Engaging Public through Citizen Science

V. Tokareva for TA7 “Education, training, outreach, and citizen science”

PUNCH4NFDI Annual Meeting 2025
19 November 2025

What is Citizen Science?

Definition. Citizen Science is scientific research conducted, in whole or in part, by members of the general public — often in collaboration with professional scientists.

Key ideas:

- Non-professionals take part in real scientific processes: data collection, analysis, or computation.
- Contributions are structured, validated, and integrated into research workflows.
- It's part of the broader Open Science movement — emphasizing transparency and participation.

In the framework on Deliverable-1 we studied popular citizen science platforms, projects, initiatives, associations in physics and beyond.

What Citizen Science is Today

Diverse, global field:

- Environmental monitoring (biodiversity, pollution)
- Health and social data (community studies)
- Astronomy and physics (Einstein@Home, CREDO, Galaxy Zoo)
- Machine learning and AI challenges (Kaggle competitions)
- Education and science communication

Trends:

- Increasing use of online platforms and gamified interfaces
- Institutional recognition: EU Citizen Science platform, UNESCO guidelines
- Still dominated by biology and ecology. Physics is underrepresented

What Citizen Science is **not** and how it does **not** work

Myth:	Reality:
"If we publish our idea, people will join."	People will join only if you invest in outreach, communication, and feedback.
"Citizen Science is easy — just give people some data to play with and read them a lecture."	Citizen Science is real science — it must yield valid, publishable research results
"Every bachelor student can come up with a project idea and run it 10 hours per week."	Successful projects need diverse, experienced teams and sustained effort.
"Citizen Science is self-organizing."	Without coordination and project management, most projects quickly fade
"It's cheap labor for data collection."	Participants expect purpose, transparency, and recognition — not exploitation.

Common Criticisms of Citizen Science

Scientific Concerns

- Data quality & validation — participant-collected or annotated data often needs extensive cleaning and verification.
- Lack of reproducibility — inconsistent methods or tools can make results hard to compare.
- Limited scientific significance — projects sometimes focus on engagement rather than meaningful research outcomes.

Social and Organizational Concerns

- Lack of diversity — participants are mostly well-educated men from developed countries; inclusivity goals remain largely unmet.
- Short-lived engagement — most volunteers lose interest quickly if feedback is missing.
- Unequal power dynamics — participants often feel “used” rather than genuinely involved in research decisions.
- Insufficient project management — many initiatives fail due to the absence of coordination and communication structure.

How to join an existing Citizen Science Project

Where to look

World	Citizen Science Association (CSA)
Europe	European Citizen Science Association (ECSA)
	Research Infrastructures FOR Citizens in Europe (REINFORCE)
DACH	Citizen Science DACH AG
Germany	mit:forschen! Gemeinsam Wissen schaffen
	Wissenschaft im Dialog (WiD)
Platforms	BOINC, CitSci.Org, SciStarter, Zooniverse, etc.
Projects	Einstein@Home, CREDO, Gravity Spy, etc.

Other national level organisations: Citizen Science Network Austria (CSNA), Australian Citizen Science Association (ACSA), etc.

WP7.4 platforms overview (2024)

