RIECS concept

Deliverable 2.1:

Challenges assessment

Issue date V2: 22/12/2025

Work Package Leader: CSIC

Project Title: Towards a pan-European Research Infrastructure for Excellent Citizen Science

Project Acronym: RIECS-Concept

Contract Nº. 101188210

Programme: Horizon Europe (HORIZON)

Topic: Research infrastructure concept development

HORIZON-INFRA-2024-DEV-01-01

Project start date: 01 January 2025

End date: 31 December 2027

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that

may be made of the information it contains.

Deliverable title	Challenges assessment
Deliverable number	2.1
Deliverable version	1.0
Contractual date of delivery	31/12/2025
Actual date of delivery	V2. 22/12/2025
Nature of deliverable	R
Dissemination level	PU
Work Package	2
Task(s)	Challenges assessment
Partner(s) responsible(s)	CSIC
Author(s)	Karen Soacha-Godoy, Ana Álvarez & Jaume Piera

Description	
Key words	Citizen science, research infrastructures, participatory information systems

Reviewers

Reviewer(s) name	**Reviewer(s) organisation**
Linda See	IIASA
Franziska Stressmann	ECSA
David Cuartilles	MAU
Fermín Serrano	Ibercivis

Version History

Version	**Date**	**Comments**
0.1	15/06/2025	Preliminary version, content, structure and initial inputs
0.2	21/07/2025	First draft discussed in WP2 meeting
1.0	30/10/2025	First version of the deliverable to review
2.0	22/12/2025	Deliverable second version

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

How to cite:

Soacha-Godoy, K., Alvarez, A., & Piera, J (2025). Challenges assessment | Deliverable D2.1 | V2.0. RIECS-Concept Consortium: Towards a Pan-European Research Infrastructure for Excellent Citizen Science. Horizon Europe Grant No. 101188210. doi: https://doi.org/10.5281/zenodo.17841819

ACRONYMS

Acronym	Description
AI	Artificial Intelligence
API	Application Programming Interface
CARE	Collective Benefit, Authority to Control, Responsibility, Ethics (data governance principles)
CAP	Common Agricultural Policy
CC	Creative Commons
CO	Citizen Observatory
CSV	Comma-Separated Values
CS	Citizen Science
CSP	Citizen Science Platform
DIONE	Complementing EO data with farmer-based monitoring to inform CAP regulations
DOI	Digital Object Identifier
DwC	Darwin Core (biodiversity data standard)
EC	European Commission
ECSA	European Citizen Science Association
EGI	European Grid Infrastructure
EOSC	European Open Science Cloud
EO	Earth Observation
ERIC	European Research Infrastructure Consortium
ESFRI	European Strategy Forum on Research Infrastructures
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable (data principles)
FHIR	Fast Healthcare Interoperability Resources (health data standard)
FP7	Seventh Framework Programme (EU funding programme)
GA	General Assembly
GDPR	General Data Protection Regulation
GBIF	Global Biodiversity Information Facility
GEOSS	Global Earth Observation System of Systems
GHGS	Greenhouse Gases
GitHub	Distributed version control and code-sharing platform
HISP	Health Information Systems Programme

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

HPC	High-Performance Computing
HORIZON	Horizon Europe Research and Innovation Framework Programme
HMIS	Health Management Information System
ICOS	Integrated Carbon Observation System
ICT	Information and Communication Technology
IPR	Intellectual Property Rights
JRC	Joint Research Centre (European Commission)
JSON	JavaScript Object Notation
KPI	Key Performance Indicator
LMIC	Low- and Middle-Income Countries
NGO	Non-Governmental Organization
OECD	Organisation for Economic Co-operation and Development
ODbL	Open Database License
OGC	Open Geospatial Consortium
OSM	OpenStreetMap
OSS	Open Source Software
PC	Project Coordinator
PM	Particulate Matter
R&I	Research and Innovation
REINFORCE	REsearch INfrastructures FOR Citizens in Europe
RI	Research Infrastructure
RRI	Responsible Research and Innovation
SLA	Service Level Agreement
SH	Stakeholder
SSO	Single Sign-On
STAplus	SensorThings API Plus
STILT	Stochastic Time-Inverted Lagrangian Transport (ICOS model)
TRL	Technology Readiness Level
TRUST	Transparency, Responsibility, User Focus, Sustainability, Technology (repository principles)
UI	User Interface
UNESCO	United Nations Educational, Scientific and Cultural Organization
URI	Uniform Resource Identifier
UX	User Experience
XML	Extensible Markup Language

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Table of contents

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

List of figures

List of tables

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Datasets and repositories

Datasets	Link
Inventory of european citizen science digital platforms	Inventory_consolidated*
Characterization of european citizen science digital platforms	Characterization*
Matrix B: Needs, challenges, solutions from EUfunded projects	Matrix B

Repositories	Link
Zotero: publications, reports and all the documentation used in the desk research	https://www.zotero.org/groups/5960620/riecs_tech
Github: scripts developed to run data analysis (e.g. characterization of platforms)	https://github.com/pynomaly/RIECS_CS_infrastructures/tree/main

*Datasets remain under embargo while the academic publication is being prepared and will transition to open access once the embargo period expires.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

1. Executive Summary

This executive summary offers an overview of the main findings from the desk research conducted for Deliverable 2.1. Given the complexity of the topic, the points presented here remain high-level. Readers are encouraged to consult the full document, or at least the sections most relevant to their interests, for a detailed understanding.

Deliverable 2.1, Challenges Assessment, provides a foundational analysis of the technical hurdles that must be overcome to build a world-class CS infrastructure. The deliverable identifies a set of major challenges hindering the development of a unified European CS infrastructure. One central issue is the **lack of interoperability among CS systems.** Different projects use widely **varying data formats, metadata standards, and validation methods, making it difficult to aggregate or reuse data across platforms.** Incompatibilities in software and data schemas mean that valuable datasets remain siloed within individual projects. **Furthermore, scalability limitations are noted: many existing CS platforms struggle to handle increasing volumes of data or growing numbers of users, indicating that current architectures may not easily scale to a Europe-wide level. Another challenge is the short project-based lifespan of many platforms—**most CS tools are developed in the context of time-limited projects and lack sustainable maintenance and updates beyond the project's end. This leads to *link rot* and obsolescence, undermining the continuity of services for citizen scientists and researchers.

Beyond these capacity issues, several emerging challenges relate to the evolving nature of data and technology in CS. **The integration of artificial intelligence (AI) and big data techniques into CS projects brings new ethical, legal, and privacy considerations.** For example, using AI to analyze citizen-contributed data can yield powerful insights, but it also raises questions about data protection (e.g. compliance with GDPR), algorithmic transparency, and potential biases. The deliverable notes that guidelines and safeguards specific to AI use in CS are not yet well-established.

Additionally, **citizen engagement and inclusivity gaps (e.g. digital gaps) persist.** Not all demographic groups are equally reached or empowered by current platforms. Some projects lack features for accessibility or multilingual support, which can exclude segments of the public and hinder the *citizen* aspect of CS. **There is also uneven geographic and disciplinary coverage—certain countries and scientific domains have many active platforms, whereas others are under-served or in development, indicating a need to bridge these disparities.**

An overarching challenge is the absence of a common governance model across the myriad CS initiatives. Currently, each platform or project tends to have its own governance and data management practices. **There is no unified framework to coordinate efforts, share resources, or set joint policies at the European level. This lack of coordinated governance means issues like data standards, quality control, and integration are handled ad hoc, if at all, and opportunities for synergy**

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

are missed. For instance, projects in different countries might collect similar environmental data but have no mechanism to consolidate their findings.

The deliverable acknowledges that the organizational and social dimensions of a research infrastructure (such as governance structures, community management, and long-term funding strategies) are crucial, but these aspects are addressed in the deliverable 3.1 focusing on organizational challenges. Deliverable 2.1 remains centered on technical challenges—software, hardware, data pipelines, and digital components—while noting that these technical issues often intersect with social considerations in practice.

1.1 Insights from domain-specific contexts

In addition to general issues, the deliverable considers examples in specific domains (environment, health, and climate) to illustrate unique needs and successful strategies. In the biodiversity field, numerous platforms exist for species observations, but efforts like using the Darwin Core data standard and aggregators such as GBIF show how fragmented datasets can be integrated, in particular the Cos4Bio (integrating biodiversity observations from multiple platforms) and Cos4Env (integrating environmental observations of different nature, water quality, air quality) examples provide evidence that this integration is possible. In the health sector, the long-running DHIS2 platform demonstrates the value of a modular design, open APIs, and strong community support for sustaining a large-scale data system. In climate monitoring, the ICOS ERIC infrastructure highlights how a formal federated model with early institutional backing can secure long-term operations. These cases reinforce the broader lessons for RIECS, underlining the importance of standardization, flexible architecture, community engagement, and formal governance structures.

1.2 Key takeaways: The fragmentation challenge

The analysis identifies fragmentation as the meta-challenge manifesting across five interconnected dimensions:

1. **Technological fragmentation**: Projects develop isolated technical solutions using incompatible technologies, programming languages, and architectures, partly influenced by technological obsolescence. This results in duplicated effort and inability to share components or integrate systems.
2. **Data fragmentation**: Lack of common data standards and formats prevents data sharing and aggregation. While some domains have adopted standards (e.g., Darwin Core for biodiversity), most citizen science data remain in proprietary formats inaccessible to wider use.
3. **Standards fragmentation**: Absence of agreed protocols for data quality, validation, and metadata creates barriers to data reuse and scientific credibility. Each project develops its own quality assurance mechanisms without coordination.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

4. **Resource fragmentation**: Knowledge, tools, and best practices remain trapped within individual projects. Documentation is scattered across various repositories, and valuable lessons learned are not systematically captured or shared.
5. **Community fragmentation:** beyond storing data, CS technologies also sustain communities, learning processes, and relationships between volunteers and scientists, not only the storage of data. When technologies emerge in isolation due to incompatible programming languages, architectures, communication channels, competition among projects and platforms, and incentives linked to funding schemes, the fragmentation reaches the community layer as well. Participants struggle to move between projects, collaboration and engagement weaken, and knowledge exchange becomes difficult.

1.3 Proposed solutions and approaches

To address these challenges, the deliverable suggests several strategic directions for RIECS. First is a **federated, modular infrastructure design that links together existing citizen science platforms rather than replacing or duplicating them**. RIECS would act as a **network layer connecting diverse projects via shared standards and protocols. Adopting common metadata schemas, open APIs, and data principles (e.g. FAIR for data management and CARE for ethical data use) would enable interoperability** so that observations and resources can flow across different tools and communities.

Another key solution is to invest in community building and participatory co-design. This involves **training citizen scientists and developers, providing shared tools and open-source modules,** and actively engaging stakeholders in shaping the infrastructure's features. By strengthening the practitioner community and involving end-users in design and governance, RIECS can ensure the platform meets real needs and earns broad trust. It also means **building capacity to support emerging technologies (such as AI-driven analytics or managing big data)** in a responsible way, so that the infrastructure stays innovative yet ethical and user-centric.

Figure 1 synthesises the main challenges and potential solutions for building a citizen science research infrastructure, grounded in lessons from close to twenty years of technological evolution in citizen science.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

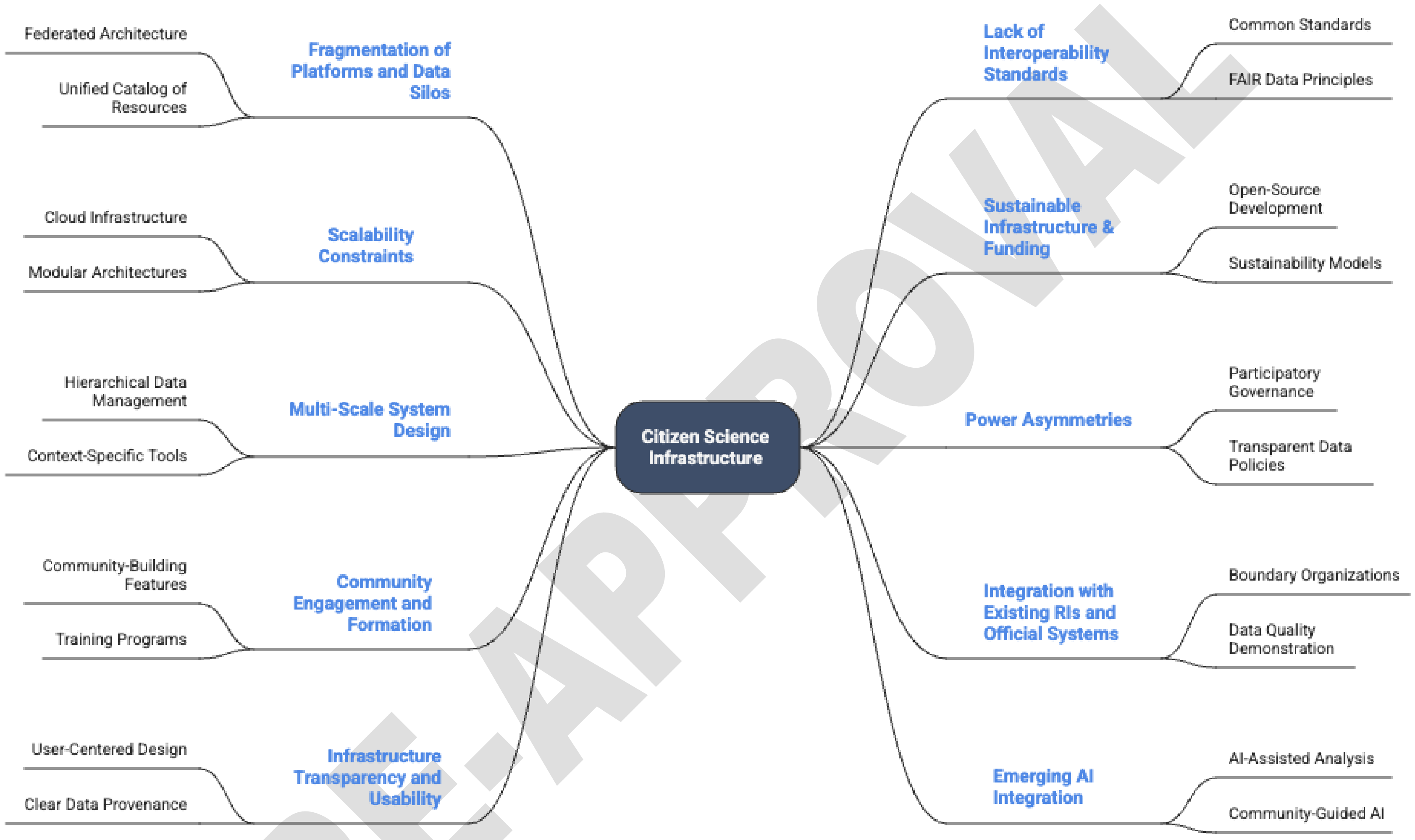

Figure 1 Citizen science infrastructure: Main challenges and solutions

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

2. Introduction

The RIECS-Concept project is conceptualizing the European Research Infrastructure for Excellence in Citizen Science, addressing the growing need for a robust, integrated technical infrastructure to support the transformative potential of citizen science (CS) in European research and public engagement. This deliverable, **D2.1 Challenges Assessment**, serves as a cornerstone component by identifying, analysing, and categorising the technical challenges that must be overcome to establish a world-class research infrastructure for CS.

The rapid evolution of CS across Europe has highlighted both tremendous opportunities and significant challenges within the current technical landscape, where numerous tools, platforms, and technologies have emerged through EU-funded programmes such as FP7 and Horizon 2020, yet the landscape remains fragmented and characterised by an increase in technologically siloed solutions that has resulted in reduced efficiency in data management, and diminished overall impact of CS contributions to scientific knowledge and policy-making. This limitation becomes particularly evident when considering that the absence of shared standards, common metadata practices, and coordinated infrastructures has produced large volumes of citizen-generated data that currently cannot be connected, compared, or reinterpreted across projects, which weakens their scientific value and restricts their usefulness for institutions and policymakers seeking robust and scalable evidence [1], [2].

The scope of D2.1 follows the criteria outlined in Section 4.1 on methodology, which define its thematic and conceptual boundaries.

2.1 Objectives

According to the Grant Agreement, Task 2.1 led by CSIC and the participation of ECSA, UNIMIB, CSZ, MAU and IBE the primary and specific objectives of the Challenges Assessment in the deliverable 2.1 are: to **analyse the current technical landscape for Citizen Science by examining both domain-specific challenges relevant to environmental observations, health, and climate change, and cross-domain technical issues affecting CS infrastructure more broadly**, while **methodically uncovering, understanding, and classifying the technical barriers associated with developing a robust research infrastructure for excellent CS.** This will be achieved through extensive desk research drawing on academic and grey literature, industry reports, and existing case studies to identify the challenges that influence the establishment and operation of CS technical infrastructure across diverse applications, and to **examine existing solutions with attention to development status, licensing, support structures, GDPR compliance, and alignment with FAIR principles**. The assessment will further **identify critical gaps in current tools and infrastructures, highlighting where available solutions do not adequately respond to community needs**, and will culminate in the

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

construction of a systematic matrix that organises each challenge together with corresponding solutions to provide a clear and structured overview of the present technical landscape.

The findings from this process will serve as foundational input for Task 2.2, which focuses on defining the technical requirements for RIECS, while being systematically contrasted with inputs from extensive stakeholder engagement activities (T4.2, T4.3 & T4.4) to ensure that the technical analysis remains grounded in real-world user needs and community requirements.

2.2 Structure of the document

Deliverable 2.1 comprises three main components: the main analytical document, four datasets, and two collaborative repositories. The datasets include: (1) an **Inventory of European CS digital platforms** cataloging all identified platforms, tools, and services with their descriptions, links, and categorizations; (2) a **detailed characterization of European CS digital platforms** expanding information of a subset of platforms including their development status, licensing agreements, support structures, GDPR compliance, and adherence to FAIR principles; (3) a **Matrix of needs, challenges, and solutions** derived from EU-funded projects that maps identified technical challenges to existing and proposed solutions; and (4) a **Body of Knowledge** compiling all documentation used in the desk research phase. These resources are supported by two repositories: a **Zotero group repository** (RIECS group) providing open access to all the documentation reviewed, and a **GitHub repository** containing scripts and analytical tools developed for platform characterization and data analysis.

The document itself is structured in eight sections that synthesize findings from the extensive desk research and analysis: (1) **Objectives** defining the assessment goals, (2) **Scope** delineating boundaries and focus areas, (3) **Key Takeaways** highlighting critical findings, (4) **Guiding conceptual approach** establishing the theoretical framework, (5) **Methodology** detailing the research processes, (6) **Mapping of CS platforms, tools and services** presenting the technological landscape analysis, (7) **Challenges and solutions** examining both cross-disciplinary technical challenges and domain-specific challenges in environment, health, and climate sectors..

2.3 Use of the document and future updates

Deliverable 2.1 is the starting point for a shared understanding of the technical challenges that the RIECS infrastructure will need to address, drawing on the trajectory of CS technologies over the past two decades. Although submitted in M12 (December 2025), it is conceived as a living document.

Its elements — the challenge matrix, the identification of gaps, and the mapping of platforms, services, and tools — will be reviewed and enriched through workshops, interviews, and other exchanges with RIECS stakeholders, with technology providers as a priority group. These iterative

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

processes will shape future project outputs in multiple formats, including reports, academic articles, and technical briefs.

The document also feeds directly into ongoing tasks, particularly the service catalogue (Task 2.3), the requirements analysis (Task 2.2), and the organisational challenges work (Task 3.1). CSIC will continue leading the development of subsequent products related to technical challenges, ensuring continuity and coherence across these efforts. Deliverable 2.1 should therefore be seen as an initial reference point that will evolve throughout the project and guide future discussions and refinements.

3. Framework for assessing technical challenges in CS research infrastructures

This section sets the foundation for the technical challenges assessment by defining what constitutes a research infrastructure (RI) for CS, clarifying the type of infrastructure addressed, and situating it within established frameworks such as the European Strategy Forum on Research Infrastructures (ESFRI) and global open science initiatives like the ones lead by the Organization for Economic Cooperation and Development (OECD) and the United Nations Education, Scientific and Cultural Organization (UNESCO). The **theoretical lens** establishes a conceptual foundation by drawing on ESFRI definitions and recent scholarship, framing CS as participatory infrastructure. The **policy lens** clarifies how European and global policy frameworks define platform functions, domains, and data governance relevant to CS infrastructure. The **operational lens** focuses on the types of infrastructure—data, computing, and digital—relevant to RIECS, referencing ESFRI classifications and alignment with FAIR, CARE, and TRUST principles. These lenses guide the scoping and classification of technical challenges across domains, ensuring coherence with both RI standards and open science values.

3.1 Theoretical perspective: Insights from infrastructure frameworks

A central concern for RIECS-Concept is how a RI is defined. ESFRI, reflecting the Horizon 2020 regulation (EU No 1291/2013), defines RIs as ***facilities, resources and services used by research communities to conduct research and foster innovation, including major equipment, knowledge-based resources like collections and archives, e-infrastructures (data and computing systems, networks) and other tools essential to achieve excellence in research and innovation [3]***. This definition highlights the breadth of infrastructures: from physical laboratories and instruments to digital platforms and data repositories. In practice, RIs can take different organizational forms – centralized facilities, distributed observatories, virtual platforms, or networks of resources – as long as they support excellence in research and innovation[3].

This foundational definition grounds ESFRI's work, including the Landscape Analysis 2024 [4], which guides RI development across Europe. The Landscape Analysis identifies strengths and gaps in the RI

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

ecosystem and emphasizes integration, responsiveness to future needs, and cross-domain capability as key features of next-generation infrastructures. This strategic view serves as an essential reference for conceptualizing RIECS within the European research landscape.

At its core, a RI is more than just hardware or IT – it is the ensemble of facilities, resources, and services enabling a community to produce new knowledge. Crucially, modern perspectives treat RIs as socio-technical systems [5] . Infrastructure is not merely a technical artifact but is deeply embedded in social contexts and practices. As Star and Ruhleder (1996) observed in a classic socio-technical analysis, infrastructure is a fundamentally relational concept, something that *becomes real infrastructure in relation to organized practices*, emerging through use and integration in daily work [6]. In other words, what counts as infrastructure depends on who uses it and how [6]. This insight guides our conceptual approach: **the CS infrastructure must be conceived not just as a technical platform, but as an enabler of collaborative practices, connecting scientists, citizens, and data in a seamless way.** The notion of *knowledge infrastructure* is especially relevant – a system of people, technologies, and institutions that together facilitate knowledge production and sharing across traditional boundaries.

