage and curation, the setting of standards and protocols to ensure the interoperability of data, accurate models on ecosystem dynamics, and e-Infrastructures where models devoted to specific environmental components can be coupled.**

Detailed explanation of LifeBlock Architecture



- Any organization can join LifeBlock in two different ways:
  - **Existing databases**.- Initially they will be joined through import / export tools to progressively integrate completely into the architecture
  - **New databases**.- They will have tools to perform a native integration in the environment.
- The LifeBlock architecture will provide access to data on ecosystem services, both free and private
  - **Free data**.- Organizations will make it available to the public through the tools provided by Lifeblock with automatic integration
  - **Private data**.- Accessible through payment used Lifetoken
- LifeBlock tools
  - LifeBlock will provide tools for the integration of ecosystem data in the blockchain natively in a simple and accessible way for all participating agencies
  - Likewise, LifeBlock will provide tools for the integration of existing databases for their native integration into the system.
  - Finally, LifeBlock will have applications that will allow a simple and structured access to the data using BigData technology that will allow to classify them and structure them appropriately for a simple and fast access

Which will be the usage of LifeBlock ?

LifeBlock is going to be a blockchain infrastructure whose objective will be to store, manage and give access to all kinds of data and ecosystem applications (e-Services) for all types of organizations and people who want to access/share them with the following characteristics:

- Anyone can be a node of LifeBlock
- Anyone can make information available to others, either publicly or through payment
- LifeBlock will have its own token that will allow users to quickly and easily make payments for information
- LifeBlock is understood as a self-sufficient and very scalable network. In the near future, it will grow without the need for maintenance or any investment

Will we be LifeBlock able to “engage” existing databases ?

Answer is YES !!!

There will be a series of intermediate steps to adapt the existing databases to the LifeBlock environment:

- Initially, connectors will be developed between Lifeblock and those databases so that the data is accessible even if they are outside LifeBlock
- At the same time, carrier systems will be developed that will allow the transfer of information with its format within LifeBlock
- Likewise, the systems for updating information can be modified so that it can be carried out directly on LifeBlock

Guaranteeing the long-term sustainability of RI: LifeToken, the token of LifeBlock

- LifeToken is the token of LifeBlock
- It is a private utility token (non ERC-20 or similar). Its value will be always 1 LifeToken=1 €-Life for example, concrete “virtual” monetary value **to get aware Decision Makers of importance of preserving Biodiversity (Circular Economy connection)**
- It will be used by holders to buy private data and by sellers to change them by fiat money or buying private data from other sellers
- It will be an internal token only to facilitate information exchange and monetization. It won´t be available in exchanges and its price will be always the same
- Tokens will be generated automatically by LifeBlock when users demand it. Money received from this generation will be withdrawal to users that want to exchange their existing token by fiat money

Scalability issues

- LifeBlock is a network that **can scale autonomously** and independently due to the characteristics of distributed databases of blockchain architectures
- Each new organization that joins the environment will store its data and make it available to the community, while it will store other data from other organizations using the distributed characteristics of Blockchain
- Therefore it is a network that can grow autonomously, independently and without the need of any central control body
- Additionally, the token management will be carried out automatically by the Blockchain (Smart Contracts) applications of LifeBlock
- LifeBlock tools for aggregation / access / data management will be in the public domain and can be used by any organization for both data availability and navigation/download

Blockchain chosen: Hyperledger Fabric

Blockchain selected for Lifeblock will be Hyperledger Fabric for the following reasons:

- Hyperledger Fabric comes with a full permission system. Organizations can choose who will be able to access their data and at what level they get access. So basically different actors see different things. This is not possible in most of blockchains
- It has the necessary level of maturity for the developments needed for the project
- It provides a very stable development roadmap
- It is free and open. All its resources and tools are free for any entity
- Hyperledger Fabric has been build for a variety of tasks with nodes not being identical. This means there is not a one solution fits all consensus. It lets you choose which consensus algorithm will fit best
- As for now Hyperledger fabric supports NodeJs, Java or Go code which are standard programming languages
- Hyperledger doesn't need computers solving problems day and night, Hyperledger doesn't need people to have cryptocurrency to reach consensus
- **And it is the only way Global Protection Data Regulations –GPDR– !!!**

PART III

A future demostration case: Non-indigenous & Invasive Species (NIS)