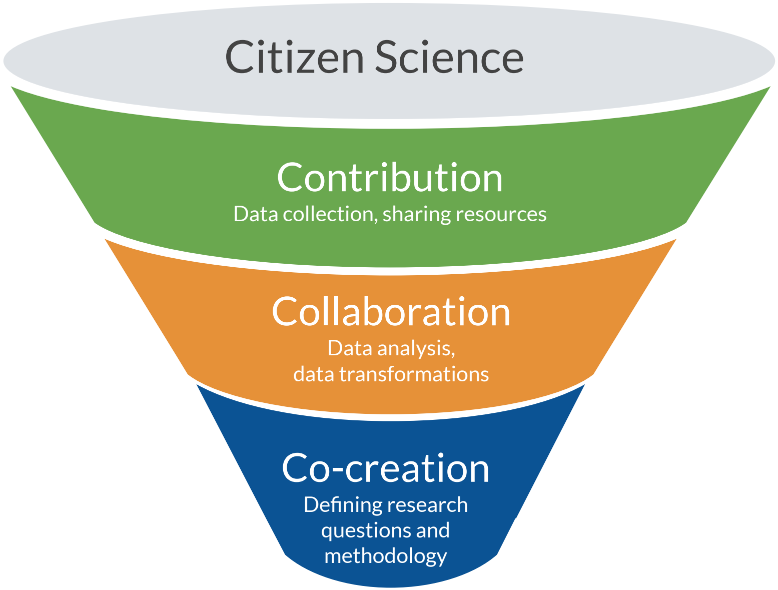
- 11 platforms/projects: Anecdata, CitSci.Org, FieldScope, GLOBE Observer, iNaturalist, Zooniverse, SciStarter, BOINC, Bürger schaffen Wissen, EU-Citizen.Science, Kaggle^
- Platforms with presence of CitSci projects in physics (8):

Platform Name	Anecdata	CitSci.Org	EU-Citizen. Science	Bürger schaffen Wissen	Zooniverse	BOINC-based	SciStarter	Kaggle
NumAllProj	48	1142	321	168	97	29	1393	165 ("Research")
NumActPhysSpaceProj	2	2	16	2	25	6	101	12
NumActPUNCH-rel	0	0	5	2	23	6	≈65	7

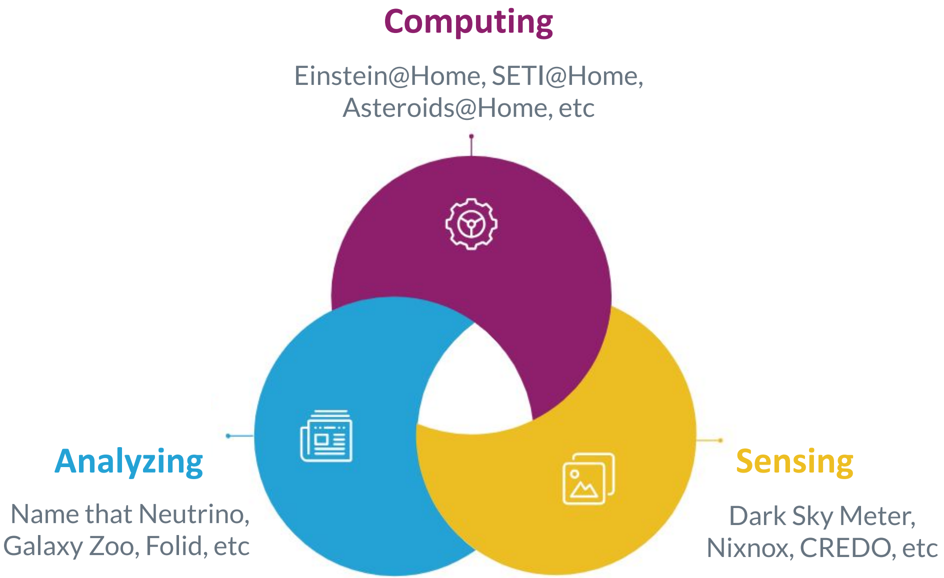
^ Kaggle is AI and ML platform. It's main goal is not CitSci

Citizen Science activities

Types of engagement[1]



Activities classification[2]



[1] Levesque V. Citizen Science– Sustainability Methods & Perspectives. Pressbooks. Retrieved 2025 Nov 11, from https://pressbooks.pub/sustainabilitymethods/chapter/citizen-science/

[2] Strasser BJ et al. (2023) Quantifying online citizen science: Dynamics and demographics of public participation in science. PLoS ONE 18(11): e0293289. https://doi.org/10.1371/journal.pone.0293289

Who is the audience of Citizen Science?