CS plays a transformative role in such transdisciplinary knowledge infrastructures. Active involvement of non-professional scientists in data collection, analysis, and problem-solving allows CS to blur the line between researcher and public, creating a bridge between formal scientific institutions and society. Indeed, CS platforms and observatories themselves are now viewed as *rapidly expanding RIs… that support the growth of CS*, significantly boosting data gathering capacity and public engagement [7]. **This implies that a dedicated CS RI must accommodate a wide diversity of actors and data streams, spanning multiple scientific domains and community contexts. It should support co-creation of knowledge, where volunteers and researchers work together on scientific questions.** For example, UNESCO's Open Science Recommendation highlights the importance of *platforms for exchanges and co-creation of knowledge between scientists and society*, calling for sustained support to CS organizations as part of open science infrastructure [8]. Thus, **our guiding approach conceptualizes the CS infrastructure as a socio-technical backbone for transdisciplinary research – one that is technically robust while socially inclusive, facilitating excellence in science through broad participation.**

Another concept that is fundamental for D2.1 is what we mean by *CS platform*. Recent research has revealed a proliferation of overlapping terms in this domain – A review by Soacha-Godoy et. al. (2025) **found 98 distinct terms referring to CS platforms or observatories, indicating conceptual fragmentation [7]**. To bring clarity, the authors have suggested a **purpose-based taxonomy of CS platforms that categorizes platforms by their foundational mission [7]. This taxonomy defines nine distinct platform categories (e.g., project discovery portals, knowledge resource hubs, on-site data collection apps, online data analysis portals, educational gamification platforms, decision-support systems, etc.), each representing a specific way that a platform supports CS projects** as shown in Figure 2. For example, *project discovery platforms* serve as directories helping volunteers find and join

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

projects, thereby boosting participation (e.g., SciStarter). Such a taxonomy is valuable in the context of Task 2.1. It provides a shared language to describe CS infrastructures and ensures that when we assess technical challenges, we consider the full range of platform types and purposes. For instance, both data-centric infrastructures (data collection, curation, sharing) and engagement-centric ones (community building, learning, outreach) are considered, as they may face different technical requirements.

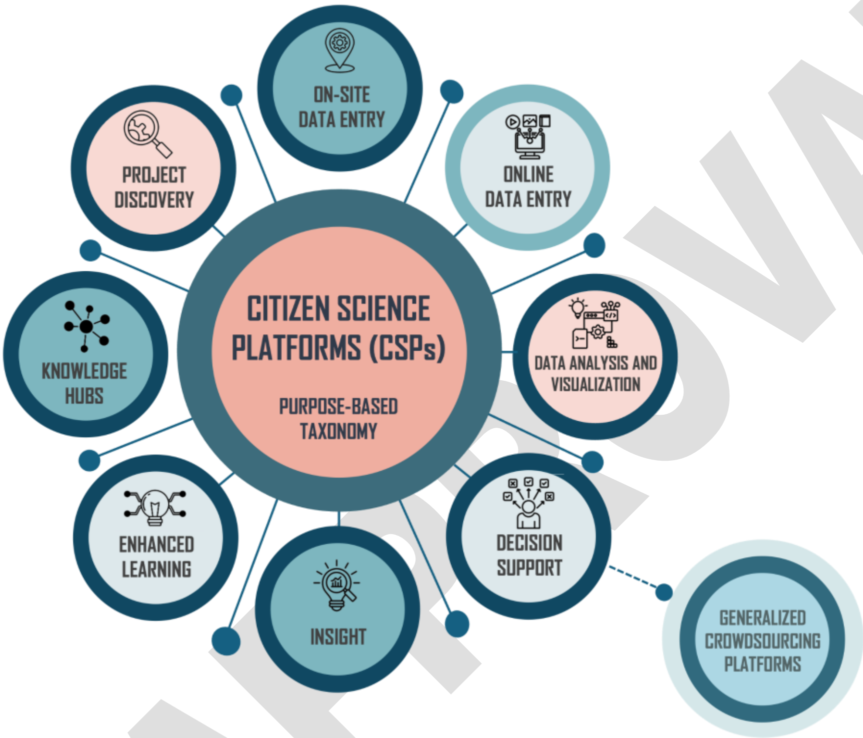

Figure 2 Citizen science platform (CSP) purpose-based taxonomy. Nine categories defined for classifying CSPs. Source: Soacha-Godoy, K. et. al. (2025). *Research Infrastructures in Citizen Science: State of Knowledge and Taxonomic Framework As a Pathway to Sustainability | Citizen Science: Theory and Practice*. https://doi.org/10.5334/cstp.831

Building on the preceding taxonomical framing, Soacha-Godoy et al. (2025) [9], working within the RIECS-Concept development, highlight the role of one of the categories and present an analytical elaboration focused on *on-site data entry*. The authors introduce a functional and structural characterisation of citizen observatories (COs). **While traditionally viewed as time-bound or project-specific monitoring initiatives, COs are redefined as lasting participatory RIs.** Drawing on and extending the ESFRI definition, the authors argue that COs fulfil all core functions of RIs when approached from a socio-technical and participatory science perspective. They offer stable facilities (e.g., digital platforms, sensors), essential services (including data collection, validation, and visualisation), and governance mechanisms that facilitate research and innovation through citizen involvement. The authors further stress the processual and evolving nature of infrastructures in CS, noting that these systems are co-constructed over time through iterative socio-technical practices, community participation, and shared knowledge processes.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

In practical terms, this perspective allows us to **assess COs and CSPs not merely as discrete tools, but as potential building blocks of a distributed European RI.** For example, the authors propose a taxonomy of infrastructure functions displayed in Figure 3—including data management, participatory governance, capacity building, and interoperability—that parallels both ESFRI's operational criteria and the OECD's ecosystem-based RI thinking. Such an approach is directly applicable to the scope of RIECS, which aims to integrate these fragmented initiatives into a federated, technically coherent, and socially legitimate infrastructure.

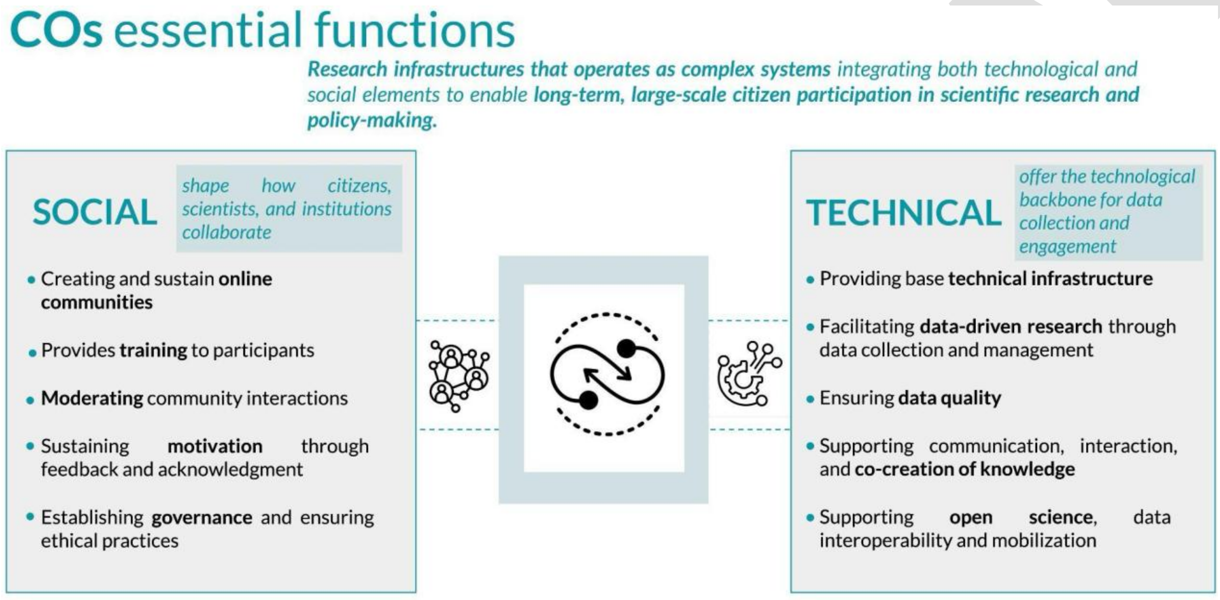

Figure 3 Essential functions of Citizen Observatories (COs). Source: Soacha-Godoy, K., López-Borrull, A., Serrano, F., & Piera, J. (2025). The Backbone of Participatory Science: Reframing Citizen Observatories as Research Infrastructures. Sustainability, 17(10), Article 10. https://doi.org/10.3390/su17104608

Taken together, these theoretical foundations reinforce the legitimacy of conceptualizing RIECS as a participatory, distributed, and evolving RI. They provide the analytical basis for identifying technical challenges not as isolated bottlenecks, but as critical junctures where architecture, governance, ethics, and social engagement intersect. **Task 2.1 will build on this multifaceted definition to provide the necessary inputs for conceptualizing an infrastructure that is designed as a functional, scalable, and inclusive ecosystem responsive to both scientific and societal needs.**

3.2 Policy perspective: Scope and strategic alignment

In framing the technical challenges, we draw on established frameworks and analyses of RIs at the international level. European policy initiatives increasingly recognize CS as a pillar of research and innovation, which in turn influences how CS infrastructures are scoped. The European Commission's Open Science policy explicitly names *Citizen Science* as one of the eight ambitions of Open Science, aiming to recognize citizens as creators of knowledge and to integrate their contributions into the

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

scientific enterprise. A dedicated European CS RI would build on these policy directions, demonstrating how an infrastructure can make possible ***free of charge services***, ***resources and expertise for all researchers and citizen scientists*** as envisioned in policy roadmaps [10]. In fact, in 2023, the EU funded project REINFORCE (REsearch INfrastructures FOR Citizens in Europe) released a ***Policy Roadmap on Research infrastructures for citizens science in Europe***, which stressed the need for open access, interoperability, and sustainability as pre-conditions for normalizing CS usage in RIs [11]. It also points out current gaps – for instance, many RIs are not readily accessible to the public or lack interfaces for citizen contributions – and calls for policy measures to open up infrastructures to non-traditional actors while ensuring data quality and governance issues (i.e., ownership, privacy, etc.) are addressed [11].

The ESFRI Landscape Analysis 2024 provides a strategic overview of the European RI ecosystem and emerging trends. Notably, the Landscape Analysis (LA) identifies strengths and gaps in Europe's RIs, emphasizing integration and responsiveness to future needs [4]. The ESFRI LA 2024 is positioned as a framework for the next ESFRI Roadmap, outlining critical gaps, synergies, and opportunities for clustering and interoperability across RIs. It explicitly aims to foster a *fully functional and interoperable European RI ecosystem* by exposing deficiencies and proposing improvements in areas like accessibility, networking, and cross-fertilization of services. **This perspective informs our approach by emphasizing that a CS RI should not be built in isolation – it must integrate with the broader ecosystem (for instance, linking to the European Open Science Cloud and domain-specific infrastructures like OPERAS, LifeWatch or ELTER-RI) and address known gaps (e.g., data interoperability, service fragmentation).** The LA's inclusion of a cross-domain trends and challenges section reinforces the idea that infrastructures must increasingly support interdisciplinary and transdisciplinary research, which is central to CS.

CS infrastructure intersects with existing ESFRI domains and EU research priorities. ESFRI categorizes RIs into domains (e.g., Environment, Health & Food, Social & Cultural Innovation, etc.) [4], yet CS cuts across all these thematic areas. A platform supporting citizen participation in science must be inherently cross-domain. **Our challenge assessment will thus consider requirements for domain interoperability – for example, the ability to manage biodiversity observations (environmental domain) alongside health data or cultural heritage contributions – reflecting the interdisciplinary scope of CS.** This aligns with ESFRI's recognition that some infrastructures provide *thematic or interdisciplinary services* and the need for deeper integration across domains in a fully functional European RI ecosystem. Furthermore, European open data policies (such as the INSPIRE directive for environmental data, or the general push for FAIR data in Horizon Europe) will shape standards that a CS infrastructure should adopt.

Similarly, OECD's work on RIs and open science guides our theoretical framing. A recent OECD policy report (2023) states that **tackling complex scientific and societal challenges requires RI ecosystems – dynamic partnerships across infrastructures, disciplines and borders – because no single infrastructure can provide all the tools needed for today's interdisciplinary problems [12]** . In other words, connectivity and complementarity among infrastructures are key. The OECD calls for

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

broadening user communities and incentivizing collaboration across disciplines, embedding this *ecosystem thinking* into strategic planning [12]. **For a CS infrastructure, this suggests that we should design for openness and interoperability from the outset, allowing data and tools to flow between our CS platform and other RIs, and allowing new user groups (e.g., citizen scientists, community organizations, educators) to access resources traditionally confined to professional labs.** The OECD's emphasis on sustainable funding and inclusion also reminds us that technical solutions must be coupled with governance models that encourage cooperation (e.g., shared platforms, standards) and ensure long-term viability.

The UNESCO Recommendation on Open Science (2021) [8] provides additional theoretical grounding, especially regarding principles for infrastructure development. UNESCO urges Member States to *invest in open science infrastructures and services* and to ensure these are accessible for all and as interoperable as possible. Crucially, the Recommendation specifies that open science infrastructures should follow core specifications like the FAIR principles (making data Findable, Accessible, Interoperable, Reusable) [13] and the CARE principles (data governance oriented to Collective Benefit, Authority to Control, Responsibility, Ethics) [14]. This aligns the infrastructure with global best practices in data management and ethics. **The open science framework also advocates that infrastructures be community-owned or driven, not-for-profit, and sustainable in the long term** [8]– values highly pertinent to a CS RI, which must center on community participation and trust.

In summary, these international frameworks (EU Open Science Policy, ESFRI, OECD, UNESCO) shape our framing of technical challenges by highlighting several key theoretical imperatives: interoperability and standards compliance (FAIR), ethical and inclusive data practices (CARE), collaboration across disciplines and borders, sustainability, and an ecosystem mindset rather than siloed development. Our assessment in Task 2.1 will align with these dimensions, ensuring that the identified challenges and proposed solutions resonate with the broader landscape of RI development.

3.3 Operational perspective

Finally, we ground our framework in an operational perspective, outlining how we define the technical challenges themselves and how this will guide the methodology of Task 2.1. From a pragmatic standpoint, we focus on several key dimensions of challenge: **architecture, scalability, standards (data and software), and socio-technical integration.** Each of these corresponds to practical requirements for a robust CS RI:

Architecture & Modularity: The infrastructure's architecture must be designed for interoperability and flexibility. Given the diversity of CS tools (i.e., apps, sensors, databases, analytics platforms), an effective architecture will likely be distributed and modular, allowing integration of heterogeneous components and services. The RIECS-Concept envisions an architecture that *leverages both citizens' resources and existing scientific resources (platforms, data collections, RIs)*, essentially a federation or network-of-networks [28]. A challenge here is defining the right interfaces and middleware to connect volunteer-provided inputs (e.g., observations from a mobile app) with institutional systems

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

(e.g., a national data repository). This also involves addressing security, user management, and knowledge management across the architecture. Scoping these architectural challenges will dictate what kind of technical blueprint and standards are needed (e.g., use of APIs, microservices, cloud infrastructure, etc.).

Scalability & Performance: CS projects can engage tens of thousands of participants and produce massive datasets (e.g., biodiversity records, astronomical images). The infrastructure must scale in terms of users, data volume, and processing load. Scalability challenges include ensuring cloud or High Performance Computing (HPC) resources are available to analyze crowdsourced data in near real-time, and that the system can handle peak loads during mass participation events. We will assess technologies for scalable data ingestion, storage and retrieval (for instance, leveraging big data frameworks or distributed databases). This ties closely to the requirement of pan-European reach – the RI should serve users across Europe (and potentially globally), **which means addressing multilingual support, network distribution, and robust performance and uptime.** Scalability is not purely technical; it also has a cost dimension (i.e., sustainable funding for expansion) that will be considered as part of operational feasibility.

Data and Software Standards: Ensuring interoperability is a foundational technical challenge. Adhering to community standards for data formats, metadata, and protocols will make it easier to integrate with other infrastructures and to reuse CS data in research. We will be guided by the FAIR principles – for instance, evaluating how to make citizen-contributed datasets more findable (through common metadata and registries), accessible (through open APIs and clear licenses), interoperable (using controlled vocabularies and data models), and reusable [8]. Similarly, for software, adopting open standards and open-source components will be emphasized, as per UNESCO's recommendation that digital infrastructures use open technologies and facilitate *community control* of tools [8].

Socio-Technical Integration: Perhaps uniquely for a CS infrastructure, the technical design must be tightly coupled with user engagement and social factors. We frame socio-technical integration as a core challenge: how to design systems that are user-friendly, inclusive, and support collaboration between scientists and volunteers. This includes user interface/experience design (lowering barriers for public participants), community tools (e.g., forums, feedback mechanisms, reward systems), and training and documentation for diverse users. It also extends to data governance questions – respecting privacy, attributing contributions, and implementing ethical practices when volunteers share data. Here we will invoke principles like CARE (which reminds us that data practices should ensure collective benefit and ethical use, particularly when involving community data or Indigenous knowledge) [16]. Additionally, we consider the TRUST principles for repositories – Transparency, Responsibility, User focus, Sustainability, Technology – as benchmarks for trustworthy infrastructure operation [28]. For instance, transparency in how data are validated and used, user focus in the design of tools, and sustainability in governance and funding all need to be taken into consideration.

Landscape of infrastructures: From an operational perspective, the ESFRI Landscape Analysis (2024) delineates distinct yet interrelated categories of infrastructures as displayed in Figure 4: **data**

infrastructures, computing infrastructures, and digital infrastructures. Data infrastructures are understood as systems primarily concerned with the long-term stewardship, sharing, and preservation of research data, covering both curated repositories and distributed data services that support FAIR principles and open access policies. **Computing infrastructures**, by contrast, provide the high-performance computational capacity and processing capabilities necessary for advanced modelling, simulation, and large-scale data analytics. These include high-throughput computing clusters, cloud services, and specialised platforms that enable intensive scientific computation. **Digital infrastructures** integrate both data and computing layers but go further in supporting broader interoperability, access, and scalability across disciplines and communities. In ESFRI's classification, digital infrastructures act as enablers of cross-domain integration, combining data services, computational tools, middleware, and user interfaces into cohesive, reusable, and accessible ecosystems.

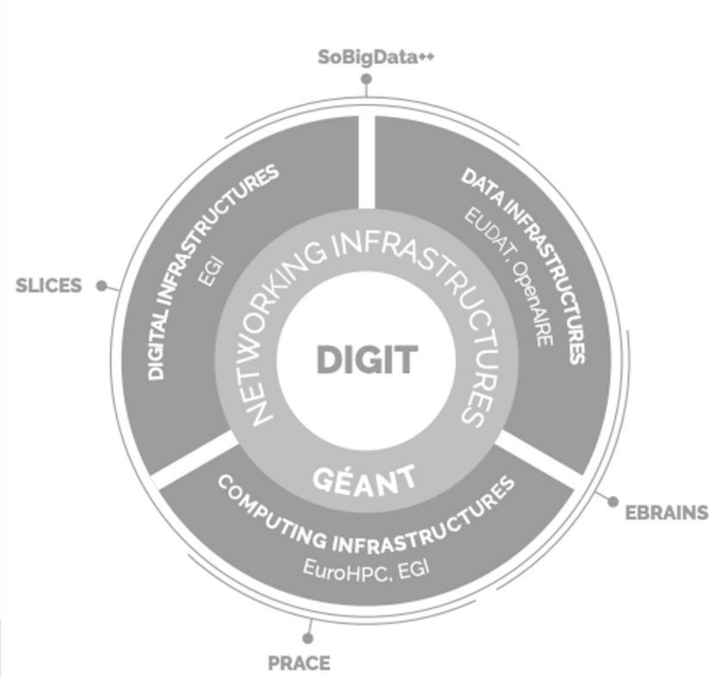

Figure 4 The Landscape of the Data, Computing & Digital Research Infrastructures domain. Source: ESFRI. (n.d.). European Strategy Forum on Research Infrastructures ESFRI LANDSCAPE ANALYSIS 2024. Retrieved April 24, 2025, from https://landscape2024.esfri.eu/media/coqdoq0q/20240604_la2024.pdf

This operational perspective is not just a list of technical topics; it directly shapes our methodology for the challenges assessment. We have structured Task 2.1's analysis around the above dimensions – examining architecture, scalability, standards, and socio-technical features in turn – to ensure wide-ranging coverage. For each category, we will gather evidence through a literature review and case studies of existing platforms. Finally, our pragmatic lens keeps sight of real-world infrastructure requirements and domain-crossing capabilities that the CS RI must fulfill. The outcome of Task 2.1 will ultimately support the scoping of a feasible infrastructure design, so each challenge we assess will be linked to potential solutions for the future system.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

4. Methodology

The desk research methodology of D2.1 is divided into three main components each of them with their corresponding subcomponents as displayed in 5: (1) the mapping of CSPs, tools and services; (2) the challenges assessment and (3) the identification of gaps. The three components are interconnected but are designed to produce specific outcomes.

In addition to the in-depth review of the documentation listed in the body of knowledge dataset and available in the RIECS Zotero group, interviews with RI managers and technicians were conducted to complement the desk research.

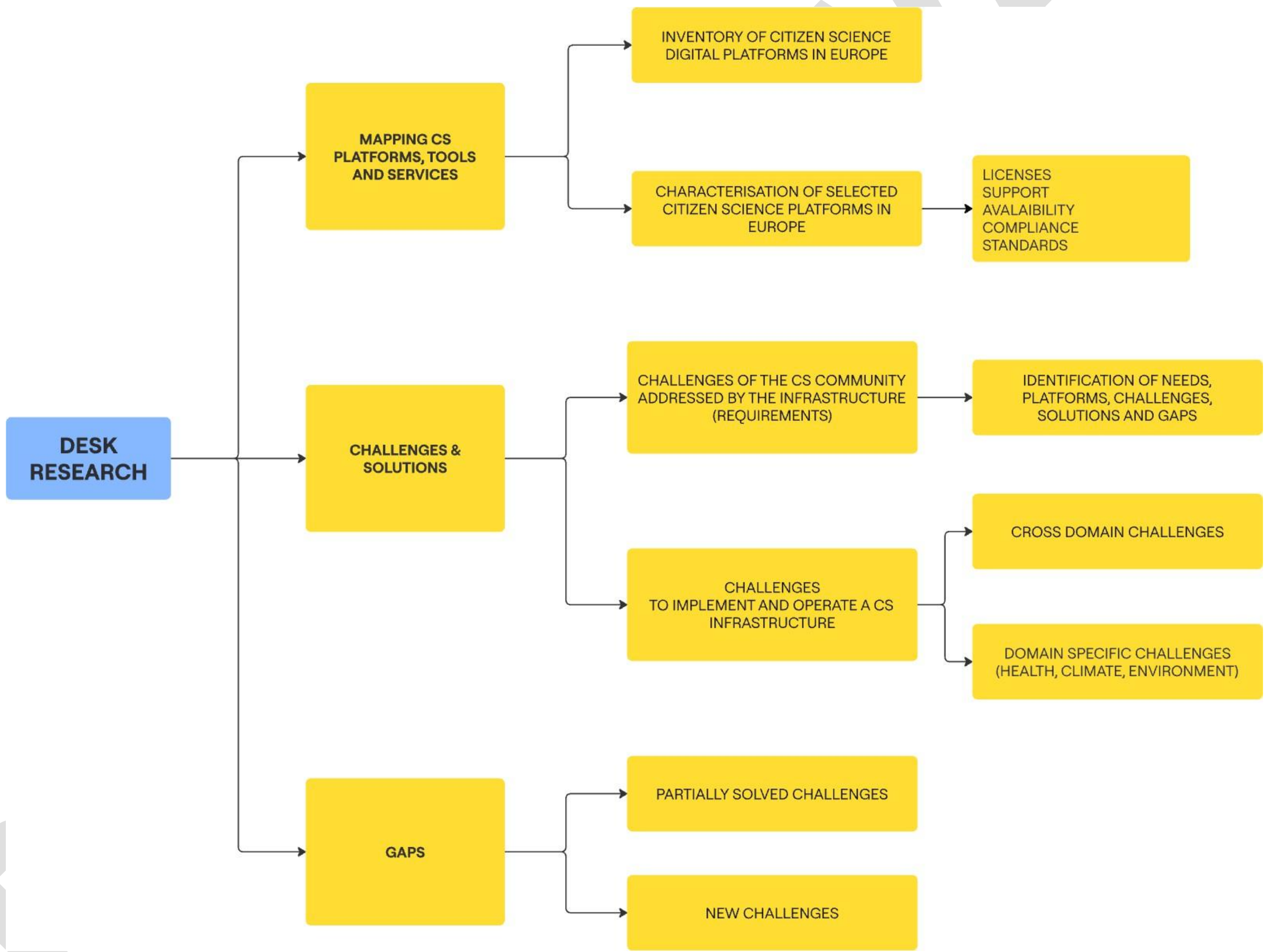

Figure 5 Desk research components.