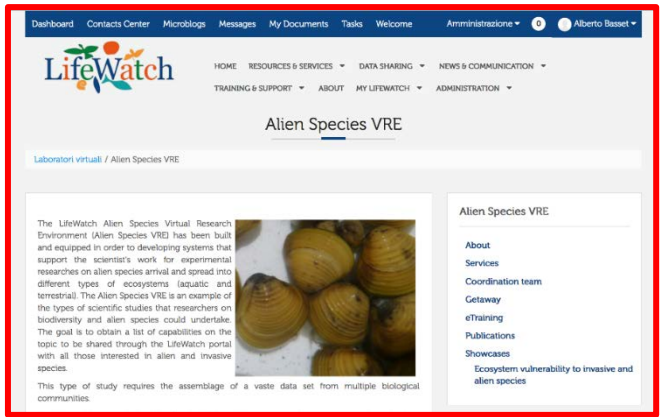
Since non-indigenous and invasive species are considered a major threat to biodiversity, the knowledge produced on their current and future distribution, on their impact on native species and genetic diversity, ecosystem functioning and services as well as on species of economic relevance is of paramount importance to institutions, administrations and managers in finding solutions to major environmental problems at different scales.

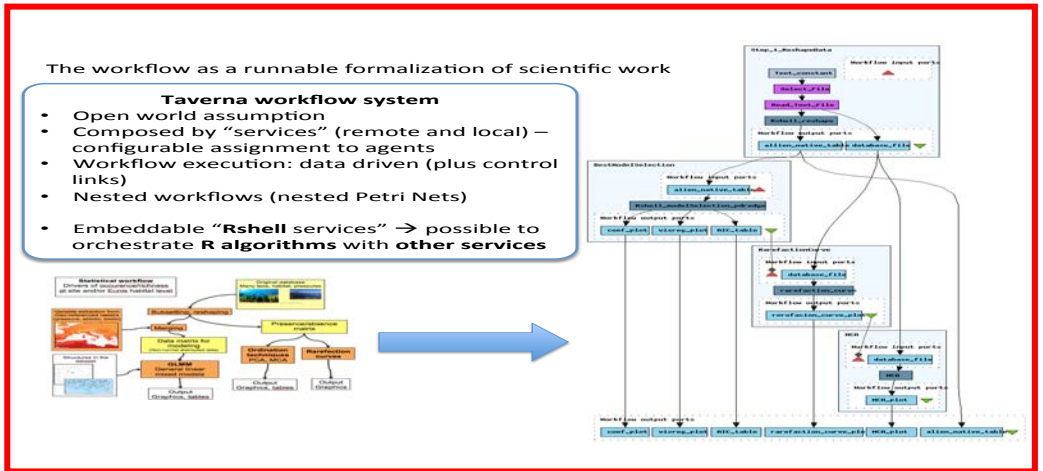
- In fact, biological invasions are listed as one of the five major causes of biodiversity loss, alongside habitat destruction, over-exploitation, climate change and pollution (Millennium Ecosystem Assessment, 2005), and measures to control their introduction and establishment are urgently needed (Aichi Target 9; CBD 2010).

- A number of key issues are still open and challenging for both the research and the management fields. They deal with the phenotypic traits affecting invasiveness of species, their pathways of range expansion at the global scale, their impact on community organisation, ecosystem functioning, services and socio-economic benefits, as well as on the habitat and ecosystem type vulnerability to alien and invasive species and the core factors determining their vulnerability at the global scale.

The European and global scale databases available on alien and invasive species distribution, abundance and traits (e.g., DAISIE http://www.europe-aliens.org/ ; EASIN and its Data Partners https://easin.jrc.ec.europa.eu/easin; IUCN GISD http://www.iucngisd.org/gisd/;WoRIMShttp://www.marinespecies.org/introduced/online_sources.php; AquaNIS, the aquatic non-indigenous species database http://www.corpi.ku.lt/databases/index.php/aquanis/ ; European Regional Databases https://www.nobanis.org/regulation-on-ias/other-resources/regional-databases/; ELNAIS database https://elnais.hcmr.gr/)…

…and the tools, e-Services & VREs already developed and operational on LifeWatch ERIC national nodes for data resources integration, analysis and modelling offer an unique opportunity for a LifeWatch-ERIC Proof of Concept on this topic of such great scientific, social and economic interest.





How will it work?

- LifeBlock will be used to manage the identity and storage of data related to non-indigenous and invasive species (NIS).
- Being a distributed network, data related to NIS will be stored locally with full traceability.
- This will allow to provide data on any NIS about geographic location and distribution with possibilities of temporary scaling.
- The use of LifeBlock will allow any identified user to add NIS data in real time and make these available to VREs and other community organizations
- Data will be stored with characteristics of identification, traceability, geographic location and temporality and may be done by any person or entity registered in LifeBlock.