The **mapping of platforms, tools and services fulfils two main aims.** The first is a **broad review of CSPs in Europe,** intended to showcase the information available about digital platforms developed for CS across multiple domains. The second mapping focuses on the **platform characterisation,** providing general assessment of the platform status in terms of licenses, GDPR compliance, FAIR

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

principles adherence, and other criteria. In the following sections, the methodology for each approach is explained in detail.

The **challenge assessment focuses on two main areas.** First, we identified the **challenges of the CS community that are addressed** or need to be addressed by the infrastructure, defining the requirements alongside the platforms created to respond to these needs. Second, we examine the **challenges to implement and operate a CS infrastructure,** considering also the solutions created by the developers of platforms, tools and services to address those challenges.

In the third component of the **gap analysis**, we identified partially solved challenges and emerging issues beyond those documented in the EU-funded projects. In this section, we have relied on literature, reports, interviews, and experience from the Consortium to highlight challenges that were not mentioned in the previous section.

4.1 Scope

Based on the requirements established in the Grant Agreement for Deliverable 2.1 this document defines the following scope:

- The mapping focuses on CS platforms (CSPs) specifically designed for the CS field. General crowdsourcing platforms like OpenStreetMap are mentioned as part of the broader landscape but not included in detailed mapping.
- The desk research prioritizes documentation spanning the last 10 to 12 years, with particular emphasis on the period from the initiation of FP7 EU-funded projects to the present day. Given the rapid pace of technological advancement over the past decade, analysis of challenges beyond this temporal scope may yield insights with limited applicability to contemporary CS infrastructure requirements.
- The assessment covers CSPs from the three main domains prioritised for the RIECS-Concept: environment (including biodiversity), health, and climate (including Earth observation platforms). These domains were chosen for strategic, scientific, societal, and policy-aligned reasons as initial demonstration domains, but the infrastructure concept is intentionally cross-disciplinary and domain-neutral. They also contain a significant share of existing CSPs, which supports their relevance for this stage of the analysis.
- The challenges assessed in the document serve as valuable entry points to address the wider ecosystem of CS and will shape the core capabilities and flexibility of the RIECS-Concept to support additional disciplines in the future. This document remains within these domains and does not map platforms from other fields connected to CS, which will be addressed in data and metadata interoperability deliverables, specifically D3.3 (Data and metadata criteria).
- The assessment of organisational and social challenges of RIECS will be covered separately in deliverable D3.1 (Organisational challenges), ensuring this document focuses exclusively on technical challenges.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

- The technical challenges are focused on technological development in CS, considering software, hardware, sensors, and related devices. Infrastructure in CS also includes laboratories, fab labs, libraries, and other physical infrastructures that provide support to CS activities. Although these broader infrastructure elements are acknowledged as important components of the CS ecosystem, this challenge assessment focuses primarily on the digital and technological aspects mentioned initially.

4.2 Definition of body of knowledge

The body of knowledge presented in Deliverable D2.1 brings together evidence from peer-reviewed and grey literature on technologies used in CS. The work draws on established research efforts, including the 2023 literature review *Research Infrastructures in Citizen Science: State of Knowledge and Taxonomic Framework as a Pathway to Sustainability*[1], which examined 74 documents addressing technology in CS and provided a structured view of technological and methodological models. Further input comes from the ECS project's Deliverable D3.1 *Best Practices - Citizen Science Infrastructures*[2], which analysed 70 references and described existing technological platforms, services, and integration mechanisms, with emphasis on technical architectures and platform services for CS.

In addition, this deliverable integrates findings from 704 documents and project outputs associated with 24 EU-funded initiatives focused on CS technologies, complemented by snowball-referencing and expert recommendations. The resulting corpus forms a robust evidence base that gives us a structured understanding of current practices, challenges, solutions and gaps in the technological landscape of CS.

The mapping of CS platforms, tools, and services draws on six sources of data. These include eight existing databases linked to meta-reviews on CS technologies, results from CS applications available in the Android and iOS stores, the GitHub repository, the 24 EU-funded projects, and additions gathered through snowballing. Figure 6 brings together all sources and indicates the number of records (e.g., documents, URLs) associated with each.

[1] https://theoryandpractice.citizenscienceassociation.org/articles/10.5334/cstp.831
[2] https://doi.org/10.5281/zenodo.10635857

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

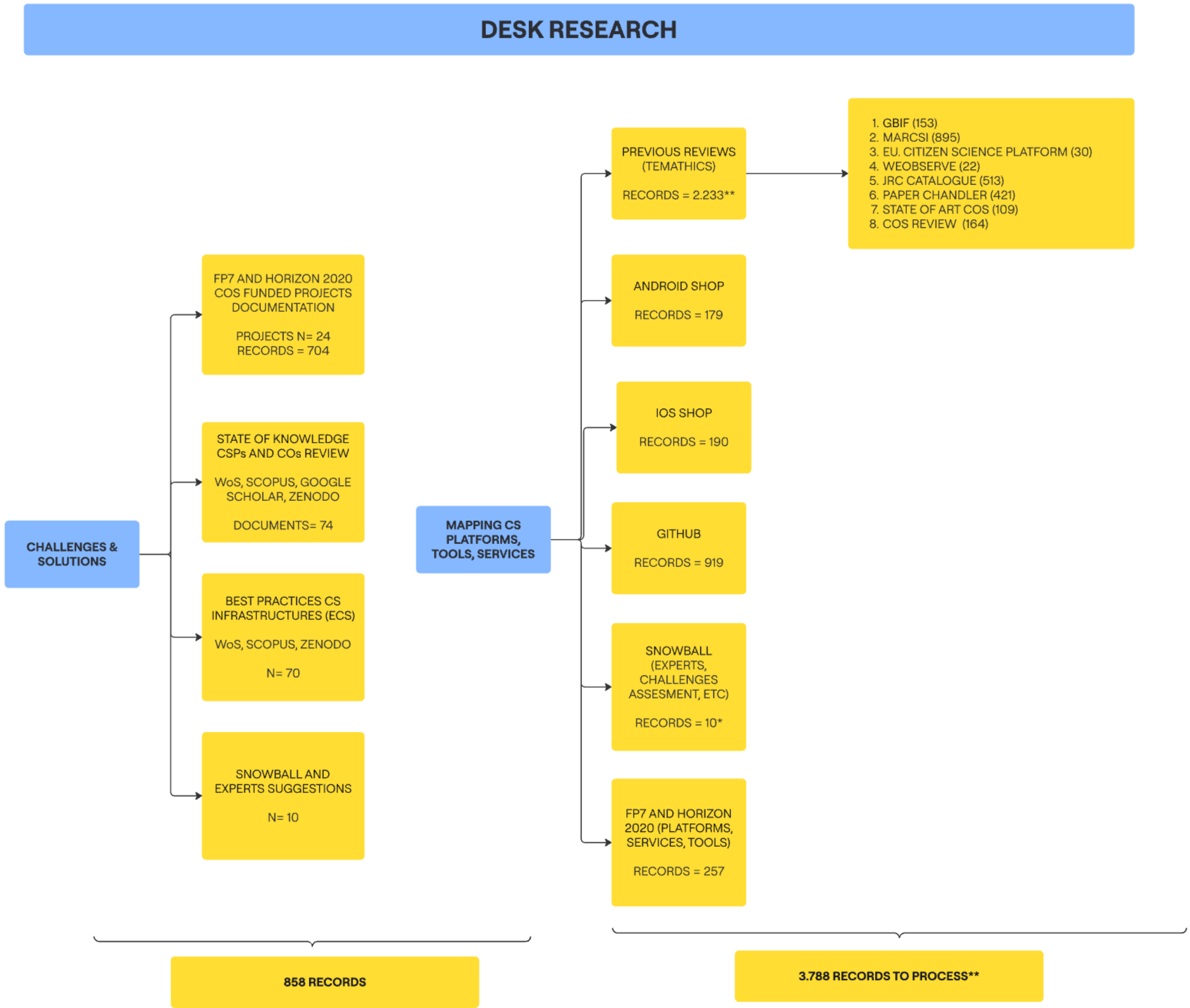

Figure 6 Sources of information for the D2,1 body of knowledge.

4.2.1 EU Funded CS infrastructures

We identified 24 EU-funded projects that developed citizen observatories, infrastructures, and platforms between 2012 and 2025. Selection was informed primarily by the *Roadmap for the Uptake of Citizen Observatories*, a document curated by a consortium of specialists in citizen science as part of the WeObserve project. The roadmap extends beyond citizen observatories and compiles the main EU initiatives that have advanced technological components for citizen science, which makes it a robust point of reference. The authors of this deliverable reviewed the list to validate its relevance and coherence; it aligned well with our own assessment. A few gaps emerged during the process—for instance, the initial absence of the REINFORCE project—which have now been addressed.

While the WeObserve study analysed a consolidated set of 23 Citizen Observatory related projects, we fully acknowledge that many additional EU-funded citizen science initiatives exist across other programmes (e.g. LIFE, Interreg, EMFF) and within Horizon Europe and Horizon 2020 under themes

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

not explicitly labelled as Citizen Observatories. The choice to rely on this specific set is intentional: these projects represent the most coherent, comparable and mature body of work focused on large-scale, technology-enabled participatory data generation. They also provide continuity in methods, infrastructures and lessons learned, making them particularly relevant for a comparative technical analysis. This subset is used as a representative sample to extract patterns, gaps and solutions that are broadly applicable across the wider European CS ecosystem.

The 24 projects have generated 704 research records (i.e., documents, reports, papers and other related documentation). The projects focus on environmental monitoring across different domains: air quality (CITI-SENSE, hackAIR, CitieS-Health), water monitoring (Citclops, WeSenseIt, Scent, Monocle, Ground Truth 2.0), odour detection (OMNISCIENTIS, DNoses), soil observation (GROW Observatory), biodiversity (COBWEB, FRAMEwork), land use (LandSense, DIONE), urban resilience (SMURBS, WeCount) and climate adaptation (CAPTOR, TeRRIFICA).

Five projects have addressed infrastructure and integration: WeObserve consolidated knowledge from citizen observatories, MICS measured citizen science impacts, Cos4Cloud integrated citizen observatory data with the European Open Science Cloud, Making Sense designing digital maker practices and REINFORCE developed research infrastructures for citizen engagement.

Table 1 presents these projects together with their focus areas, timelines, and number of associated records. The 704 records informed the identification of challenges, solutions, and gaps derived from the technical documentation of the CS technologies developed in these projects. These records also served as inputs for the *mapping of platforms, tools and services*, which compiles the developments generated across the projects.

Table 1 EU-funded projects related to the development of citizen science technologies

No.	Project Name	Focus	Project timeline	Records
1	OMNISCIENTIS	Odour monitoring	2012-2014	7
2	Citclops	Coastal and marine water quality monitoring	2012-2015	27
3	CITI-SENSE	Air pollution monitoring	2012-2016	62
4	COBWEB	Biosphere monitoring	2012-2016	30
5	WeSenseIt	Flood and drought monitoring	2012-2016	33
6	Making Sense	Open design and digital maker practices for monitoring environment	2015-2017	29

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

7	CAPTOR	Collective Awareness Platform for Tropospheric Ozone Pollution	2016-2018	19
8	hackAIR	Development of an open technology toolkit for citizens' observatories on air quality	2016-2018	43
9	Ground Truth 2.0	Flood risk management	2016-2019	26
10	GROW Observatory	Soil	2016-2019	41
11	Scent	Water supply and quality	2016-2019	15
12	LandSense	Land use and land cover monitoring	2016-2020	32
13	SMURBS	Integration of EO and citizen observations for a common approach to enhance urban environmental and societal resilience	2017-2021	29
14	WeObserve	Knowledge consolidation and mainstreaming of Citizen Observatories	2017-2021	41
15	DNoses	Odour monitoring	2018-2021	29
16	Monocle	Water quality monitoring	2018-2022	38
17	CitieS-Health	Assessing urban air and noise pollution and the link to health impacts	2019-2021	21
18	MICS	Measuring impacts of citizen science	2019-2021	30
19	WeCount	Urban road transport monitoring	2019-2021	26
20	TeRRIFICA	Adaptation processes to climate change through living labs	2019-2022	12
21	Cos4Cloud	Interoperability and integration of Citizen Observatory technology and data with European Open Science Cloud	2019-2023	64
22	DIONE	Complementing EO data with farmer-based monitoring to inform Common Agricultural Policy (CAP) and decision-making at farm level	2020-2022	24
23	FRAMEwork	Citizen Observatory for monitoring biodiversity in farmland landscapes	2020-2025	20

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

24	REINFORCE	REsearch INfrastructures FOR Citizens in Europe	2019-2022	6
				704

*Records: Number of documents, reports and papers available for each project.

*EO: environmental observations

4.2.2 Platforms, tools and services

The **mapping of CS platforms, tools, and services** was built through a structured, multi-source methodology integrating **13 complementary datasets** listed in Table 2, that together capture the diversity of digital infrastructures supporting CS in Europe. Each source contributed different types of evidence —ranging from institutional databases and peer-reviewed inventories to software repositories and expert inputs— allowing a robust and multidimensional representation of the current ecosystem.

Table 2 List of sources for the mapping of CS platforms, tools and services.

Source	Description	No. records
Citizen science data in GBIF (Global Biodiversity Information Facility)[3]	Provides a baseline of datasets explicitly tagged as CS, enriched using its API to retrieve metadata about institutions and data accessibility.	147
MARCSI (Inventory of Marine CS Initiatives)[4]	Contributes marine-focused projects with detailed documentation about FAIR data practices.	894
EU-Citizen.Science[5]	Contains official records of platforms operating in Europe, obtained by scraping individual platform pages.	28
WeObserve[6]	Provides a public registry of citizen observatories, showing a mapped overview of the COs established primarily across Europe.	21

[3] https://data-blog.gbif.org/post/gbif-citizen-science-data/
[4] https://zenodo.org/records/14260016
[5] https://eu-citizen.science/platforms/
[6] https://www.weobserve.eu/knowledge-base/

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

The Joint Research Centre (JRC) Catalogue[7]	Provides a repository of datasets from CS initiatives, categorized by themes and sources, with many datasets originating from structured platforms, offering insight into technological infrastructures.	512
Contribution of Citizen Science Towards International Biodiversity Monitoring, journal paper. [8]	Includes a database that lists CS programs including their geographic and thematic coverage, many of which operate via online platforms or structured systems.	420
State-of-the-Art Study in Citizen Observatories.[9]	Analyzes current technological trends, tools, and system architectures in citizen observatories, providing detailed mapping of platforms and challenges in their development.	108
Research infrastructures in citizen science, journal paper.[10]	Metareview of peer-reviewed research on citizen science technologies, accompanied by a dataset listing citizen science platforms.	110
Android app searches	Provided data on CS-related mobile applications, extracted through targeted keywords and categorized by domain (environment, biology, health, astronomy, etc.).	178
iOS app searches	Replicated the exploration done in the Android store within the Apple Store ecosystem, using the term CS.	48
GitHub	Collected popular open-source CS repositories through API queries ranked by stars.	918
EU-funded project documentation	Gathered information from reports, deliverables, and websites of 24 European projects developing CS tools or infrastructures.	256
Snowball expert inputs	Included manually reported platforms and tools emerging from interviews and consultations with domain experts between	26

[7] https://data.jrc.ec.europa.eu/collection/citsci
[8] https://www.sciencedirect.com/science/article/pii/S0006320716303639#ec0010
[9] https://helda.helsinki.fi/items/e17bc53c-8434-4573-bbb5-a1280c7696d2
[10] https://theoryandpractice.citizenscienceassociation.org/articles/10.5334/cstp.831

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

	May and November 2025, as well as internal benchmarking of citizen observatories by EMBIMOS research group.	

*No. records: Platform, tool or service (or potential one) identified in the database.

Together, these thirteen sources generated a **harmonized dataset of 3,666 entries**, systematically cleaned, normalized, and enriched to further develop outputs like the **catalogue services, tools and platforms** or deeper analysis about the state of CS technology.

The **characterisation of platforms** builds on the consolidated list of tools, services, and technologies extracted from the thirteen sources described before. From this larger landscape, a subset of 38 platforms widely used within the European citizen science community was selected for deeper analysis. Selection also depended on their current operational status, since only active platforms allowed systematic extraction of publicly available information through web-scraping procedures. This approach ensured that the platforms assessed represent both the most established solutions in practice and those with sufficient technical stability to support reliable data collection.

For each selected platform, a set of six analytical criteria guided the extraction of information as shown in Figure 7. **Development status** included identifiers such as platform name, URL, domain of application, descriptive information, operational status, and year of creation. **Licensing agreements** covered data licences, software licences, and references to code repositories. **Support structures** captured the presence of explicit governance arrangements. **Compliance with data protection and responsible data** use was examined through GDPR statements, terms of use, and privacy policies. **Adherence to FAIR principles** was assessed through indicators such as data download options, availability of APIs, and the use of standards. **Additional criteria** included interface languages, country and organisation of the managing entity, and contact information such as platform email.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

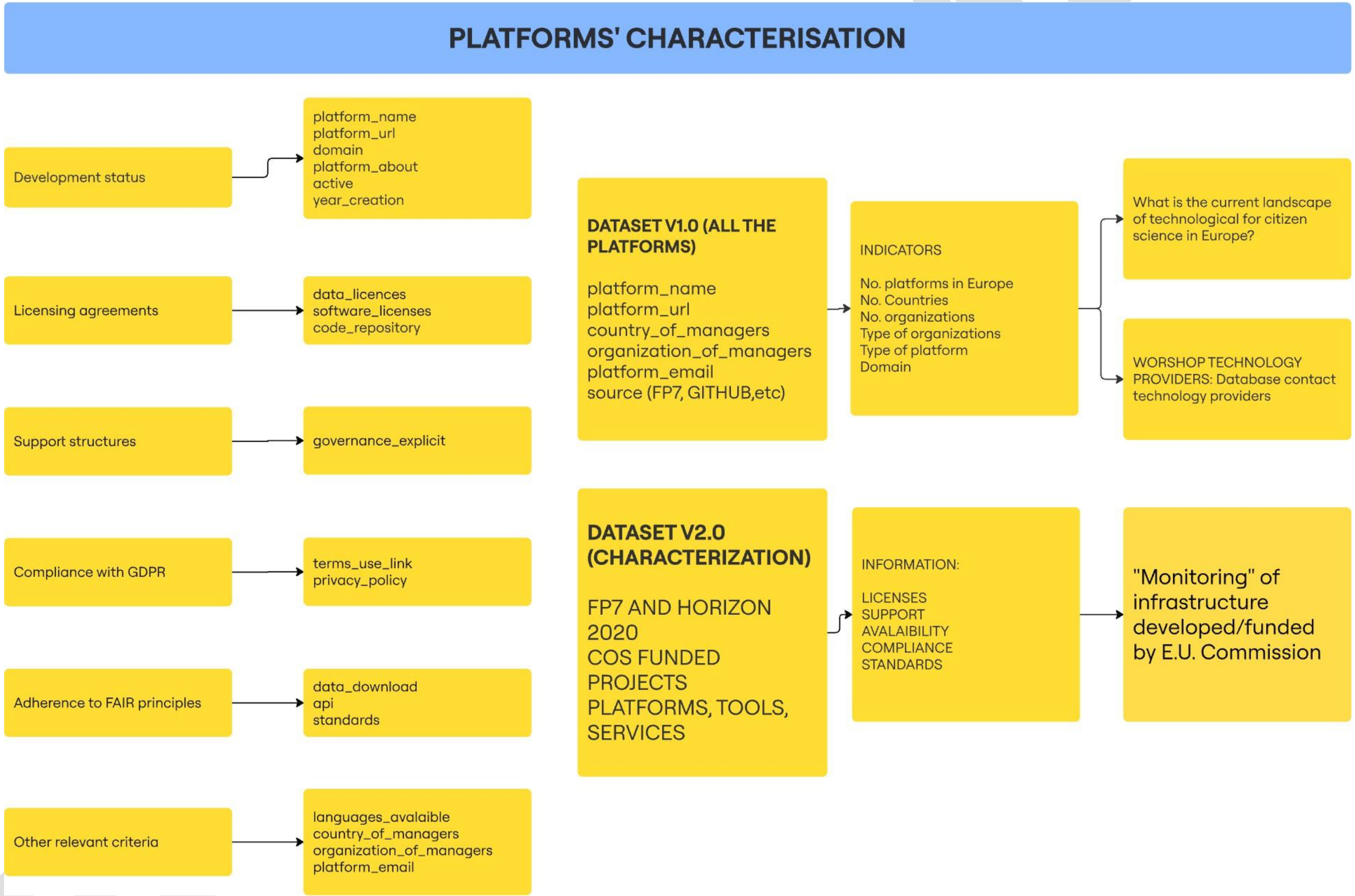

Figure 7 Characterisation of platforms. Criteria and data extracted to characterise citizen science platforms.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

5. Mapping of CS platforms, tools and services

The landscape analysis is based on the ~3,666 entries of CS technology related initiatives in Europe. These include a mix of digital platforms, tools, and services, as well as some project-specific entries and organizations from the sources listed in Table 2. In the context of CS technology, platforms are broadly defined as sociotechnical systems that bring together communities and functionality – "both means and outcomes…the 'things' or boundary objects in a design process – generating spaces where communities of practice can form" [15]. In other words, a **platform** is an environment (usually web or app-based) that enables volunteers and scientists to collaborate, contribute data, and share results. By contrast, **tools** tend to be standalone applications or instruments serving a specific function (e.g. a sensor kit or a single-purpose app), and **services** are supporting components such as data APIs or online services integrated into platforms.

Using these definitions (including Cuartielles' notion of platforms as community-centered digital constructs [15]), the majority of entries in the inventory can be categorized as *platforms*. In fact, among the entries that had an explicit type classification, about 95% were labeled as platforms, with only a small fraction identified as tools or services. This suggests that Europe's citizen science technology landscape is dominated by platform-type initiatives – comprising mobile apps and web portals where citizens submit observations, classify data, or otherwise participate in research. Examples range from biodiversity recording websites to environmental monitoring apps. The *tools* in the inventory are comparatively few (dozens out of thousands) and often correspond to hardware sensor kits or specialized software developed within projects (e.g. DIY air quality sensor nodes, AI-based species identification apps, or data analysis toolkits). Similarly, *services* entries (also only a handful) tend to be back-end services or infrastructure components – for instance, an API for plant identification (like Pl@ntNet-API) or data upload and visualization services used by multiple projects.

It's worth noting that some entries were also tagged as *projects* or *organizations*, reflecting the overlap between specific citizen science projects and the platforms/tools they produce. In many cases, a single citizen science initiative might consist of an organization running a platform with associated tools. For the purposes of this analysis, however, our focus is on the technological facets – hence we treat project-based platforms similarly to standalone platforms. Overall, the European CS tech ecosystem is rich but skewed heavily toward platforms (in the sense of full-featured applications or portals supporting communities), with relatively fewer isolated tools or services. This aligns with the idea that successful CS often requires an integrated platform providing multiple functions – data collection, management, community engagement, etc. [16]. Indeed, CSPs are seen as "web-based infrastructures that provide a single point of access to multiple functions…designed to enable and broaden citizen science practice."[16]

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

5.1 Current active platforms and long-term functioning

A critical aspect of this landscape is the sustainability and current status of these platforms and tools. CS technologies often emerge from time-limited projects (e.g. a research grant or a pilot funded for a few years). As a result, many platforms face sustainability challenges once initial funding ends. One proxy for sustainability is whether the platform's URL is still functional (i.e. the website or app remains accessible and maintained). We examined the inventory's entries to determine how many are currently active versus likely defunct.

Based on URL functionality, a substantial subset of the 3,666 entries appears to be inactive or no longer accessible. Many older project websites have gone offline or are no longer updated, reflecting the common pattern that short-term projects often do not transition into long-term infrastructures. For example, numerous citizen observatory initiatives from the early 2010s – often set up with EU funding – have since had their domains expire or content stagnate once the project concluded. Entries with dedicated .eu project websites or niche URLs are particularly prone to link rot. By contrast, platforms backed by enduring institutions or communities tend to remain online. For instance, national biodiversity data portals and well-supported community platforms still have functional websites years later, whereas one-off apps from small research projects can disappear if not sustained.

It's estimated that on the order of 20–30% (or more) of the listed initiatives may be defunct or dormant at present – especially among those launched in the 2000s and early 2010s. This underlines a key challenge in the European CS tech landscape: maintaining platforms beyond the initial enthusiasm. Recent analyses emphasize "the need for long-term institutional support, shared services, and coordinated policies to ensure [platform] sustainability" [9]. The inventory's existence is itself a testament to growing awareness of this issue; by cataloging projects, RIECS and related efforts can identify which platforms are thriving and which have faded away. Encouragingly, many newer platforms (and updates of older ones) are moving toward more sustainable models – either by integrating into larger infrastructures, open-sourcing their code, or securing ongoing funding via institutions or communities. Projects that have transitioned into long-term infrastructures (for example, a one-time project evolving into a permanent citizen observatory run by a museum or NGO) are far more likely to have active platforms today. Robust, established platforms (often community-driven or institution-backed) tend to persist, whereas many ad-hoc tools and apps see usage dwindle after their initial project phase.

5.2 Geographical distribution

The inventory also sheds light on the geographical distribution of CSPs in Europe. Entries are associated with initiatives from across Europe and beyond, reflecting a broad spread but with notable clusters. In terms of where these platforms/tools originate or are managed, a few countries stand out:

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

- **United Kingdom**: The UK accounts for one of the largest sets of entries in the inventory. This is due in part to its active citizen science community and early adoption of CSPs. The UK hosts national platforms like iRecord (for biodiversity recording) and has been involved in global projects (the Zooniverse platform has strong UK roots). Many UK-based organizations (museums, agencies, NGOs) have developed citizen science tools, contributing to the high count.

- **Western/Northern Europe:** Countries such as Spain, France, Germany, the Netherlands, Ireland, and Austria are well-represented. Spain, for example, has contributed platforms like MINKA and various apps from EU projects, while France is known for platforms like Pl@ntNet. The Netherlands is notable for Observation.org and the suite of national portals feeding into it. Ireland's presence is interestingly large in the inventory, likely bolstered by the inclusion of data from the Irish EPA and biodiversity projects (e.g., Ireland's National Biodiversity Data Centre platform). Austria hosts the Spotteron platform (which powers many apps under a commercial model) and other CS hubs.

- **Global and Non-European**: Importantly, the inventory isn't strictly limited to EU-origin projects; it includes global platforms used in Europe. For instance, US-based platforms like iNaturalist, eBird, and Zooniverse (as well as the CitSci.org) are included due to their European user base and project participation.

This geographic spread highlights that Europe's CS infrastructure is partly home-grown and partly woven into a global network. CS initiatives in Europe make a significant use of platforms developed outside the region. Conversely, some European platforms (like those in the biodiversity domain) have global reach or user communities beyond Europe. For example, the Dutch-led Observation.org platform now serves users worldwide and gathers observations globally, not just in Europe.

At a high level, Western Europe dominates the landscape in sheer numbers of platforms/tools – likely reflecting greater funding and activity in countries like the UK, France, Spain, Germany, and the Netherlands. Meanwhile, fewer entries originate from Eastern or Southern Europe except via international collaborations. This imbalance indicates that capacity and investment in CS infrastructure have been stronger in some countries, while in others CS is still emerging, which partly explains the limited number of platforms and related services. Initiatives such as the European Citizen Science Association (ECSA) and EU funding are helping to reduce these differences across the region. The inventory's breadth (including contributions from pan-European projects and networks) hints at growing cross-border collaboration. Still, national silos remain common – many countries developed their own platforms for similar purposes (especially in biodiversity monitoring), leading to parallel systems in different languages and contexts.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

5.3 Domain focus and fragmentation trends

The European CS tech landscape spans multiple domains. However, there are clear focal areas where CSPs are especially prevalent:

- **Biodiversity and Nature Observation:** This is arguably the largest domain represented. A very high number of platforms in the inventory deal with recording observations of species (birds, insects, plants, etc.) and biodiversity data collection. Examples include national biodiversity portals (like Sweden's Artportalen, Spain's Observado, or Belgium's NaturaList), global apps like iNaturalist and eBird, and thematic platforms such as Butterfly-monitoring schemes or bird atlas projects. The inventory data (and external studies) confirm that biodiversity platforms constitute a major portion of CS infrastructure. Chandler et al. (2016) [17] identified over 100 biodiversity data portals worldwide, and Europe alone contributes dozens of those. Fragmentation in this domain is high – many countries or even regions have their own systems for similar tasks (e.g., separate bird monitoring apps), and multiple platforms often coexist (for example, a birdwatcher in Europe might choose between entering sightings in eBird, a local ornithological society database, or Observation.org). While this diversity allows tailoring to local needs, it also leads to siloed data. Efforts like GBIF (Global Biodiversity Information Facility) act as integrators by aggregating data from many of these platforms. Indeed, major platforms such as eBird, iNaturalist and Artportalen are **major contributors of CS data to GBIF** [18], illustrating how data from fragmented sources can be pooled at a higher level.

- **Environmental Monitoring (Air/Water/Climate):** The next prominent domain is environmental observation beyond biodiversity. This includes citizen observatories for air quality, water quality, weather and climate, noise, and other environmental parameters. Europe has seen a proliferation of environmental sensor platforms and crowdsourced monitoring projects – for instance, the Luftdaten project in Germany expanded into Sensor.Community, a platform hosting thousands of low-cost sensors measuring air pollution across Europe and globally. Similarly, projects like FreshWater Watch, Water Frame, or CrowdWater focus on water monitoring, while others track phenology or climate indicators. Many of these started as independent projects (often EU-funded under programs like Horizon 2020 or national grants). The result is a patchwork of environmental apps and sensor networks. Some are *tailored* citizen observatories for specific communities or topics, while others are more *open* multi-purpose platforms that host multiple projects [9]. The fragmentation trend here is characterized by duplication of effort – e.g. several groups building similar DIY sensor kits or apps – though there is movement toward consolidation. Projects like the European WeObserve initiative attempted to network these observatories, and platforms such as OpenSenseMap or Smart Citizen allow different sensor projects to upload data in a common space. Still, compared to biodiversity, the environmental monitoring domain has fewer truly massive platforms and more scattered, smaller-scale ones. Many remain active only at local scales. Sustainability is a concern: some sensor networks thrive on community uptake (as in sensor.community, which persists as a grassroots effort), while others faded once the pilot ended.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

- **Health and Biomedical:** This is a smaller but growing domain. The inventory contains a few entries related to health, such as Influenzanet (a Europe-wide platform where citizens self-report influenza symptoms for epidemiological tracking) and various crowd health or biohacking initiatives. In general, health-related citizen science in Europe tends to be less platform-centric (often organized as research studies collecting data via surveys or wearables, rather than public-facing apps). Nonetheless, some platforms do exist (for example, portals for patient-led research or apps for tracking disease outbreaks). One notable trend in recent years was the use of citizen science tools during the COVID-19 pandemic (e.g. apps to report symptoms or take part in COVID data collection), though those were often built on existing infrastructures or as ad hoc tools. Compared to biodiversity and environment, health citizen science platforms are few, and many are linked to specific research programs.

- **Other Domains:** The inventory also includes platforms in areas like astronomy, cultural heritage, social science, and education. For instance, Zooniverse hosts astronomy projects (e.g. Galaxy Zoo) and humanities projects; some European platforms focus on history or linguistics (transcription projects, etc.), and a variety of educational CS toolkits appear as well (such as Arduino-based science kits or apps for school projects). These tend to be scattered and often leverage either global platforms (like Zooniverse for crowd classification tasks) or smaller bespoke apps. The fragmentation here is mostly by discipline – each niche tends to have its own platform or uses a general one like Zooniverse.

The European landscape is highly fragmented by both geography and domain. There are many overlapping platforms serving similar purposes, often developed in parallel. For example, in the biodiversity realm a volunteer might have to navigate multiple apps depending on the taxon or country – one app for birds (perhaps a national ornithology app or eBird), another for insects, a separate platform for plant observations, etc., despite the possibility of unified approaches. This fragmentation has downsides (duplication of effort, small user bases per platform, data silos), but it also reflects healthy innovation and community-specific customization. There are signs of consolidation in some areas: aggregator platforms and data standards are knitting some of these pieces together. The fact that platforms like iNaturalist or Observation.org can host "hundreds of biodiversity projects and users" under one roof [9] is an example of countering fragmentation by aggregation. Likewise, multi-project infrastructures like CitSci.org or Spotteron allow many projects to live on a single platform, sharing technology. **The emerging concept of Citizen Science as a Service (offered by platforms like Spotteron or Anecdata) is essentially an attempt to reduce fragmentation – projects don't each build from scratch, they use a common service. Nonetheless, Europe's citizen science tech will likely continue to feature a rich mosaic of platforms, given the diverse languages, cultures, and scientific communities involved.**

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

5.4 Reliance on US-Based Platforms vs. European Infrastructures

A striking aspect of this landscape is the interplay between European initiatives and major US-based CSPs. Many of the world's largest CSPs were developed in the US (often with global scope), and European projects frequently utilize these instead of (or in addition to) Europe-originated tools. The inventory data and external analyses point to three especially prominent US-born platforms in European citizen science:

- **iNaturalist (USA)**: A global biodiversity observation platform (jointly run by California Academy of Sciences and National Geographic) which has become hugely popular worldwide, including across Europe. Many European naturalists and projects use iNaturalist for recording species observations and crowdsourced identification. Several EU countries have iNaturalist *nodes* or communities (e.g., iNaturalistUK, iNaturalistEU) leveraging the central infrastructure. Its impact in Europe is significant – millions of observations from European users are on iNaturalist. Notably, however, Europe also has parallel platforms (like Observation.org and national systems) as alternatives. A recent study highlighted that platforms like iNaturalist and eBird have "generated millions of biodiversity observations, transcending geographical boundaries", underpinning participatory science on a large scale [9]. European reliance on iNaturalist is strong in the sense that many projects find it easier to plug into this existing network than to build something new. That said, iNaturalist's data is openly shared (e.g. to GBIF), meaning European science benefits from it even if infrastructure is US-hosted.

- **eBird (USA):** Run by the Cornell Lab of Ornithology, eBird is the world's largest CSP for bird observations. It has a devoted user base in Europe (especially Spain, the UK, and increasingly other countries where birdwatchers log their sightings via the eBird app). European ornithological societies historically had their own databases, but eBird's powerful tools and global reach have attracted many European users. Several European country bird atlases now incorporate or draw from eBird data. The downside is a dependency on a US institution's platform, but so far eBird has proven sustainable and cutting-edge. European projects (and EU policy) acknowledge eBird as a key resource for biodiversity monitoring.

- **Zooniverse (USA/UK)**: Zooniverse is somewhat unique: it was co-founded by teams in the US and the UK, and is the world's largest platform for online CS *crowdsourcing* projects. It supports projects in astronomy, ecology, medicine, history, etc., by providing a portal where volunteers classify images or data. Europe relies heavily on Zooniverse for any project that requires volunteer classification at scale. Many European research projects (from searching for exoplanets to transcribing archives or identifying wildlife in camera trap photos) have chosen to create a Zooniverse project rather than develop a custom platform. Zooniverse's scale is immense – it boasts over 120 projects and more than 2 million registered volunteers worldwide [19]. European contributors and projects form a significant part of that community. The platform's infrastructure is largely hosted and managed by a team that includes US institutions (like Adler Planetarium) and UK's Oxford University,

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

so it's a transatlantic effort. European initiatives benefit from Zooniverse by gaining instant access to a large volunteer pool and a proven platform, but at the same time this illustrates Europe's partial outsourcing of infrastructure: instead of dozens of separate EU-built crowd-classification platforms, Zooniverse became the go-to solution.

Beyond these three, other US-based infrastructure also touches Europe: for example, the CitSci.org platform (a general project management tool for citizen science) is US-based but hosts some European projects; Anecdata (from a US lab) similarly hosts a number of European environmental projects on its platform [9].

To what extent do European initiatives rely on US-hosted infrastructures? In practice, quite extensively in certain domains. For biodiversity data collection, Europe's community scientists often use iNaturalist and eBird by the tens of thousands – these platforms have become part of the global CS commons, and Europeans are among the top contributors. For online crowdsourcing tasks, Zooniverse is virtually a default option used in Europe. This reliance means that a share of Europe's citizen science data and participation flows through servers and organizations outside Europe. While this is not inherently negative (science participation is global), it does raise considerations about data sovereignty, long-term availability, and alignment with European values (e.g., GDPR compliance, localization, etc.). So far, the collaboration has been positive – eBird and iNaturalist have European chapters/partners, and Zooniverse was co-developed with European input. However, European stakeholders are indeed interested in building home-grown capacity and alternatives to ensure not all roads lead across the Atlantic.

5.5 European alternatives and emerging infrastructures

In recent years, Europe has been actively developing robust CSPs of its own to complement or offer alternatives to the major US-based ones. A number of these European-led platforms are showing strong growth and could be considered the backbone of Europe's CS infrastructure moving forward. Here we highlight a few notable examples and their domains:

- **Observation.org (Observation International):** Based in the Netherlands, Observation.org is a powerhouse platform for recording biodiversity observations (similar in purpose to iNaturalist). It originated from the Dutch community site Waarneming.nl and now serves a global audience with multi-language support. Observation.org allows enthusiasts to log sightings of birds, mammals, plants, insects – essentially any wildlife – and has an open-data ethos. It has become a European alternative to iNaturalist. In fact, by some metrics Observation.org holds its own against iNaturalist: it has amassed on the order of hundreds of millions of observations. (One report noted over 269 million records on Observation.org, comparable to the ~226 million on iNaturalis t[9].) The platform's strength lies in its community roots and integration with European monitoring schemes. Many national portals (in Belgium, the Netherlands, Spain, etc.) feed into the Observation.org database,

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

reducing fragmentation among those participants. It exemplifies a sustainable infrastructure, maintained by a nonprofit and used by thousands of Europeans daily.

- **Spotteron:** Hailing from Austria, Spotteron is a different kind of platform: it's a professional platform-as-a-service for citizen science projects. Rather than one site for all data, Spotteron offers customizable apps and web portals for individual projects, all built on a common technology stack. Dozens of European citizen science projects (many related to environment and sustainability) have chosen Spotteron to host their apps – for example, projects like ClimateWatch, CrowdWater, Nature Calendar, and many others carry the SPOTTERON tag in their names. Each project gets its own branded app, but users can often use a single account across Spotteron projects. By aggregating multiple projects on one platform, Spotteron helps projects avoid reinventing the wheel. It's a sustainable model (projects pay for service packages, ensuring maintenance) and has grown particularly in central Europe. Spotteron's focus is often environmental monitoring and community mapping, but it spans various domains (even archaeology or astronomy projects could, in theory, use it). This platform illustrates a path to sustainability through a service model, and it reduces fragmentation by providing a common framework for many small initiatives. While Spotteron is a private company effort (not a government infrastructure), it fills a niche for those who want a stable, long-term home for their citizen science apps without building one from scratch. It stands as a strong European-grown alternative to ad-hoc development, even if it doesn't directly compete with the huge global community sizes of iNaturalist or Zooniverse (Spotteron's strength is depth of customization and steady support rather than sheer scale of a single database).

- **sensor.community (Luftdaten):** This is a grassroots European platform that has become a globally adopted infrastructure for environmental sensing, especially air quality. Started in Stuttgart, Germany, as *Luftdaten*, it enabled citizens to build low-cost particulate matter sensors and share air quality data to an open map. The initiative expanded and rebranded as sensor.community, now encompassing thousands of sensors in over 70 countries. In Europe, sensor.community nodes are widespread, forming one of the largest open-air quality monitoring networks in the world. The platform is open-source and community-driven, emphasizing local empowerment to gather data on air pollution. Its success shows that European initiatives can lead in the DIY citizen science hardware+platform space. Notably, sensor.community data has been used by scientists and even policymakers to complement official monitoring. It's a robust platform in the environmental domain and continues to grow (with expansions into noise sensors, weather sensors, etc.). In terms of fragmentation, sensor.community has actually reduced fragmentation in its domain: before, various small groups had their own sensor projects; now many have coalesced around the sensor.community standard and infrastructure.

- **MINKA:** MINKA is an emerging European platform worth highlighting for its innovative approach. Developed EMBIMOS research group in ICM-CSIC, MINKA is a citizen science observatory platform geared towards environmental and biodiversity data for the UN Sustainable Development Goals (SDGs) [20]. In essence, it's designed to help trained volunteers collect and share observations

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

that are aligned with specific indicators (e.g., marine biodiversity for SDG14 Life Below Water, etc.). It can be seen as a European attempt to provide a unifying platform that not only gathers observations but ties them into a larger framework of sustainability tracking.

- **Others and Domain-Specific Platforms:** Europe has numerous other notable platforms – e.g., Pl@ntNet (a French-led plant identification app and repository, extremely popular globally for flora observation – effectively an alternative to iNaturalist for plants, using AI to identify species from photos), Artportalen (the Swedish biodiversity database which is a long-standing national infrastructure feeding into international networks), CrowdHealth platforms (like InfluenzaNet in health), and many science-specific portals (for instance, Einstein@Home and other BOINC-based citizen science computing projects have European contributors, though BOINC itself is US-developed). In the humanities, European archives and libraries have set up crowdsourcing platforms (e.g., Transcribathon for WWI letters, or Europeana's citizen engagement portals). These might be smaller in scale, but collectively they indicate a broadening of citizen science infrastructure into diverse domains with European leadership. Also relevant to point out is the EU-Citizen.Science platform that functions as a knowledge hub led by the European Citizen Science Association (ECSA), offering training, project resources, and coordinated access to citizen-science initiatives across Europe.

The emergence of robust European platforms is gradually addressing some fragmentation issues. For example, where once dozens of apps for plant identification existed, Pl@ntNet now aggregates a huge user base and dataset (with over 20,000 plant species recognized and millions of users). Observation.org unifies wildlife recording communities under one umbrella in a way that previously was split by country. The trend seems to be that successful platforms become aggregation points, either by being open and extensible or by offering superior functionality that attracts users away from smaller tools. **However, Europe's landscape still has a long tail of very niche or localized platforms that likely will remain in use for specific communities.** The concept of **CS research infrastructures** is essentially to provide stable, scalable European platforms that outlast individual projects. **In the long run, one might envision a more interconnected European ecosystem where data flows between platforms seamlessly and a few key platforms serve most needs, complemented by specialty tools.**

5.6 Macro-level trends

Taking a step back, this analysis reveals several macro-level dynamics in the European CS technology landscape:

- **Explosion of platforms (Quantity vs Quality):** Europe has a vast number of CSPs indicating vibrant activity and innovation. However, this has often meant many small platforms rather than a few large ones. The sheer quantity has led to overlap and redundancy. A positive trend is the gradual shifting from quantity toward quality and interoperability – recent efforts encourage projects to reuse existing platforms or at least make their outputs interoperable (via standards or data portals). The

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

RIECS inventory itself, alongside platforms like EU-Citizen.Science (which catalogues resources), shows an increasing desire to map and coordinate these efforts rather than let them exist in isolation.

- **Sustainability and lifespan:** The historical pattern was that many CS tech initiatives were ephemeral – created for a one-off purpose and not maintained. This is slowly changing. More platforms are being designed as long-term infrastructures from the start (often with open source code, community governance, or institutional backing). The awareness of high mortality among past projects has led to calls for better sustainability planning [21] [22]. We now see examples of projects transitioning to permanence (e.g., an EU pilot that is handed over to a university to run indefinitely). Still, funding mechanisms need to adapt to support maintenance, not just innovation. The presence of inactive entries in any inventory is a cautionary tale – one that European research infrastructure initiatives are trying to address head-on [21], [23], [24].

- **Global integration vs European autonomy:** Europe is both a contributor to and beneficiary of global CSPs. The integration is evident – European observations feed into global databases, and Europeans are key users of global apps. Yet, there's a parallel push for European autonomy in infrastructure – partly for strategic and political reasons (data governance, digital sovereignty), and partly to tailor tools to European languages/cultures. The result is a dual landscape: European platforms thriving in some niches (e.g., biodiversity, where Observation.org or Pl@ntNet are as integral as iNaturalist), while in other areas Europeans rely almost entirely on global platforms (e.g., Zooniverse for crowdsourcing, or eBird for bird data). Future developments like RIECS may further strengthen Europe's own platforms, but it's likely that a healthy collaboration with global platforms will continue. Interoperability will be key – ensuring data can move between US and EU platforms smoothly (for example, observations exchanged via GBIF or projects listed across multiple portals).