Advantages of using LifeBlock

- Traditional databases store data, but they do not describe how they got there.
- LifeBlock will contain, in a native way, information about the origin of the data, its traceability, generation, geographical location and the entity that generated it.
- The data, being a distributed database, are automatically available to any entity.
- Fault tolerance of the network is excellent as it does not depend on any central node.
- The maintenance cost of this great registry infrastructure for NIS will be practically zero since the nodes will reside in the interested organizations and the communication will be made through public networks without dependency
- Any entity can join at any time the data network as a generator and/or data consumer
- There will be a library of APIs for connection and BigData tools for data management and search

PART IV
Some conclusions

- ✓ **LifeBlock will deploy an “anti-tampering” provenance and traceability e-Services based on Blockchain technologies, the so-called “LifeBlock" (LW ERIC Blockchain) platform.**

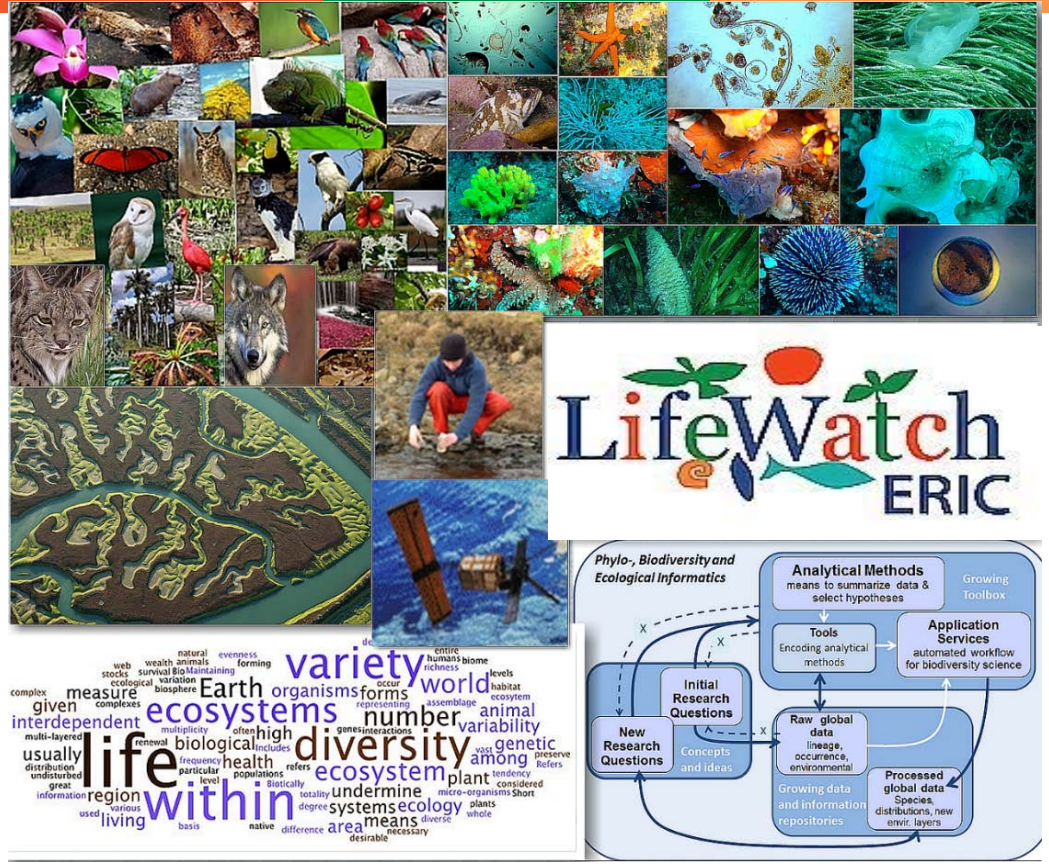
- ✓ **LifeBlock will also guarantee the availability of FAIR-compliant data. This new approach represents a change in accessibility of Big Data for biodiversity research and also creates a new type of e-Science panEuropean RI facilities.**

- ✓ **LifeBlock will contain, in a native way, information about the origin of the data, its traceability, generation, geographical location and the entity that generated it. Its use will allow any identified user to add NIS data in real time and make these available to VREs and other community organizations. There will be a library of APIs for connection and BigData tools for data management and search, among other available e-Services.**

- ✓ **Valuation of Biodiversity Ecosystem Functioning + Species Traits -> Ecosystem Services in a distributed way. Supporting Circular Economy mechanisms, by also respecting GPDR.**

- ✓ **Non-indigenous & Invasive Species (NIS) will be one of the first demonstration case uses to put LifeBlock in practice.**

- ✓ **NOTE: Do not worry about energy consumption levels (e.g. Bitcoin case). There are new and efficient consensus mechanisms in this regard !.**



Thank you very much !

https://www.lifewatch.eu

cto@lifewatch.eu