- **Community and infrastructure fusion:** A notable theme is that technology and the community are intertwined. Platforms that have succeeded in Europe usually pair technology with a strong community or institution. For example, Observation.org grew from a passionate community of naturalists; Spotteron sustains itself by building a client community among project owners; sensor.community thrives on volunteer builders. This echoes Cuartielles' point that platforms are sociotechnical constructs *managed by communities [15]* – the human element is crucial. Therefore, one sees that purely technical solutions without community uptake fade away, whereas those embedded in networks of people persist. Europe's citizen science movement, through organizations like ECSA and national citizen science centers, is helping nurture these communities, which in turn bolster the platforms.

- **Selected case illustrations:** To illustrate these broader dynamics, consider a specific case: the air quality citizen science domain in Europe. Several years ago, various projects (e.g., Citi-Sense, AirSensEUR, Smart Citizen Kit in Barcelona, and Luftdaten in Stuttgart) each developed their own sensors and platforms. Initially, this was highly fragmented – each had its own data portal and community. Over time, sensor.community (Luftdaten) emerged as a de-facto platform because it was

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

open, simple, and community-driven; many others started contributing to it. Now, a large portion of European DIY air sensors report through sensor.community, while some other platforms like Smart Citizen (by FabLab Barcelona) also continue but with a smaller user base. The trend in this case was consolidation around a robust European platform, improving sustainability and data consistency. On the other hand, consider a case from biodiversity: bird monitoring in Europe. Historically, each country had a scheme; some still do (e.g., the UK has BirdTrack/iRecord, Sweden has Artportalen, etc.), but many individual birders now use eBird because of its convenience and global community. Here we see a tilt towards a US platform despite local options, driven by network effects (seeing one's data alongside global data, the superior tools for eBird). The outcome is that European bird data largely flows into an American system (though often mirrored back via GBIF). This case underscores reliance on external infrastructure when it offers clear benefits.

Moving forward, we can expect the lines between standalone project platforms and long-term infrastructures to continue blurring: more project outputs will feed into enduring platforms, and more platforms will position themselves as flexible infrastructures that can host many projects (reducing the need for each project to start from scratch). The technological landscape for CS in Europe is thus trending toward a more sustainable, interconnected ecosystem, albeit one that must balance local diversity with global integration. As this happens, both macro-level trends (like consolidation and networking of platforms) and the on-the-ground examples (like a small citizen science app finding new life by plugging into a larger platform) will likely shape the next generation of CS practice in Europe.

6. Challenges, solutions and lessons learned

Section 6 frames Challenges and solutions through two lenses: cross-disciplinary challenges (6.1), and domain-specific cases (6.2). The cross-disciplinary part details recurring constraints in hardware and sensing, software and apps, data quality and governance, interoperability and standards, connectivity and scaling, system architecture, AI and analytics, device compatibility, user capacity and training, and long-term sustainability. Each theme is grounded in EU-funded practice, showing how heterogeneous data models, calibration drift, API fragmentation, and hosting choices shape project outcomes.

The domain section then moves from general patterns to concrete cases: Cos4Bio and Cos4Env as paired biodiversity–environment portals that operationalise the integration of observations and measurements across multiple citizen observatories; DHIS2 as a mature digital public good focused on health that demonstrates offline-first operation, open APIs, and country ownership at national scale; and climate-focused threads to connect observation continuity with European services and ICOS, an ERIC infrastructure, which offers a reference model for long-term, standardised, high-precision environmental monitoring through its distributed stations, thematic centres, and central carbon portal, showing how rigorous calibration chains and common data workflows support trust and reuse.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

6.1 Crossdisciplinary challenges

6.1.1 Technical challenges: EU-funded citizen science projects

The analysis of cross-disciplinary technological challenges draws on evidence from 24 EU-funded projects that developed citizen observatories, infrastructures, and data-collection systems between 2012 and 2025. Cross-disciplinary CS infrastructures encounter system-level interoperability challenges that become visible across the categories outlined in Sections 6.1.1–6.1.10. These difficulties emerge from heterogeneous data architectures, incompatible calibration protocols, and divergent quality-assurance approaches. For instance, Section 6.1.1 shows how environmental observatories such as CITI-SENSE, CAPTOR, and Making Sense struggled with sensor precision, drift, and environmental durability, revealing how hardware limitations cascade into later stages of data handling. In Section 6.1.2, usability and cross-platform development issues documented in hackAIR, D-NOSES, and WeCount illustrate how software decisions influence data completeness and user engagement. Likewise, Section 6.1.3 highlights the mixed quality and fragmented metadata produced in projects like Citclops, COBWEB, and Ground Truth 2.0, demonstrating the strain placed on validation and processing pipelines.

Across sections 6.1.4–6.1.6, the absence of common data models, uneven API integration, and complex system architectures—encountered in LandSense, SMURBS, Cos4CLOUD, and MONOCLE—show how attempts to merge distinct technological stacks often introduce new tensions, particularly when aligning with national standards or legacy tools. Section 6.1.7 adds a further layer, as AI-supported analytics in hackAIR, LandSense, and Cos4CLOUD require training datasets and computational capacity that many platforms cannot sustain. These constraints interact with device compatibility barriers (Section 6.1.8) documented in SCENT, WeCount, and MONOCLE, and with the substantial training and support needs (Section 6.1.9) identified in Making Sense, CitieS-Health, and TeRRIFICA.

Rather than isolated domain-specific issues, these examples reveal emergent system complexities. Solutions devised for one component—such as calibration routines, workflow orchestration, or privacy safeguards—often create dependencies or incompatibilities elsewhere. Projects must therefore handle measurement instrumentation, data governance, GDPR compliance, and infrastructure scalability in parallel, while responding to a broad spectrum of methodological expectations. Section 6.1.10 further shows that long-term sustainability concerns in GROW Observatory, MICS, and FRAMEwork exacerbate these tensions, as evolving standards and technology lifecycles require continuous adaptation.

Taken together, these insights point to a structural gap: the absence of shared APIs, harmonised data models, and flexible architectures capable of supporting interdisciplinary work across multiple areas, environment, health, social aspects, among others.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Hardware and sensor technology

Challenge category	Specific challenges	Projects related
Sensor accuracy & reliability	- Low-cost sensors produce questionable data quality - High variability between sensors from same manufacturer - Metal-oxide sensors have upper detection limits - Sensors struggle with environmental interference	CITI-SENSE, CAPTOR, Making Sense, SMURBS, CitieS-Health
Sensor calibration	- Large unresolved calibration issues - Factory vs. in-field condition mismatches - Need for regular recalibration due to drift - Lack of suitable general calibration models	OMNISCIENTIS, CITI-SENSE, CAPTOR, Making Sense
Environmental durability	- Severe environmental conditions prevent installation - Water damage and theft issues - Bio-fouling corruption in marine environments - Temperature effects on battery life and performance	OMNISCIENTIS, Citclops, GROW Observatory, SCENT
Hardware Failures	- Shortened sensor lifespans - Hardware malfunctions (camera modules, SD cards) - Connection issues (breadboard disconnections, Wi-Fi)	CAPTOR, GROW Observatory, WeCount, MONOCLE

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

	- Component unavailability and discontinuation	
Device compatibility	- Issues with older smartphones - Different camera and sensor capabilities - White balance procedure variations - Hardware specification limitations	Citclops, SCENT, WeCount, MONOCLE
Installation difficulties	- Complex sensor installation procedures - Requirement for technical supervision - Physical access challenges - Equipment security concerns	CAPTOR, WeCount, SCENT, WeSenseIt

Software and application development

Challenge category	Specific challenges	Projects related
Cross-platform development	- Android vs iOS compatibility issues - Bluetooth connectivity problems - Performance penalties with cross-platform tools - App crashes on older devices	CITI-SENSE, COBWEB, hackAIR, D-NOSES
User interface & usability	- Complex onboarding processes - Difficult taxonomy selection for images - Poor smartphone functionality	CITI-SENSE, LandSense, MICS, TeRRIFICA, WeCount

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

	- Need for extensive training and documentation	
Application stability	- App crashes and freezing - Data loss and transmission failures - Version compatibility issues - Battery drain problems	OMNISCIENTIS, SCENT, hackAIR, WeCount
Feature limitations	- Inability to work with encrypted Wi-Fi - No offline functionality - Limited data input capabilities - Missing user feedback mechanisms	WeCount, COBWEB, GROW Observatory

Data quality and management

Challenge category	Specific challenges	Projects related
Data quality control	- Mixed quality from crowdsourcing - Difficulty validating citizen-collected data - Need for continuous quality control procedures - Manual verification inefficiency	Citclops, COBWEB, WeSenseIt, Ground Truth 2.0, CitieS-Health
Data completeness	- Missing data in pilots - Insufficient data for reliable analysis - Spatial and temporal gaps - Low volume and limited coverage	CITI-SENSE, Citclops, SCENT, SMURBS

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Data processing	- Converting crowdsourced observations to scientific units - Handling irregular data availability - Complex data fusion from multiple sources - Computational intensity of processing	WeSenseIt, Ground Truth 2.0, LandSense, SMURBS
Metadata management	- Inconsistent metadata availability - Need for standardized formats - Missing coordinate and contextual information - GDPR compliance challenges	Citclops, GROW Observatory, LandSense, D-NOSES

Interoperability and standards

Challenge category	Specific challenges	Projects related
Data standardization	- Lack of common data models - Multiple data formats and structures - Difficulty combining outputs into single platforms - Harmonization with national standards	CITI-SENSE, COBWEB, Ground Truth 2.0, LandSense, SMURBS
API and service integration	- Complex data flow chains - Interoperability with existing systems - Integration with OGC standards - Semantic interoperability challenges	hackAIR, MONOCLE, Cos4CLOUD, DIONE, SMURBS
Legacy system integration	- Adapting existing open-source tools - Fragmented code ownership	COBWEB, Citclops, LandSense, SMURBS

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

- Workflow orchestration complexity

- Connecting disparate sensor technologies

Infrastructure and connectivity

Challenge category	Specific challenges	Projects related
Network connectivity	- Poor internet connectivity in rural areas - Unstable Wi-Fi and mobile signals - Unreliable GPS signals - Underground connectivity issues	Ground Truth 2.0, SCENT, WeCount, SMURBS, FRAMEwork
Real-time data transmission	- Need for robust systems handling intermittent connections - Direct data transmission challenges - Real-time quality control requirements - Scalability of data processing	WeSenseIt, MONOCLE, SMURBS
Cloud and server infrastructure	- Long-term hosting sustainability - Server maintenance and backup costs - Migration between platforms - Green hosting requirements	hackAIR, Cos4CLOUD, MICS, FRAMEwork

System integration and architecture

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Challenge category	Specific challenges	Projects related
Platform integration	- Integrating diverse software components - Complex system architectures - Multiple technology stack coordination - Custom solution development needs	Ground Truth 2.0, GROW Observatory, Cos4CLOUD, DIONE
Scalability issues	- Handling large volumes of IoT data - Scaling human management systems - Database optimization challenges - Cost implications of scaling	SMURBS, COBWEB, Cos4CLOUD
Workflow management	- Complex workflow orchestration - Process chaining difficulties - Automation framework requirements - Resource allocation management	COBWEB, LandSense

AI and advanced analytics

Challenge category	Specific challenges	Projects related
Machine learning implementation	- AI algorithms in learning phase - Need for large training datasets - Computational complexity requirements - Model accuracy improvements needed	LandSense, hackAIR, Cos4Cloud

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Automated analysis	- Algorithm optimization ongoing - Difficulty with image processing - Automated calibration challenges - Pattern recognition limitations	Citclops, SCENT, MONOCLE, WeCount

Human capacity, user experience and training

Challenge category	Specific challenges	Projects related
Technical skills requirements	- Users lack technical setup skills - Need for extensive training - Complex calibration procedures - Digital literacy barriers	Making Sense, CitieS-Health, TeRRIFICA, FRAMEwork
User support needs	- High technical support requirements - Need for troubleshooting assistance - Documentation and resource gaps - Continuous user guidance needs	Making Sense, hackAIR, WeCount, CitieS-Health

Long-term sustainability

Challenge category	Specific challenges	Projects related
Financial sustainability	- Funding for ongoing maintenance - Cost of continuous sensor replacement - Long-term infrastructure costs - Recurring technology investments	Making Sense, hackAIR, GROW Observatory, MICS, FRAMEwork

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Technology evolution	- Rapidly changing technology landscape - Need for continuous updates - Obsolescence of components - Keeping pace with standards evolution	COBWEB, Cos4CLOUD, MICS

6.1.2 Solutions and gaps

The Matrix B compiles the analysis of the 23 funded projects reviewed for their technological elements, operational workflows, and supporting infrastructures. The matrix offers an integrated overview of the requirements expressed by users, the characteristics of the devices and platforms developed across the projects, the service components they rely on, the main implementation challenges, and the corresponding solutions proposed to address them. It also documents the gaps that remain unresolved.

Drawing from this synthesis, several cross-cutting solutions emerge:

- **Modular architectures that support heterogeneous devices and data pipelines**, allowing projects to plug into shared components without rigid dependencies.
- **Dedicated validation and calibration procedures** that reduce uncertainty in data collection and processing, especially in complex environmental or health-related contexts.
- **APIs and interoperability layers** that facilitate data exchange across platforms and reduce duplication of technical effort.
- **Clear onboarding and licensing workflows**, which improve clarity for contributors and allow projects to manage data access conditions more effectively.
- **Community-facing tools for transparency and recognition**, which strengthen participation and accountability.

Despite these advances, the analysis surfaces a set of persistent gaps:

- **Insufficient harmonisation of metadata and data models**, which limits the ability to integrate outputs from different initiatives.
- **Limited reusability of device-specific solutions**, as many prototypes remain tied to narrow use cases or lack documentation for broader deployment.
- **Fragmented approaches to quality assurance**, with no shared protocol that spans multiple domains or device categories.
- **Weak linkages between citizen-generated data and institutional data pipelines**, which restricts the potential use of these datasets in formal evidence processes.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

- **Unstable maintenance pathways**, since several technical components depend on project-based funding without long-term hosting or governance arrangements.

6.1.3 Lessons learned for RIECS

The synthesis of requirements, challenges, solutions, and gaps captured in the Matrix B provides several lessons that should inform the design of a future RIECS research infrastructure. Although the projects reviewed differ in scope, discipline, and maturity, their shared difficulties and partial solutions reveal structural needs that RIECS must anticipate from the outset.

Modular and flexible technical design is essential: Projects repeatedly rely on bespoke device configurations and data pipelines. Modular architectures reduce this dependency and allow projects to align with shared components without losing their specific methodological approaches. RIECS will need to offer adaptable building blocks rather than fixed workflows.

Interoperability cannot be achieved through software alone: Many projects introduced APIs or conversion layers, but incompatibilities persisted due to divergent metadata models and documentation practices. RIECS must address interoperability at the level of data models, metadata completeness, calibration protocols, and documentation standards if it aims to enable cross-project integration.

Validation and calibration require domain-aware workflows: Environmental, health, biodiversity, and pollution monitoring projects each confront distinct quality assurance issues. The lessons show no generic protocol works across domains. RIECS should support domain-specific pipelines while ensuring that outputs can still connect to joint repositories and discovery services.

Sustainable governance is as critical as technical robustness: Several components created during the projects remain difficult to maintain once funding ends. Without a shared governance mechanism, tools risk obsolescence. RIECS must plan for maintenance, versioning, hosting, and community support structures that extend beyond project cycles.

Contributor onboarding and licensing frameworks influence participation: Clear terms for data use, attribution, and access improve trust and reduce friction. Projects that invested in transparent onboarding, consent, and licensing achieved more stable participation. RIECS should integrate these processes as part of the infrastructure's core services rather than peripheral guidance.

Citizen-generated data still struggles to enter institutional workflows: Even mature projects find it challenging to connect their outputs to official data flows. Barriers include inconsistent metadata, missing quality indicators, and unclear validation histories. RIECS will need mechanisms that facilitate the movement of data from community settings into institutional evidence processes while preserving provenance and context.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Documentation and reproducibility remain weak points: Many device-specific solutions are only partially documented, limiting reuse. The matrix highlights that documentation quality strongly correlates with integration potential. RIECS should promote documentation standards and provide templates and tools that lower the effort required for proper record keeping.

Engagement and technical development must co-evolve: Projects that integrated user requirements from early stages achieved smoother deployment. This confirms that co-design is not a one-off exercise but a continuous governance practice. RIECS should formalise participatory mechanisms to ensure that technical evolution follows community needs.

6.2 Domain-specific challenges

Domain-specific challenges arise from the inherent technical constraints and methodological requirements of distinct scientific disciplines, each imposing unique instrumentation demands, data validation protocols, and analytical frameworks. **Environmental** monitoring projects encounter sensor drift, calibration instability, and spatiotemporal sampling biases that compromise data accuracy across heterogeneous deployment conditions. **Health-focused** initiatives face stringent privacy compliance, clinical validation requirements, and participant safety protocols that necessitate specialized data governance architectures and consent management systems. **Climate** projects struggle with long-term data continuity, temporal resolution mismatches, and phenological observation standardization across diverse geographical and ecological contexts. These domain-specific constraints create specialized technical debt within sensor networks, data pipelines, and validation workflows that cannot be easily generalized across scientific disciplines, requiring domain-expert knowledge for proper system architecture and quality assurance implementation.

To better understand the specific challenges, four case studies were selected. These case studies were chosen based on their availability of documentation, the ability to address a key infrastructure challenge, their level of innovation, and their long-term sustainability, which contributes to a significant learning curve.

6.2.1 Environment & Biodiversity: Cos4Bio and Cos4Env

6.2.1.1 Cos4Bio: Expert portal for biodiversity validation

An online portal that integrates biodiversity observations from multiple citizen observatories

Cos4Bio is the biodiversity-centric expert portal that the H2020 Cos4Cloud project released in January 2022, after an agile prototyping cycle that began in late 2021; the service reached TRL-9 and was onboarded to the EOSC Marketplace as a fully operational, FAIR-compliant resource. Built on the general-purpose integration platform delivered earlier in the project, its core mission is to give taxonomic specialists a single, real-time gateway where they can locate, download and validate citizen-science observations drawn from multiple observatories such as iSpot, Natusfera or Pl@ntNet

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

via a common Darwin Core API while feeding their identifications and comments straight back to the sources as explained [25] in Figure 8.

To achieve that mission the portal wraps a Darwin Core–driven interoperability layer with a suite of expert-oriented services: Authenix single-sign-on; a seven-language interface that mirrors familiar citizen-science workflows; dual search (by species name or place) plus faceted filters for portal, taxon rank, data-quality flags, licence and date-range; observation detail pages supporting identifications and threaded comments; CSV/JSON download pipelines that record the user's reason for download and keep a re-usable history; feedback forms; personal dashboards summarising each expert's contributions; and a public KPI dashboard for the whole service. All incoming records are normalised to Darwin Core terms and tagged with their original Creative Commons licence (CC0, BY, BY-NC, BY-SA), so users can search or export data with confidence in provenance and re-use conditions. The entire codebase is openly published in the Bineo-Consulting/Cos4Cloud repository to encourage adoption and reuse.

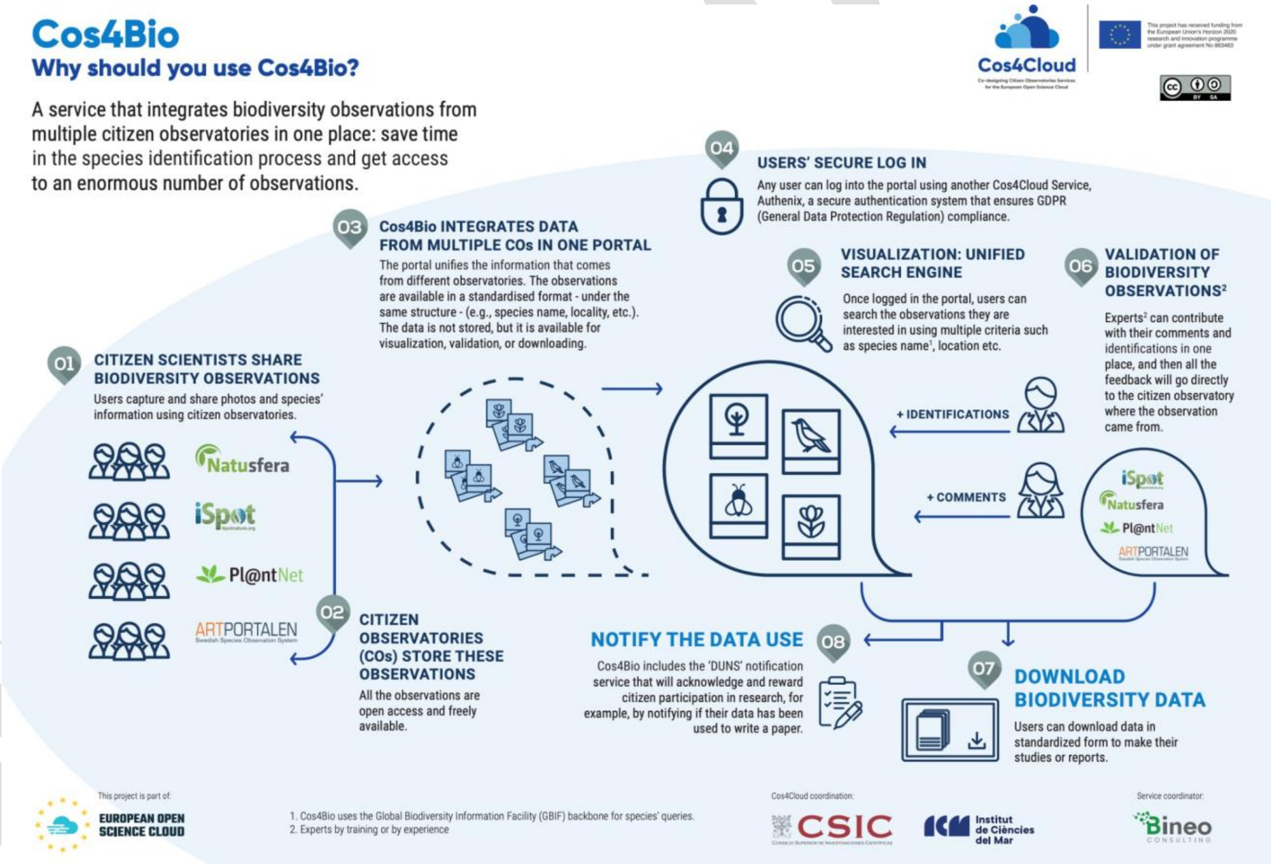

Figure 8 Cos4Bio purpose and main functions. Source: https://cos4cloud-eosc.eu/services/cos4bio/

The benefits ripple outward. For individual experts Cos4Bio radically reduces discovery time, provides one-click, standards-based downloads and lets them showcase their impact across observatories; for the observatories, it brings-in more timely, high-quality identifications; and for the wider research community, it represents a single FAIR data endpoint that unifies otherwise siloed

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

citizen-science streams. The main challenges faced by Cos4Bio and the solutions implemented are summarized in the table below.

Challenges	Solutions
Harmonising heterogenous CO data models	Each citizen observatory (iSpot, Natusfera, Artportalen, Pl@ntNet…) stored observations with its own taxon lists, metadata and licence policies. The team had to create an interoperability layer that maps every incoming record to a common Darwin Core (DwC) profile and resolves scientific names against the GBIF backbone. This required building a mapping API and a background normaliser pipeline.
Real-time aggregation at scale	Unlike GBIF's batch-publishing workflow, Cos4Bio streams fresh records from several CO APIs. Caching, request-throttling and a fail-fast strategy had to be added so that a slow or unavailable CO does not block search results.
Fine-grained search & download performance	Experts expect faceted search (taxon rank, portal, data-quality flags, licence, date-range) plus CSV export of up to millions of rows. The team introduced server-side pagination, pre-signed download files and column-projection to keep response times below ~5 s.
Federated identity & contribution tracking	The service must recognise the same expert across portals, store her identifications/comments, and display personal dashboards. Integration with Authenix SSO and a new profile module meant dealing with GDPR, multiple OAuth providers and linking user IDs to remote CO accounts.
Multilingual, lightweight front-end	A single Stencil-JS component library renders the portal in six languages and is packaged so partner COs can embed search widgets on their own sites.

6.2.1.2 Cos4Env: Expert portal for environmental data validation

A service that integrates environmental data from multiple citizen observatories in one place
Building on the same architecture as Cos4Bio, Cos4Env extended the platform to sensor-based environmental measurements, as explained in detail in Figure 9. Cos4Env targets variables such as

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

odour intensity, PM particulates, temperature, humidity and CO_2. It re-uses Authenix login, i18n and download modules, but swaps Cos4Bio's image grid for a WebGL map with clustering and bounding-box queries. Minimum-viable functionalities agreed with domain experts include location-centric search, dynamic filters (e.g., portal, measurement type, date, licence), CSV export, observation detail with time-series, feedback forms and cached query histories [26].

Because environmental data mix numeric units and diverse semantics, the team added a MeasurementOrFact extension to the Darwin Core mapping service and harmonised units server-side. Licence management had to embrace Open Data Commons ODbL v1.0 alongside the Creative Commons set, so users can filter or mix datasets under CC0, CC-BY, CC-BY-NC, CC-BY-SA or ODbL terms. The front end is written in Stencil JS, while the back end stays in Node/Express with Swagger docs and a Docker image for local deployments; the source is published at GitHub Bineo-Consulting/Cos4Env.

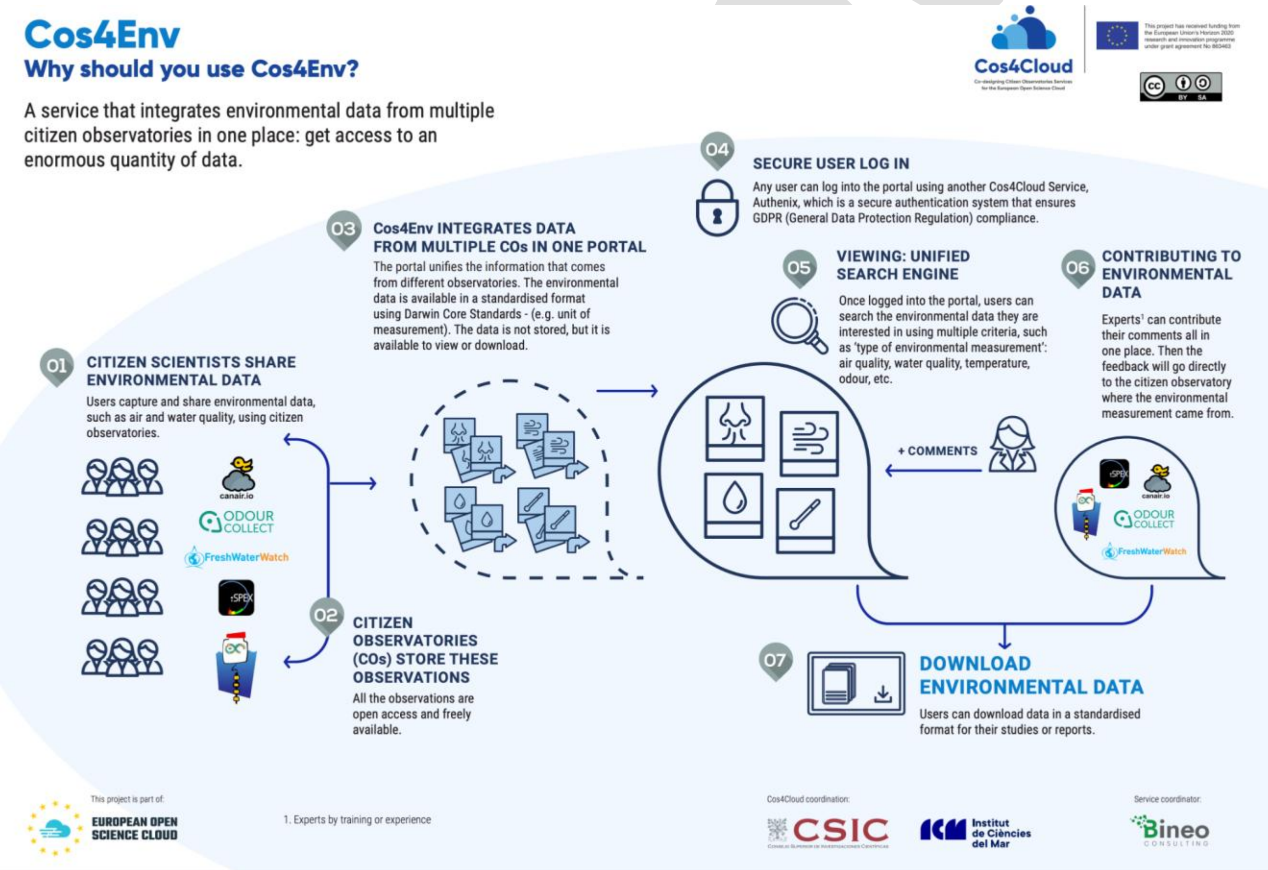

Figure 9 Cos4Env purpose and main functionalities. Source: https://cos4cloud-eosc.eu/cos4env/

The challenges faced by Cos4Env and how these were addressed are summarised in the table below.

Challenges	Solutions

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Extending DwC for sensor/odour data	Environmental COs (OdourCollect, CanAirIO, Canário) produce numeric *measurements* rather than taxa. An extra MeasurementOrFact extension had to be added to the DwC-based API and every incoming JSON record remapped accordingly.
Variable heterogeneity & units	The platform had to unify PM2.5, CO_2, VOCs, temperature, humidity, odour intensity & hedonic tone—each with its own units, precision and valid ranges—into a single search/filter vocabulary and guarantee unit consistency in downloads.
Geospatial UI and clustering	Unlike Cos4Bio's grid of images, Cos4Env needed an interactive map with dynamic clustering (tens of thousands of points) and pop-ups showing time-series of sensor readings. This required client-side WebGL rendering plus server-side bounding-box queries.
Licence diversity (ODbL vs CC-*)	Sensor networks often publish under Open Data Commons licences, while Cos4Cloud had standardised on Creative Commons. The download module had to expose mixed-license datasets and warn users when ODbL share-alike terms apply.
Same user-metrics engine, new data	Personal dashboards and global KPIs (comments count, downloads by variable, professional profile stats) had to be rewritten to consume measurement-level events instead of species identifications.

6.2.1.3 Lessons learned for RIECS

Experience from developing **Cos4Bio** and **Cos4Env** offers several lessons for designing a European research infrastructure capable of integrating heterogeneous CS data streams across domains. These two services operationalised interoperability across biodiversity observations and environmental sensor data, demonstrating both the potential and the limits of a shared architectural approach.

The experience demonstrates the need for a **modular and extensible integration architecture** capable of supporting heterogeneous data models, **co-design practices embedded in governance**, and **domain-aware pipelines** that acknowledge the methodological specificities of biodiversity, environmental monitoring, health, and other scientific areas. It also highlights the importance of **structured onboarding procedures**, **clear licensing pathways**, and **mechanisms that sustain FAIR**

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

compliance over time. Finally, the use of **community-facing services that make contributions visible** strengthens trust and supports ongoing engagement.

A common interoperability layer is feasible, but extensibility is essential

Both portals relied on a **Darwin Core–based integration layer** as the unifying standard. This choice allowed multiple observatories to be onboarded through a controlled mapping process and facilitated the creation of a unified search, validation, and download ecosystem. However, the experience also showed that **domain-agnostic standards require domain-specific extensions**.

- Cos4Bio required taxonomic name resolution, image-handling workflows, and licence propagation mechanisms.
- Cos4Env demanded a **MeasurementOrFact extension**, unit harmonisation, and semantic alignment across odour, air-quality, and microclimate variables, which are not native to DwC.

This points to a principle for future infrastructures: **a shared core standard is viable, but only if supported by a governance model that anticipates systematic extensions** and provides clear versioning, compliance tests, and backward-compatibility pathways.

Interoperability work cannot be decoupled from co-design

Both portals evolved through continuous interaction with domain experts, which directly shaped functionalities such as:

- Faceted search and high-volume CSV/JSON exports for biodiversity experts,
- WebGL geospatial visualisation and bounding-box queries for environmental variables,
- Dashboards and contribution metrics that responded to expert expectations for recognition and traceability.

The backlog tables in the co-design documentation show how expert feedback drove architectural decisions, including map services, search-by-place, mixed licensing support, and the download-reason capture service. Technical coherence alone does not guarantee usability; cross-domain infrastructures need continuous participatory design cycles, not one-off consultations, especially when integrating communities with distinct validation cultures and data-quality norms.

Domain specificity generates technical debt that must be managed explicitly

Even with a shared platform, the biodiversity and environmental portals confronted different types of technical constraints:

- Biodiversity relied on **taxonomic backbones**, photographic evidence, and expert commenting workflows tightly coupled with species concepts.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

- Environmental data required **unit conversion, sensor metadata handling, time-series visualisation**, and accommodation of calibration drift or missing metadata.

These differences produced domain-specific *technical debt* inside the integration layer—e.g., multiple schema adjustments, evolving mapping APIs, and performance tuning for high-density environmental records. A future infrastructure should therefore **budget for domain-specific pipelines** rather than assuming that a single generic workflow will serve all CS domains.

Mixed licensing regimes require careful design of user-facing services

Cos4Bio worked largely with Creative Commons licences (CC0, BY, BY-NC, BY-SA), while Cos4Env had to integrate **Open Data Commons ODbL** terms. This required redesigning the download module so that users receive clear alerts on re-use constraints and that licence terms are preserved in all derived datasets.

Handling heterogeneous licences at scale is a precondition for any future research infrastructure, particularly one that aims to support public-sector decision making and cross-border data use. Lessons from Cos4Env show that **licensing interoperability must be built into the platform's architecture and not treated as an external legal layer**.

Real-time aggregation demands robust caching, throttling, and fail-fast strategies

Cos4Bio processed high-volume biodiversity observations coming from multiple APIs. Efficient caching and a **fail-fast approach** ensured that a slow observatory did not block portal-wide search responses.

Cos4Env encountered similar challenges, but with denser record distributions due to continuous sensor measurements, requiring WebGL rendering and server-side spatial indexing. Together, these experiences highlight that **scalability is not only a backend issue**: it affects front-end design, API contracts, onboarding processes, and expectations about update frequency.

Validation workflows benefit from unified interfaces but must respect each observatory's rules

Cos4Bio allowed experts to annotate and identify species while automatically routing feedback back to each observatory, which then applied its internal validation rules. This mechanism respected observatories' autonomy while still enabling cross-platform quality improvement.

Cos4Env adapted this model for comments on environmental measurements, although the nature of validation differed (e.g., contextual interpretation rather than species identification). The broader lesson is that **a shared validation interface improves expert engagement**, but the infrastructure must

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

not override domain-specific quality-assurance paradigms. Governance must allow observatories to maintain their internal epistemic rules while participating in a shared validation ecosystem.

Contribution dashboards support transparency and community engagement

Both services integrated metrics on downloads, comments, and expert activity, giving visibility to contributions across observatories. These dashboards satisfied a long-standing request from experts for better recognition and also provided observatories with insights into how their data are reused.

This suggests that any future European infrastructure should include **native tracking and reporting services** rather than treating them as optional add-ons. Recognition metrics are not merely engagement features; they are mechanisms that **reinforce data stewardship, transparency, and accountability**.

Coherent onboarding workflows are central to sustainability

Onboarding new observatories required clear API documentation, mapping templates, licence tracking, and support for testing integration before publication. The Cos4Bio sustainability notes emphasise that **easy onboarding is a prerequisite for long-term service adoption in EOSC**. A future European infrastructure will need: formal onboarding protocols, validation of FAIR compliance, automated schema and licence checking and test sandboxes to reduce integration burden.

A shared platform enables cross-domain synergies but does not eliminate domain boundaries

Although both portals were built on the same general-purpose integration platform, the resulting services diverged due to disciplinary requirements. The experience shows that **architectural and data convergence is useful and needed**, yet **full homogenisation is neither realistic nor desirable**. A future infrastructure should adopt a **federated model**:shared core services (authentication, search, download, metrics), domain-specific modules for specialised processing and interoperable but autonomous validation workflows.

FAIR and EOSC compliance require continuous alignment, not one-time certification

Both portals invested significant effort in FAIR-aligned metadata, standardised download formats, open APIs, and open-source releases. However, maintaining FAIRness in a dynamic, multi-observatory ecosystem requires ongoing alignment, especially when observatories update their APIs, introduce new variables, or change licensing. This suggests that a European infrastructure must include **FAIR-by-design governance**, including monitoring services, automated metadata quality checks, and policies ensuring that observatory updates do not break interoperability.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

6.2.2 Health: DHIS2 Infrastructure

Digital Health Platform: https://dhis2.org/

DHIS2 is the world's largest open-source Health Management Information System, used as a national system-of-record in more than 80 low- and middle-income countries and by major global health initiatives. Designed as a flexible platform rather than a health-specific tool, **it supports integrated data collection, validation, analysis, and visualization,** allowing real-time decision-making across health programs while remaining adaptable to sectors such as education, logistics, and agriculture. Its development is coordinated by the HISP Centre at the University of Oslo within a global network of 23 regional groups that work directly with ministries to implement, customize, and maintain the system. DHIS2's modular architecture, open API, metadata-driven configuration model, and adherence to global data standards allow countries to integrate it with other systems, upgrade efficiently, and operate in low-resource environments through features like offline mobile apps and SMS reporting. Countries retain ownership of their DHIS2 instances and data, while a large community of government users, NGOs, developers, and international partners contributes to continuous improvement and widespread adoption.

6.2.2.1 Challenges and Solutions in DHIS2's Infrastructure

Over its 30-year evolution, DHIS2 has encountered and addressed numerous technical and organizational challenges. Summarised below in Table 1, key challenges and the solutions adopted include:

Infrastructure & Connectivity: Challenge: DHIS2 is heavily used in remote and low-resource settings where internet connectivity, electricity, and hardware are limited. Health facilities often had only paper forms or offline computers, making real-time data reporting difficult. **Solution:** DHIS2 embraced an offline-first approach. The DHIS2 Android app allows data entry on smartphones or tablets entirely offline, syncing later when connectivity is available [27]. Even SMS-based reporting is possible for basic phones. The web application is optimized for low bandwidth, and countries can host servers nationally to improve access speeds. For example, during Ebola outbreaks and in rural clinics, health workers could capture data on battery-powered devices and upload when online, preventing data loss. This approach effectively bridges the digital divide, allowing DHIS2 to function as critical infrastructure even in underserved areas.

Integration & Data Silos: Challenge: Health information ecosystems are often fragmented – different programs (HIV, TB, logistics, hospitals) use separate systems that don't talk to each other. This siloing hinders extensive analysis. **Solution:** DHIS2 was designed with interoperability in mind, featuring a well-documented open API and modular architecture [27]. This allows it to serve as a central data warehouse that other tools can plug into. A wide range of integrations and interoperability layers have been developed: e.g. OpenFN middleware to connect DHIS2 with external databases, and plug-ins to feed DHIS2 data into business intelligence tools like Tableau. Support for international

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

standards like HL7 FHIR means DHIS2 can exchange data with electronic medical record systems or lab systems using common formats. These integrations reduce data re-entry and ensure that DHIS2 can sit within a larger digital health infrastructure rather than operating in isolation. Countries like Ethiopia and Sri Lanka, for instance, integrate DHIS2 with logistics management systems and disease surveillance apps, creating a more unified information system.

Local Customization vs. Maintainability: Challenge: Each country or project has unique data requirements (different indicators, languages, health system structure). If everyone modified the core software for their needs, it would fragment the platform and complicate upgrades. **Solution:** The DHIS2 team adopted a *configure not code* philosophy. **Core development focuses only on generic features needed across multiple contexts, while country-specific needs are met through configuration or custom apps without altering core code** [27]. The platform's flexibility (dynamic metadata model) allows users to design forms, reports, and business rules that reflect local needs, all through the user interface or minor extensions. This means a health ministry can, say, add a new disease surveillance form or modify an indicator formula themselves, without any programming. If truly novel functionality is needed, it can often be added as a separate app/module via the DHIS2 API. This modular extensibility has been key to DHIS2's longevity: countries get tailored systems, yet all still run fundamentally the same core platform, benefitting from common upgrades. An example of this approach is how DHIS2 handled COVID-19: instead of creating a new system, countries rapidly built COVID case registries and dashboards as configurations on DHIS2's tracker module – building on the existing platform and then sharing these configurations globally.

Human Capacity & Data Use: Challenge: Deploying a national system is not just a technical job – users must have the skills and motivation to enter high-quality data and use that data for decisions. In many LMICs, there have been gaps in data analysis capacity at local levels and resistance to using data (relying on habit or hierarchy instead). Additionally, staff turnover can erode capacity. **Solution:** DHIS2's sustainability heavily relies on investments in training and capacity building. The HISP network and UiO have trained thousands of people through DHIS2 Academy courses and onsite mentoring [27]. Over 70 PhD graduates from the Global South have specialized in health information systems via HISP, many of whom now lead country implementations. These local experts provide day-to-day support to health offices and help cultivate a data-use culture. HISP groups facilitate regular workshops, on-the-job training, and a global Community of Practice where implementers troubleshoot and share best practices. The result is a growing cadre of in-country champions who understand both the technology and the health context. For example, in the Democratic Republic of Congo (DRC), a recent assessment found that many facility-level workers lacked data analysis skills, limiting their use of DHIS2 data. The proposed solution was targeted data analytics training and mentorship at those levels. **Indeed, across countries, a key lesson has been that providing user-friendly dashboards is not enough – continuous capacity building and engaging users in the system design are essential so that data is actually utilized [28]**.

Governance & Coordination: Challenge: Health information systems often span multiple programs and departments, raising questions of governance: Who *owns* the system? How to coordinate donors

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

and avoid duplicate systems? Some countries initially faced fragmented governance where different ministries or vertical programs ran parallel data systems, complicating a unified strategy [28]. **Solution:** HISP advocates for strong government ownership and coordination mechanisms. Typically, a Ministry of Health unit is designated to govern the DHIS2 platform (often an HMIS or eHealth unit), with clear roles and responsibilities defined among departments. Regular stakeholder coordination meetings and technical working groups are established to manage requirements and integrations [28]. For instance, in DRC, the HISP team recommended routine coordination meetings among ministries and partners to oversee HIS strengthening and minimize silos. **The global DHIS2 governance also plays a role: HISP UiO coordinates a software roadmap process where requirements from various countries and partners are evaluated and prioritized transparently**. This ensures the core platform evolves according to shared needs, reducing uncoordinated custom developments. Over the years, many countries have developed data policies and governance frameworks around DHIS2 – covering data standards, access control, and privacy – often with support from WHO and donor projects. As DHIS2 has become a part of national infrastructures, countries foster inter-departmental trust in the system. In practice, this means immunization, HIV, and primary care programs, for example, all agree to use the common platform and contribute to its governance, rather than each having their own databases.

Security, Privacy & Ethics: Challenge: Handling sensitive health data (like HIV status or personal details) raises important ethical and security concerns. Many implementing environments initially lacked mature cybersecurity practices or dedicated IT security officers (illustrated by research titled *Where There is No CISO*). Ensuring data privacy and patient confidentiality in a large, distributed system is an ongoing challenge. **Solution:** DHIS2 approaches this through both technical and organizational means. Technically, the platform includes robust access control, encryption for data in transit, and audit logs. Admins can configure user roles so that, for example, a district officer only sees data for their district. The DHIS2 team also provides guidance via a Trust Center outlining security best practices and privacy principles (such as compliance with data protection regulations) [27]. Organizationally, HISP has promoted the concept of data ownership and stewardship by the local authorities, meaning ministries must put in place governance that respects patient rights and secures data. Countries like Uganda and Tanzania have developed data privacy guidelines as part of their DHIS2 implementations, often supported by HISP experts. **The open-source nature of DHIS2 additionally allows code audits by the community to identify and fix vulnerabilities, contributing to more secure software over time.** While challenges remain (e.g., ensuring every health worker is trained in data confidentiality), the platform's longevity can also be attributed to trust earned through its ethical stance on country data sovereignty and privacy.

Financial Sustainability: Challenge: Maintaining a national-scale digital infrastructure requires ongoing funding – for servers, IT staff, training, and software development. Early on, many DHIS2 implementations and the core development relied on donor funding (e.g. grants from global health initiatives). A risk is that if external funding is cut (as seen recently with some donors reducing support post-COVID), systems could go offline or stagnate [29]. **Solution: One key to DHIS2's resilience has**

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

been its low cost of ownership and ability to be sustained with relatively modest resources compared to proprietary systems. Being free to license, countries mainly need to invest in hosting and staffing. Many ministries have gradually moved DHIS2 costs into their domestic budgets, at least for keeping the system running. **HISP UiO and partners are now actively encouraging a transition to locally-funded support models – for example, by having countries contribute to core development funding, and by making the costs visible in national health budgets** [28]. **This increases local commitment and reduces sole reliance on foreign aid.** Moreover, because DHIS2 focuses on essential routine data, it has proven its value to health ministries, who are more likely to allocate funds to a system that clearly supports day-to-day decisions. Indeed, a 2025 rapid survey found that while many donor-funded vertical data systems shut down when funding was pulled, all national DHIS2-based HIS in surveyed countries remained online with local staff maintaining routine data collection [42]. T**his highlights a fundamental principle: investing in what works and lasts – locally owned information systems, the digital public goods they run on, and local capacity – yields more resilience** [28]**.** Going forward, a diversified financing approach (domestic funds supplemented by multi-donor pooled funding for core development) is being pursued to ensure DHIS2's sustainability through economic ups and downs. The continued support of global stakeholders (as seen in DHIS2 Investor meetings) and alignment with international strategies (like the Lusaka Agenda for sustainable health systems financing [28]) provide a favorable environment for DHIS2 to thrive in the long term.

Table 1: Key challenges and DHIS2's approaches to sustainability

Challenge or Need	DHIS2 Approach and Solution
Limited connectivity & tech infra (rural areas, low bandwidth)	Offline-first design – e.g. Android app with offline data capture, SMS reporting [27]. Light web apps optimized for low bandwidth. Ensures functionality in remote, resource-constrained settings.
Diverse local requirements (varying data elements, languages, workflows)	Flexible metadata configuration and modular apps instead of hard-coding. Core software only implements generic features, with local customization done via configuration [27]. Allows adaptation without forked code, so all users benefit from core updates.
Data silos & external systems (need to integrate multiple systems)	Open APIs and interoperability layers for integration. Supports standards (FHIR) for data exchange. Many plug-ins and middleware (e.g. OpenFN) connect DHIS2 with other software, enabling a unified data ecosystem [27].
Low data use and skills (staff not analyzing or using data)	Massive capacity building via HISP network: on-site training, DHIS2 Academy courses (thousands trained). Developed local experts (70+ PhDs, MSc) in each region. Community of Practice forum for continuous support. These efforts improve data literacy and cultivate a data-use culture [27].

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Fragmented governance (multiple stakeholders, parallel systems)	Promotion of in-country ownership – each country runs its own instance with Ministry governance [27]. Establishing coordination committees and clear roles for HMIS management [28]. Aligning donors to use the national DHIS2 instead of creating new systems.
Security & privacy concerns (sensitive health data)	Role-based access controls, audit logs, and encryption in the platform. Guidance via DHIS2 Trust Center on best practice [27]. Emphasis on data sovereignty – countries host their data under local laws [28]. Open-source transparency allows security auditing by community.
Long-term financing (sustainability beyond donors)	Open-source (no license fees) keeps costs low. Encouraging governments to budget for HIS staff/infrastructure. Core development funded by diverse grants; moving towards country co-investment [28]. Focus on routine data that proves value to national programs (justifies budget priority) [28].

6.2.2.2 Key factors in DHIS2's 30-Year success

Over three decades, technical innovations and social strategies have combined to make DHIS2 a durable digital infrastructure. Key factors include:

Open Source & Flexibility: From the outset, DHIS2's open-source nature under a liberal license meant anyone could use or improve it [27]. This enabled broad adoption and a community of contributors. Its flexible, generic design (configurability) allowed it to meet evolving needs without requiring bespoke redevelopment for each context. Openness also fostered trust – countries knew there was no vendor lock-in and that they could own their solution.

User-Driven evolution: DHIS2's development has been continuously driven by real-world user requirements. Through an iterative process (quarterly release cycles and community feedback), features are added based on common needs across countries. For example, the demand for mobile data entry led to the Android app, and requests for GIS analytics led to integrated mapping features. This user-centered agility helped it keep pace with changing public health priorities (from HIV in the 2000s to COVID-19 in 2020).

Global Collaboration & Local Ownership: A unique socio-technical approach underlies DHIS2: a global network supporting local action. The University of Oslo's HISP Centre provides stewardship and quality assurance for the software, while local HISP groups and ministry teams ensure the system is embedded in national contexts[30]. This distributed model meant DHIS2 benefitted from international expertise and donor resources, but implementations were owned by local institutions (ministries of health, etc.) [27]. Such local ownership has been critical for longevity – countries are

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

invested in *their* system and continue using and improving it even when external projects end [29]. The HISP network's long-term presence (often via partnerships with local universities or NGOs) has ensured that knowledge and support are sustained on the ground, not just through fly-in consultants.

Capacity building and community: DHIS2 was not just delivered as software; equal emphasis was placed on building a human infrastructure. The investment in training thousands of users, cultivating local DHIS2 experts (many through academic programs), and facilitating peer support created a robust community of practice. This community became a self-reinforcing asset – when challenges arise, there is a pool of knowledgeable people to solve them, and when new staff come in, they have resources to learn. This lowers the risk of system abandonment. Notably, HISP's approach of pairing informatics and public health education (e.g. integrated MSc programs) produced professionals who bridge the gap between technology and health needs, which is vital for effective HIS design and use.

Supportive Governance & Policies: The longevity of DHIS2 also owes to supportive governance structures. **International recognition as a Digital Public Good gave it legitimacy and attracted funding**[4]. Many countries incorporated DHIS2 into their national eHealth strategies and policies, making it the official system for health data reporting. High-level buy-in (e.g. ministerial endorsements, use of DHIS2 data in health sector reviews) helped shield the system from political changes. **The governed open-source model, where stakeholders have a say in the roadmap, created a sense of collective governance.** Ethically, DHIS2's respect for country data sovereignty and privacy built trust – governments saw it as their platform, not an imposed one, encouraging long-term commitment.

Adaptability and Innovation: Technically, DHIS2 has been adaptable – both in incorporating new technologies and in scaling performance. It transitioned from a desktop software in the 1990s to a web-based system, and now to cloud-friendly deployments, showing an ability to modernize. The architecture's scalability (caching, database tuning, etc.) has been improved to handle national datasets with millions of records. The core team's professionalization in 2012 (hiring full-time developers, architects, etc. [27]) was a turning point that improved software robustness and release management as the user base grew. Equally important, DHIS2 implementers showed creativity in local innovations – for example, inventing novel uses like using DHIS2 for tracking commodity stock or climate data, which then informed new features for all. This continuous innovation mindset helped DHIS2 remain relevant and prevented obsolescence.

Sustainable financing & Partnerships: Surviving 30 years required money and partnership support. DHIS2 benefited from a coalition of donors (governments, UN agencies, NGOs) investing in its development and deployment as a global public good. However, as noted, the shift towards domestic financing is increasing. The fact that DHIS2 addresses routine health information – a fundamental need – means governments and donors see value in keeping it running. Even during funding downturns, partners rallied to keep core systems online [29]. Additionally, partnerships with organizations like WHO, University collaborations, and an ecosystem of tech companies providing DHIS2 services have created a safety net; DHIS2 is not solely dependent on one entity for survival.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

The collective commitment was evident during crises like COVID-19, where multiple partners collaborated to rapidly roll out DHIS2 modules for case surveillance and vaccination in dozens of countries. This broad partnership base and alignment with global health priorities (e.g. SDG monitoring) have been key to its endurance.

In summary, DHIS2's longevity can be attributed to this blend of technical excellence (flexible architecture, scalability, interoperability) and socio-technical foundation (community capacity, good governance, openness). DHIS2 has managed to operate and evolve over 30 years thanks to focusing on country needs and public value, an achievement few information infrastructures in global health can claim. It exemplifies how ethical, inclusive practices – like respecting local ownership and encouraging open knowledge sharing – reinforce technical robustness and adoption over time [29].

6.2.2.3 Lessons learned for RIECS

For a potential infrastructure like RIECS there are many takeaways from DHIS2's experience. Key lessons include:

Adopt an open, modular architecture: Design the system as an open platform with modular components and APIs. This ensures others can extend the system and integrate it with existing tools, increasing its utility. DHIS2's open API and plug-in model allowed it to become a data hub rather than a silo [27]. RIECS should similarly prioritize interoperability and open standards to fit into users' ecosystems, not forcing one-size-fits-all usage.

Design for low-resource environments: If targeting diverse contexts, build for the least optimal conditions (low internet, older devices, limited IT support). DHIS2's offline and lightweight features were fundamental for its global spread. RIECS can learn to invest early in offline capabilities, efficient performance, and user-friendly interfaces that do not assume high-end infrastructure. This widens the potential user base and demonstrates inclusivity.

Focus on users' needs and local capacity: A system will only be sustained if it serves the real needs of its users. Engage end-users and local stakeholders in design and iteration. DHIS2 thrived by responding to health workers' and managers' requirements and by enabling local configuration for local problems[31]. Equally, invest in building local capacity to use and administer the system. RIECS should plan training programs, create community forums, and perhaps partner with educational institutions to develop skilled practitioners. This not only improves system uptake but also creates champions who advocate for the system's continuation. The current scope of the RIECS-Concept responds to this recommendation through the active involvement of a wide range of stakeholders in shaping user requirements and contributing to the conceptual design of the infrastructure.

Ensure strong governance and ownership: Define clear governance structures that involve the primary beneficiaries (e.g., public institutions or communities) in decision-making. DHIS2's success in countries came when Ministries took ownership and coordinated partners around one system [29]. For RIECS, if the RI is meant for public or multi-institution use, establishing a governance board or

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

network (like the HISP network) could help balance centralized quality control with decentralized ownership. This federated model can maintain coherence of the core product while empowering local innovation. Also, addressing data governance, privacy, and ethics from the start builds trust – users must feel their data is safe and under their control.

Plan for sustainability (social and financial): Technology alone does not sustain an RI; consider long-term financing and community engagement. DHIS2's journey shows the importance of transitioning from donor support to local funding and including maintenance costs in budgets[40][43]. RIECS should evaluate its value proposition to stakeholders and aim to demonstrate results early (so funders see its importance). Building an active user community (through open-source contributions or user groups) creates a sense of collective investment in the tool's future. Additionally, aligning RIECS with broader policy goals or public needs can garner institutional support (similar to how DHIS2 became central to national health strategy in many countries). Ultimately, planning for longevity might mean modular growth – start with core essential features that have clear impact (the *routine data* equivalent), prove success, and then scale up with more features and funding once credibility is established.

6.2.3 Climate: ICOS ERIC

ICOS ERIC (Integrated Carbon Observation System) is a pan-European RI focused on greenhouse gas monitoring, making it highly relevant to climate science. Recognized as an ESFRI Landmark in the Environmental domain, ICOS conducts long-term observations of atmospheric, oceanic, and terrestrial ecosystems to provide high-precision data on greenhouse gas fluxes and the carbon cycle [4]. Its network comprises over 170 standardized measurement stations across 16 countries covering atmosphere, ecosystem, and ocean domains [32]. These observations support climate-change research and policy, delivering open-access data critical for understanding emissions, sinks, and trends in support of European climate objectives [4]. ICOS's inclusion in the ESFRI 2024 Landscape Analysis highlights its strategic importance for Earth-system science and its status as a mature, operational infrastructure serving transnational user communities [4].

6.2.3.1 Challenges and solutions

ICOS addresses a set of core challenges typical of large-scale environmental observation infrastructures: managing distributed data streams, ensuring consistent quality control, supporting diverse user needs, maintaining long-term governance and financial stability, and integrating with Europe's open science ecosystem. Its technical architecture responds to the difficulty of harmonizing heterogeneous measurements by enforcing a standardized flow from sensor to archive, with thematic centres and central laboratories responsible for calibration and validation. User-facing challenges linked to accessibility and reproducibility are met through the Carbon Portal, which offers open access, rich discovery tools, programmatic interfaces, and collaborative environments. Governance and sustainability issues are handled through the ERIC framework, which secures coordinated oversight and shared funding responsibilities across member countries. Alignment with EOSC resolves broader interoperability and integration concerns, as ICOS adopts FAIR practices, semantic metadata, and cloud-based services that connect its domain-specific outputs to European data and computing resources. These solutions, described in the sections below, illustrate how ICOS mitigates technical, organisational, and open science challenges to operate as a mature and reliable research infrastructure.

Technical architecture and data services

ICOS is organized as a distributed infrastructure with a well-documented technical architecture (Figure 10). Data flow in ICOS is standardized from sensor to archive: measurements are collected at national stations (operated by member countries) and immediately stored in safe repositories, then sent to thematic centers for domain-specific processing and quality control[32]. For example, atmospheric, ecosystem, and ocean Thematic Centres aggregate and calibrate the data from their respective networks, while Central Analytical Laboratories perform high-precision calibration and flask sample analyses [32]. Once quality-controlled, the data are transmitted to the central ICOS Carbon Portal – a core component of the architecture – which integrates all datasets into a single platform [32].

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

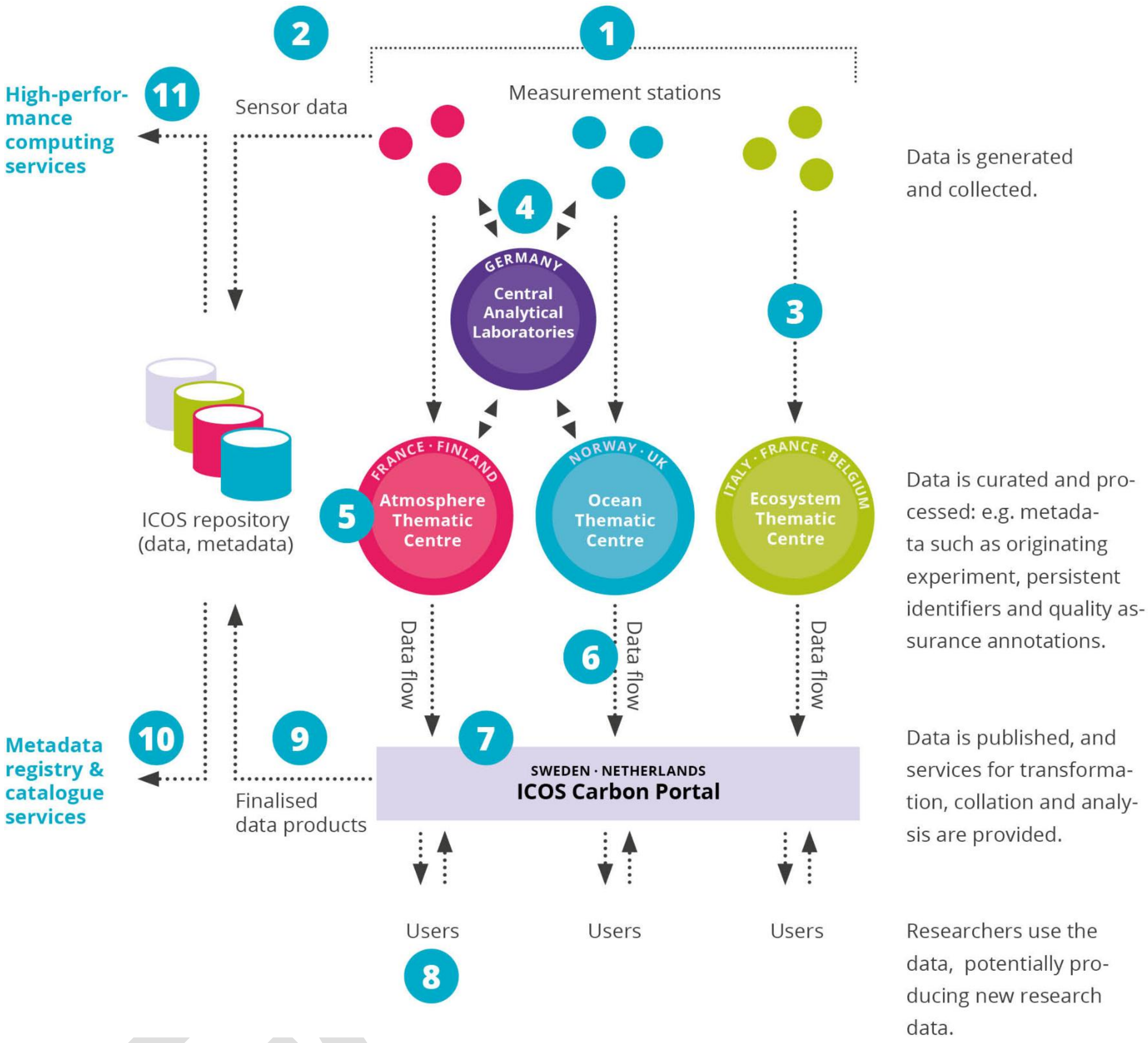

Figure 10 ICOS data flow architecture. Sensor data from distributed stations (1) are securely archived and forwarded to Thematic Centres (3,5) and Central Labs (4) for processing; curated datasets then reach the Carbon Portal (7), which serves as the one-stop shop for publication, DOI assignment, and user access [32]. The Carbon Portal employs robust e-infrastructure back-ends: all ICOS data products and metadata are replicated in a long-term repository using EUDAT B2SAFE storage, ensuring data preservation and adherence to FAIR principles. Additionally, ICOS collaborates with European e-infrastructures like EUDAT and EGI, enabling interoperability with the European Open Science Cloud (EOSC) and rapid data transfer to high-performance computing centers for analysis and modelling. This architecture showcases a FAIR and open science-aligned design, with standard data formats and cataloguing that allow integration into broader environmental data portals and services [32].

User services and open access interfaces

ICOS provides a rich suite of user-facing services and interfaces that are extensively documented and freely available. The ICOS Carbon Portal is the primary access point, offering web-based tools for data discovery, visualization (e.g. time-series plots, maps), and download of observational datasets.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Every dataset is assigned a citable DOI, and usage metrics are tracked, reflecting a strong commitment to open data and reproducible science. All ICOS data are made available under an open data license, and any user can access, view, and download data products without restriction. In fact, external researchers are even invited to contribute derived data products back to the portal, fostering a collaborative open science ecosystem [32].

Beyond the web portal, ICOS supports programmatic access via well-documented APIs and tools. For instance, the Carbon Portal provides a public SPARQL endpoint for querying metadata and semantic links [33], as well as an official Python library for convenient access to time-series data in scripts or notebooks. Advanced users can leverage Jupyter Notebook environments and analytical services hosted by the portal for on-site data processing and visualization. Moreover, ICOS offers specialized tools like the STILT atmospheric transport model for footprint analysis and a Nextcloud-based file share for community data exchange[34]. This service catalog – from graphical data portals to machine-to-machine interfaces – illustrates ICOS's alignment with open-science best practices and ensures that both human users and computational workflows can easily exploit its resources.

Governance structure and sustainability

The governance and organizational model of ICOS are well-established and transparently documented, aligning with the ESFRI guidelines for sustainable RIs. ICOS is operated as a European Research Infrastructure Consortium (ERIC), a legal framework that facilitates long-term international collaboration. The ICOS ERIC entity involves a central Head Office (for overall coordination) and the Carbon Portal (for data management and distribution), under the leadership of a Director General [35]. Strategic oversight is provided by a General Assembly of member country representatives, and scientific direction is guided by dedicated advisory boards. At the operational level, ICOS's distributed nature is reflected in its governance: National Networks in each member country manage the stations, and domain-specific Monitoring Station Assemblies ensure scientific and technical standards across the Atmosphere, Ecosystem, and Ocean networks. The Central Facilities (thematic centers and labs) work closely with these assemblies to maintain data quality and innovation in measurement techniques [35].

Importantly, ICOS has a clear sustainability plan built on its ERIC funding model. Member countries finance the core operations and long-term maintenance of ICOS: each country supports its national stations (via national research agencies) and contributes to the central ICOS ERIC budget [33]. This distributed funding approach ensures resilience and commitment, as critical functions are shared among countries rather than relying on a single host. The ICOS data policy and governance documents (available publicly) further codify its open-access mandate and the responsibilities of participants [33]. Through periodic evaluations and a five-year scientific review cycle, ICOS also adapts its strategy to secure ongoing relevance and financial support [35]. Overall, the governance structure of ICOS balances international coordination with national commitments, providing a stable foundation for sustainable operations and growth.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Alignment with EOSC and Open Science

ICOS is exemplary in embracing open science principles and integrating with the European Open Science Cloud ecosystem. As noted, all ICOS data are openly accessible under CC licenses, and the infrastructure invests in tools to improve data Findability, Accessibility, Interoperability, and Reusability (FAIR) [32]. ICOS actively participates in cross-RI and EOSC projects (e.g. ENVRI-FAIR, ATMO-ACCESS), ensuring that its data and services can be discovered and combined with other environmental data through EOSC portals and standards.

The Carbon Portal's use of DOI identifiers and semantic metadata allows ICOS datasets to be indexed in EOSC's catalogs and global repositories [32]. Furthermore, by using EUDAT cloud storage (B2SAFE) and connecting to EGI computing, ICOS exploited Europe's federated e-infrastructure – effectively bridging its domain-specific data into the broader EOSC cloud and HPC resources. ICOS also shares its data descriptions with Copernicus and other climate-data hubs, improving interoperability between in-situ observations and satellite-based services[32]. This positions ICOS as a key contributor to the EOSC, both as a provider of high-quality data and as a consumer of EOSC core services for authentication, storage, and processing. The commitment to open access, community engagement, and technical interoperability highlights why ICOS is frequently cited as a model in the ESFRI landscape for open science integration [4].

6.2.3.2 Lessons learned for RIECS

A key lesson for RIECS from the ICOS ERIC experience lies in the value of early institutional anchoring through the ERIC framework, which has allowed ICOS to sustain long-term operations, coordinate distributed infrastructures, and formalize national commitments. RIECS can adopt a similar federated model, where national nodes contribute to a common European infrastructure under shared governance, supported by strong central coordination and clear service mandates. ICOS demonstrates how alignment with EOSC, adherence to FAIR principles, and provision of machine-actionable metadata ensure that data produced in a specialized domain becomes reusable and impactful across disciplines.

Given its clear climate focus, extensive documentation, and mature service offerings, ICOS ERIC also illustrates what constitutes a robust, open infrastructure model for transnational science. It provides publicly accessible technical documentation, governance materials, and service descriptions, which together provide structured analysis of both operational and strategic components. For RIECS, the example of ICOS offers guidance on developing scalable data architectures, integrating with EOSC environments, and aligning services with policy goals. Moreover, ICOS's active support for user communities—through tools, notebooks, APIs, and curated access—can inform how RIECS builds participatory, user-responsive infrastructure.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

7. Matrix of prioritisation of challenges

Table 1 provides a high-level synthesis of key challenges, their main dimensions, and the corresponding solution approaches, together with an indication of priority derived from the full evidence base, including project documentation, case studies, and interviews. The table serves as a navigational aid that highlights the most salient challenges and solutions identified across the material, offering a concise synthesis to the themes addressed in Section 6.

This section is intended as an orientation tool rather than a substitute for the detailed analysis that was previously exposed. Readers are encouraged to consult the whole of section 6, which presents the specific challenges with richer context, methodological nuances, and concrete examples drawn from EU-funded initiatives and operational experiences.

Table 3 Key challenges for a CS Infrastructure – Dimensions, solutions, and priorities

Challenge	Dimension	Solution approaches	Priority
Fragmentation of platforms and data silos – Disconnected CS projects and tools obstruct collaboration and data reuse.	Architecture	Federated architecture integrating existing CS platforms and RIs; create a unified catalog of resources. Connect with open science clouds (e.g. EOSC) to link distributed data and services.	High – Fundamental to unlock full capacity of CS.
Lack of interoperability standards – Heterogeneous data formats and software APIs prevent seamless data exchange.	Data & Standards	Develop and adopt common standards (e.g. extended SensorThings API, Darwin Core for biodiversity). Enforce FAIR data principles (findable, accessible, interoperable, reusable). Establish cross-domain metadata schemas and APIs for interoperability. Mapping and translating between—national calibration models is essential to guarantee interoperability, reliable	High – Needed to enable integration across domains.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Challenge	Dimension	Solution approaches	Priority
		data integration, and meaningful cross-border analyses.	
Scalability constraints – Platforms struggle with massive data volumes and participant growth (e.g. astrophysics projects generating huge datasets).	Scalability	Utilize cloud infrastructure and distributed computing for data processing. Optimize platforms with modular, microservice architectures that auto-scale. Employ AI-assisted data analysis to handle data deluge while maintaining quality.	High – Essential as CS data and users rapidly increase.
Sustainable infrastructure & funding – Maintaining platforms and services beyond project lifetimes is uncertain.	Architecture / Governance	Adopt open-source development and community-maintained tools to reduce costs. Develop sustainability models (e.g. institutional backing, integration with national RIs). Include governance frameworks for long-term operation.	High – Critical for continuity and trust in the infrastructure.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Challenge	Dimension	Solution approaches	Priority
Multi-scale system design – Need to support local grassroots projects and large-scale EU-wide observatories simultaneously.	Architecture	Implement federated architectures that allow hierarchical data management – local nodes feeding into national/regional hubs and then into a European platform. Ensure flexibility to accommodate context-specific tools alongside global standards	Medium – Important for inclusivity, though initial focus is on higher-level integration.
Power asymmetries – Imbalances between professional scientists, platform providers, and citizen contributors (e.g. who controls data and decisions).	Socio-Technical	Promote participatory governance models that give community representatives a voice in decision-making Ensure transparent data policies and equitable benefit-sharing (e.g. open data access and citation credit for volunteers).	Medium – Addresses fairness and uptake, though indirectly affects infrastructure success.
Community engagement and formation – Turning ad-hoc crowds into sustainable communities; maintaining volunteer motivation and diversity.	Socio-Technical	Invest in community-building features (forums, feedback loops, recognition systems) Provide training and *cs tech literacy* programs to empower participants. Use co-design sessions to align platforms with user needs and values.	High – Without engaged communities, technical infrastructure will not be fully utilized.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Challenge	Dimension	Solution approaches	Priority
Integration with existing RIs and official systems – Difficulty aligning CS with established RIs and data frameworks	Socio-Technical / Standards	Use boundary organizations and liaisons to bridge CS with institutions Demonstrate data quality to encourage uptake in official monitoring (e.g. linking observatory data to GEOSS/Copernicus). Develop policies for data sharing with government and scientific databases.	High – Necessary for mainstream acceptance and policy impact.
Infrastructure transparency and usability – The platform should be "invisible" when not needed and visible when it aids users, to encourage adoption.	Socio-Technical / UX	Apply user-centered design so that tools are intuitive. Provide clear information on data provenance and uncertainties (e.g. dashboards showing sensor accuracy) to build trust. Use progressive disclosure – simple interfaces for newcomers, advanced options for experts.	Medium – Improves user trust and inclusion, though secondary to core interoperability issues.
Emerging AI integration – Utilizing AI for CS while preserving human engagement and addressing ethical concerns.	Emerging Tech (AI)	Integrate AI tools to assist with analysis (e.g. species identification, anomaly detection) but keep humans in the loop. Co-create AI solutions with citizen input (e.g. amai! model for public-guided AI) to ensure alignment with community values. Develop guidelines for transparent and fair AI use (preventing bias, protecting privacy).	Medium – Growing importance; can greatly expand capacity, but requires careful governance.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Key takeaways

1. Integration and standards are the foundation: Efforts in architecture, standards, and interoperability collectively shape the enabling layer of RIECS. These domains hold the highest priority because every other capability—scaling, AI integration, institutional uptake—depends on resolving fragmentation.

2. Scalability and modularity protect the system against growth shocks: The evidence shows that CS infrastructures often fail when participation or data volumes increase suddenly. Cloud-based and microservice approaches are not optional enhancements but structural safeguards.

3. Community engagement is the determining factor for actual use: Technical solutions only succeed when socio-technical conditions support adoption. Sustained participation, trust, and inclusivity drive impact more reliably than any specific digital service.

4. Governance determines sustainability: Long-term operation depends on stable governance arrangements and clear ownership models. Without these, even well-designed platforms risk becoming short-lived project outputs.

5. AI is an opportunity multiplier, not a structural pillar: AI expands analytical capacity and accessibility but does not reduce the need for human oversight, shared standards, or ethical governance. It should be treated as an enhancement layer that becomes valuable after foundational issues are resolved.

6. Lessons from ESFRI highlight the importance of early formalisation: RIECS can avoid common pitfalls by establishing legal frameworks, metadata standards, and access rules early in its lifecycle. Participatory governance and flexibility remain the differentiating opportunity.

8. Gaps and future needs

Even as solutions to the above challenges are being developed, certain gaps remain unresolved in the current landscape of CS infrastructure. These gaps point to areas where additional research, tooling, or policy intervention is needed to achieve a truly robust and inclusive system. Key unresolved issues include data interoperability shortcomings, inclusivity and equity gaps, long-term sustainability questions, and emerging AI governance needs. Each of these is discussed below, along with their implications for the future of CS in Europe.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

8.1 Data interoperability and standards

Despite progress on common standards, full data interoperability across CS domains remains a work in progress. Many projects still use inconsistent data schemas and metadata, and there is no universal exchange format adopted by all CS platforms. This means that a lot of citizen-generated data cannot easily be aggregated or used outside its original context – a missed opportunity for cross-domain research and policy use. The lack of unified standards also obscure the integration of CS data into authoritative databases. For example, while some environmental observatories feed data into systems like GEOSS, others remain siloed. Policy experts note that standardization efforts, though underway, need dedicated acceleration to truly harmonize technological tools, data quality criteria, and data sharing protocols in CS [10]. In practice, this could mean establishing a formal working group or consortium to define core data models for CS (building on efforts like OGC's SensorThings API or Darwin Core) and to promote their adoption.

Another interoperability gap is the absence of a central discovery mechanism – researchers and policy-makers can struggle to find what CS data even exist on a given topic. Without better indexing, valuable datasets remain underutilized. Closing this gap will likely require both technical tools (e.g., APIs, metadata registries) and policy incentives (requirements or encouragement for projects to publish data in open, standard formats). Until these measures take root, data fragmentation will persist, limiting the evidence base that CS can offer to science and policy. This gap in interoperability has broad implications: it affects scientific knowledge integration, complicates data-driven decision making (since datasets cannot be easily combined), and can lead to duplication of efforts. Tackling it is paramount to maximizing the return on investment in CS projects.

8.2 Inclusivity and participation

Ensuring that the benefits and participation opportunities of CS are inclusive for all communities is an unfinished task. While many projects are conscious of inclusivity, gaps remain in reaching diverse populations and lowering barriers to entry. Certain groups – for instance, people from minority ethnic backgrounds, rural communities, or those with lower internet access and digital skills – are often underrepresented in CS activities. Additionally, gender imbalances and the lack of accessibility for those with disabilities can occur if not explicitly addressed. Existing solutions (like multi-language platforms, smartphone apps for those without computers, or outreach programs in schools) are steps in the right direction but need scaling up. A related gap is the digital and scientific literacy divide. As projects incorporate more advanced tools (e.g., data analysis platforms or AI assistants), there is a risk that only highly educated or tech-savvy individuals can fully participate, leaving others behind. The REINFORCE roadmap highlighted the need to improve citizens' AI literacy and IT skills as a prerequisite for broader inclusion in cutting-edge CS [10]. Recent efforts in this direction include the creation of the European Citizen Science Academy [36], which provides courses and training programmes designed to strengthen skills across the community.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

There is no widespread curriculum or certification to train *citizen data scientists*, although some initiatives (like MOOCs and workshops) have started. The implication of these inclusivity gaps is that CS could inadvertently deepen knowledge inequalities if not corrected – those already empowered and educated gain even more opportunities, while marginalized groups see less benefit. To address this, further actions are required: e.g., dedicated funding for community organizers in underserved areas, co-created projects that align with the needs of marginalized communities, improved usability and accessibility standards in CS platforms (for example, ensuring tools work on low-end devices and follow accessibility guidelines). Inclusivity is not just a moral imperative but also practical: a more diverse participant base means more diverse data and perspectives, which can improve the science itself. Closing this gap will likely need policy support (for example, calls in Horizon Europe that require an inclusivity plan for CS projects) and sharing of best practices internationally. This is an area where social innovation is as important as technical innovation.

8.3 Long-Term sustainability and governance

Technical challenges in RIECS cannot be separated from organisational conditions. Decisions about architecture, standards, scalability, and AI integration are shaped by governance arrangements, funding pathways, and institutional commitments. Deliverable **2.1** focuses on the technical dimensions of the future CS infrastructure, while Deliverable **3.1** will analyse the organisational, institutional, and governance aspects in depth. Even so, this deliverable must highlight the main points where technical and organisational domains intersect, since many design choices depend on institutional responsibilities that are still undecided.

One of the most significant unresolved questions is who will maintain and govern the CS infrastructure in the long run, and how. RIECS is producing a concept and an implementation plan, but beyond that horizon lies the challenge of sustaining a permanent and shared infrastructure. This challenge has multiple facets: financial sustainability, organizational governance, and legal frameworks. Financially, most CS platforms today rely on short-term project grants or volunteer effort. There is a gap in stable funding models – unlike traditional RIs (telescopes, laboratories) which might receive institutional or government funding, CS infrastructures are not yet institutionalized in the same way. WeObserve and others have recommended exploring business models and funding schemes (such as integrating CS infrastructure into national research budgets, or developing services that could generate revenue) but concrete mechanisms remain undeveloped. In terms of governance, it's not yet decided what entity (or entities) will run a Europe-wide CS platform. Will it be a consortium of universities and NGOs? A new European legal entity or perhaps an ERIC? How will different countries and stakeholder groups be represented in decisions? These questions point to a gap in the governance model that needs resolution for implementation. Coordination and support from policymakers is still required – as noted in REINFORCE's findings, *fostering a supportive ecosystem for CS is a key task and challenge for policymakers [10]*, especially since CS projects have funding and operational needs distinct from conventional research.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Without clear governance, there is also a risk of fragmentation re-emerging: multiple overlapping platforms could vie for primacy, or some countries might invest in their own infrastructures while others lag behind. A unified governance approach would mitigate this, ensuring the infrastructure remains open, free of charge, and accessible as a public resource (as one policy objective suggested) [10]. Additionally, long-term governance must address legal and ethical policies: e.g. ensuring compliance with data protection laws (GDPR), clarifying data ownership rights (do citizen contributors retain any rights to their data or its products?), and establishing policies for handling misuse of the platform. While RIECS and similar projects are laying groundwork in these areas, the actual enactment of governance structures is still pending. The implication of not resolving this gap is severe: without a sustainable governance plan, the infrastructure could fail to thrive or even collapse after initial project funding ends. Therefore, this gap calls for attention from high-level decision-makers. Inclusion of CS infrastructure in strategic research agendas (like ESFRI roadmaps) and committing national co-funding could be part of the solution. In summary, the vision of a lasting CS infrastructure will only be realized if we create robust institutions or consortia to steward it beyond the prototype stage, and that remains an open task.

8.4 AI Governance and ethics

As artificial intelligence becomes more entwined with CS, a new category of gaps revolves around AI governance and ethics. While we see pilot uses of AI, there is not yet a consensus or framework specifically guiding how AI should be deployed in CS contexts. Key questions include: How do we ensure AI tools used on citizen-sourced data are transparent, fair, and accountable? What measures confirm that AI is not inadvertently introducing bias or misleading participants? And how can citizen scientists have a say in the use of AI that might affect their contributions? Currently, these questions are only partially answered. Some projects have developed internal ethical guidelines, but a broader governance approach (potentially an extension of existing AI ethics frameworks to CS) is lacking. Experts have underscored that critical issues like data privacy, algorithmic bias, and responsibility for AI decisions *must be addressed* to make sure AI augments rather than diminishes CS [37]. For instance, if an AI model flags a volunteer's data as low-quality, there should be a transparent rationale and an opportunity for review; otherwise users may lose trust without understanding the AI's decision. Inclusion is another ethical aspect: there's a gap in ensuring AI tools cater to different languages and local contexts so as not to exclude anyone – something that general AI products often fail at. Additionally, while the EU is moving forward with the AI Act to regulate AI broadly, it's unclear how CS tools will be categorized or regulated under it.

The CS community may need to proactively define standards for *CS AI* that align with the principles of open science and public participation. The cost and resource aspects are also pertinent: advanced AI could be out of reach for community groups without significant funding or technical expertise, which could centralize power with those who do have resources (a governance concern). Addressing AI governance in CS likely requires multi-stakeholder dialogue – involving AI experts, CS practitioners,

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

ethicists, and the citizen scientists themselves. This could result in guidelines or a code of conduct for AI in CS, and in best-case scenario, toolkits to help projects implement ethical AI (e.g. bias checking tools, consent management systems for data used in training AI). Until such measures are established, the gap persists and carries a risk: improper use of AI could erode volunteer trust or even cause harm (for example, if an AI misidentifies something critical like a health-related observation). In conclusion, integrating AI responsibly is an area needing further research and clear policy action. By preemptively creating governance mechanisms now – during the infrastructure's conceptual phase – the community can ensure that as AI usage grows, it does so under guidelines that protect and empower participants. This will help avoid future pitfalls and reinforce the credibility of CS in the age of AI.

References

[1] A. Bowser *et al.*, "Still in Need of Norms: The State of the Data in Citizen Science," *Citizen Science: Theory and Practice*, vol. 5, no. 1, Art. no. 1, Sept. 2020, doi: 10.5334/cstp.303.

[2] D. Fraisl *et al.*, "Mapping citizen science contributions to the UN sustainable development goals," *Sustain Sci*, vol. 15, no. 6, Art. no. 6, Nov. 2020, doi: 10.1007/s11625-020-00833-7.

[3] *Regulation (EU) No 1291/2013*, vol. 347. 2013. Accessed: Dec. 23, 2025. [Online]. Available: http://data.europa.eu/eli/reg/2013/1291/oj

[4] ESFRI, "European Strategy Forum on Research Infrastructures ESFRI LANDSCAPE ANALYSIS 2024." Accessed: Apr. 24, 2025. [Online]. Available: https://landscape2024.esfri.eu/media/coqdoq0q/20240604_la2024.pdf

[5] B. Latour, *Science in Action: How to Follow Scientists and Engineers Through Society*. 1987. [Online]. Available: https://books.google.com.co/books?hl=es&lr=&id=sC4bk4DZXTQC&oi=fnd&pg=PA19&dq=bruno+latour&ots=WamFzsedTz&sig=njWltI3QSxsoGs1Y1bBtNrFXXpc&redir_esc=y#v=onepage&q=bruno latour&f=false

[6] S. Star, "The Ethnography of Infrastructure," *American Behavioral Scientist*, vol. 43, no. 3, pp. 377–391, Nov. 1999, doi: 10.1177/00027649921955326.

[7] K. Soacha-Godoy *et al.*, "Research Infrastructures in Citizen Science: State of Knowledge and Taxonomic Framework As a Pathway to Sustainability | Citizen Science: Theory and Practice," Jan. 2025, doi: 10.5334/cstp.831.

[8] UNESCO, "UNESCO Recommendation on Open Science." UNESCO, 2021. Accessed: Sept. 18, 2020. [Online]. Available: https://en.unesco.org/science-sustainable-future/open-science/recommendation

[9] K. Soacha-Godoy, A. López-Borrull, F. Serrano, and J. Piera, "The Backbone of Participatory Science: Reframing Citizen Observatories as Research Infrastructures," *Sustainability*, vol. 17, no. 10, Art. no. 10, Jan. 2025, doi: 10.3390/su17104608.

[10] REINFORCE, "What are the Policy Challenges for the Development of Research Infrastructures for Citizen Science in Europe?" Accessed: Dec. 23, 2025. [Online]. Available: https://roadmap.reinforceeu.eu/

[11] Francesco Mureddu (LC), Alessandro Paciaroni (LC), Nathan Carvalho (LC), "D10.3 Policy Roadmap on Research infrastructures for citizens science in Europe." REsearch INfrastructures FOR Citizens in Europe (REINFORCE).

[12] "Very Large Research Infrastructures: Policy issues and options," OECD Science, Technology and Industry Policy Papers 153, July 2023. doi: 10.1787/2b93187f-en.

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

[13] M. D. Wilkinson *et al.*, "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data*, vol. 3, no. 1, Art. no. 1, Mar. 2016, doi: 10.1038/sdata.2016.18.
[14] S. Carroll *et al.*, "The CARE Principles for Indigenous Data Governance," *Data Science Journal*, vol. 19, no. XX, Art. no. XX, 2020, doi: https://doi.org/10.5334/dsj-2020-042.
[15] C. Ruiz and D. Joaquin, "Platform Design : Creating Meaningful Toolboxes When People Meet," 2018, Accessed: Nov. 24, 2025. [Online]. Available: https://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-7425
[16] H.-Y. Liu, D. Dörler, F. Heigl, and S. Grossberndt, "Citizen Science Platforms," in *The Science of Citizen Science*, K. Vohland, A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht, Eds., Cham: Springer International Publishing, 2021, pp. 439–459. doi: 10.1007/978-3-030-58278-4_22.
[17] M. Chandler *et al.*, "Contribution of citizen science towards international biodiversity monitoring," *Biological Conservation*, 2016, doi: 10.1016/j.biocon.2016.09.004.
[18] J. Waller, "Citizen Science on GBIF - 2019." Accessed: June 11, 2021. [Online]. Available: https://data-blog.gbif.org/post/citizen-science-on-gbif-2019/
[19] "CitizenScience," Zooniverse. Accessed: Aug. 30, 2022. [Online]. Available: https://blog.zooniverse.org/category/citizenscience/
[20] EMBIMOS research group, *MINKA citizen observatory*. (July 08, 2024). Institut de Ciències del Mar (ICM-CSIC), Barcelona, Spain. Accessed: July 08, 2024. [Online]. Available: https://minka-sdg.org/home
[21] W. consortium, "Roadmap for the uptake of the Citizen Observatories' knowledge base," Mar. 2021, Accessed: July 14, 2025. [Online]. Available: https://zenodo.org/records/4646774
[22] G. Hager *et al.*, "Onto new horizons: insights from the WeObserve project to strengthen the awareness, acceptability and sustainability of Citizen Observatories in Europe," *JCOM*, vol. 20, no. 6, Art. no. 6, Oct. 2021, doi: 10.22323/2.20060201.
[23] J. Masó and S. Fritz, "EuroGEO 2019 Citizen Science roadmap - 'Lisbon Declaration,'" Zenodo, July 2020. doi: 10.5281/zenodo.3946506.
[24] J. Masó and U. Wehn, "A Roadmap for Citizen Science in GEO - The essence of the Lisbon Declaration. WeObserve policy brief 1.," Aug. 2020, Accessed: July 14, 2025. [Online]. Available: https://zenodo.org/records/4001683
[25] S. Martinez de la Riva, M. Giuffrida, M. Willems, and Á. Justamante Rodríguez, "Supporting knowledge creation and sharing by building a standardised interconnected repository of biodiversity data," May 2022, doi: 10.5281/zenodo.6516724.
[26] M. P. López, R. G. Ribera, and S. M. de la Riva, "Deliverable 4.3 Experts portal for environmental data validation," Nov. 2022, doi: 10.5281/zenodo.7708596.
[27] "About DHIS2," DHIS2. Accessed: Oct. 30, 2025. [Online]. Available: https://dhis2.org/about-2/
[28] M. Krafft, "Supporting effective health data use in the Democratic Republic of the Congo with DHIS2," DHIS2. Accessed: Dec. 23, 2025. [Online]. Available: https://dhis2.org/drc-data-use/
[29] M. Krafft, "Now is the time to invest in locally-owned health data systems," DHIS2. Accessed: Dec. 23, 2025. [Online]. Available: https://dhis2.org/invest-in-routine-systems/
[30] "HISP History," DHIS2. Accessed: Dec. 23, 2025. [Online]. Available: https://dhis2.org/hisp-history/
[31] "About DHIS2," DHIS2. Accessed: Dec. 23, 2025. [Online]. Available: https://dhis2.org/about-2/
[32] "ICOS data flow | ICOS." Accessed: Dec. 23, 2025. [Online]. Available: https://www.icos-cp.eu/data-services/data-collection/data-flow
[33] "Organisation and governance | ICOS." Accessed: Dec. 23, 2025. [Online]. Available: https://www.icos-cp.eu/about/organisation-governance

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

[34] "Community services | ICOS." Accessed: Dec. 23, 2025. [Online]. Available: https://www.icos-cp.eu/data-services/community
[35] "Structure of ICOS | ICOS." Accessed: Dec. 23, 2025. [Online]. Available: https://www.icos-cp.eu/about/organisation-governance/structure
[36] "Home | ECS Academy." Accessed: Dec. 23, 2025. [Online]. Available: https://moodle.citizenscience.eu/
[37] "Inclusive and accessible science: citizen science in the age of AI." Accessed: Dec. 23, 2025. [Online]. Available: https://citizenscience.eu/blog/2025/08/11/inclusive-and-accessible-science-citizen-science-in-the-age-of-ai/

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Annex

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Matrix A: Inputs of challenges, platforms, services and tools

The following list of challenges, functionalities, and solutions is a *workshop ready synthesis*. These elements serve as inputs for the upcoming validation and prioritisation process, during which stakeholders will review and rank the identified challenges, solutions, gaps, functionalities, services, and tools. This collaborative effort will guide the prioritisation of requirements and inform the next stages of RIECS development.

Challenge	**Functionalities, services and tools**
Diversity of formats and lack of data standards	Integrated portal for projects, data, and resources
Limited integration between existing platforms and services	FAIR data repository (Findable, Accessible, Interoperable, Reusable) with consistent metadata
Insufficient scalability for large data volumes	Knowledge repository in multiple formats (protocols, videos, guides, conceptual maps)
Irregular connectivity in many areas	Library of reusable components (code, templates, workflows)
Variability in data quality and validation	Open-hardware repository for sensors and kits
Difficult long-term technical support and maintenance	Open APIs and interoperability services between platforms
Complex management of licenses and data rights	Validation tools combining automated criteria and expert review
Fragmented governance among actors	Computing capacity for processing big data, including images and time series

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Uneven participation and accessibility barriers	Infrastructure to train and evaluate AI models safely and under supervision
Use of AI with risks of bias and lack of transparency	Common authentication system with identity and permission management
	Collaboration environments and support spaces for communities, teams, and initiatives
	Dashboards and indicators for analysis and communication of results

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Matrix B: Needs, challenges, solutions from EUfunded projects

This matrix consolidates the needs, challenges, and solutions identified across the EU-funded projects and will be used in the workshop with technology providers involved in these developments. It supports a shared understanding of how challenges have evolved and what requirements emerge for RIECS moving forward: Matrix B

This project has received funding from the European Union's research and innovation programme Horizon Europe under the grant agreement No. 101188210. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.