Computing

- Monthly audience, Jun 2021: 71k
- 10% contributing 71%
- > 60% are in the field of science or IT
- SETI - the biggest project: >14mio people total
- Demographics (SETI):
  - 90% male
  - med age 34
- BOINC: 15 PFLOPS/24h -> "comparable to fastest supercomputers" (Horeka: 17 PFLOPS - 15th in Europe)

Analyzing

- Monthly audience, Jun 2021: 23k
- 10% contributing 79%
- Gender: 68% male (Galaxy Zoo)
- Age :
  - projects, such as Galaxy Zoo - many school-age participants (35%)
  - Gamified projects (FoldIt): 65% of the participants are < 30 y.o. (median age =25)
- Occupation (FoldIt): 80% in science or IT
- Background (FoldIt): 90% in science or IT

Sensing

- Monthly audience, Jun 2021: 665k
- 10% contributing 69%
- Gender:
  - iNaturalist: 56% male
  - eBird: 51% male
- Highest participation (iNaturalist): Europe, USA and countries with rich biodiversity - Costa Rica, Panama, Taiwan, etc.
- Only biological projects

[2] Strasser BJ, Tancoigne E, Baudry J, Piguet S, Spiers H, Luis-Fernandez Marquez J, et al. (2023) Quantifying online citizen science: Dynamics and demographics of public participation in science. PLoS ONE 18(11): e0293289. https://doi.org/10.1371/journal.pone.0293289

[3] Füchslin, Tobias; Schäfer, Mike S; Metag, Julia (2019). Who wants to be a citizen scientist? Identifying the potential of citizen science and target segments in Switzerland. Public Understanding of Science, 28(6):652-668. DOI: https://doi.org/10.1177/0963662519852020

How You Can Engage

- Participate as a volunteer — experience the user perspective firsthand.
- Explore and analyze existing Citizen Science datasets for your own research or teaching.
- Mentor student projects using Citizen Science data.
- Advise or collaborate with ongoing projects — contribute domain expertise and feedback.
- Provide data or tools — share datasets, validation pipelines, or visualization frameworks.
- Co-develop new methodologies — work on FAIR data, reproducibility, or engagement research.

Einstein@Home

- Long-running, well-structured citizen science project that uses distributed volunteer computing to search for pulsars and binary neutron stars.
- Built on BOINC, a mature and secure distributed computing platform.
- Uses a stable, scalable data processing workflow.
- Has a large, motivated volunteer community already in place.
- Produces scientifically significant results (e.g., newly discovered pulsars).

Lessons from Einstein@Home: what we asked

- What scientific challenges did you face, and how did you address them?
- How do you maintain motivation and long-term engagement among volunteers?
- What organizational challenges did you encounter?
- Which common misconceptions have you encountered?
- Which initial expectations turned out differently in practice?

Acknowledgement: big thanks to Baida Achkar for help with designing these questions!

Einstein@Home Q&A - 1

Question	**Answer (condensed)**
What **scientific challenges** did you face, and how did you address them?	Citizen Science is diverse; in our case, the challenge is preparing data so it can be distributed as many small, independent work units.
How to **recruit participants**?	We benefited from joining an existing motivated volunteer community rather than trying to build one from scratch.
What **organizational challenges** did you encounter?	We avoided most of them by building on BOINC, an established distributed computing infrastructure.

Acknowledgement: thanks a lot to Michael Kramer and Ramesh Karuppusamy for sharing their experience with us

Einstein@Home Q&A - 2

Question	**Answer (condensed)**
How to **maintain engagement**?	Leaderboards, digital badges, and public recognition significantly help sustain participation
Which **common misconceptions** have you encountered?	For example, some volunteers might suspect the citizen science compute software is some kind of virus or spyware. This can be mitigated by making the software stack public as in BOINC. Transparency and open-source is essential
Did the project meet **expectations**?	Yes, largely because we joined a stable, long-running system rather than starting from zero.

Acknowledgement: thanks a lot to Michael Kramer and Ramesh Karuppusamy for sharing their experience with us

How to Start Your Own Citizen Science Project

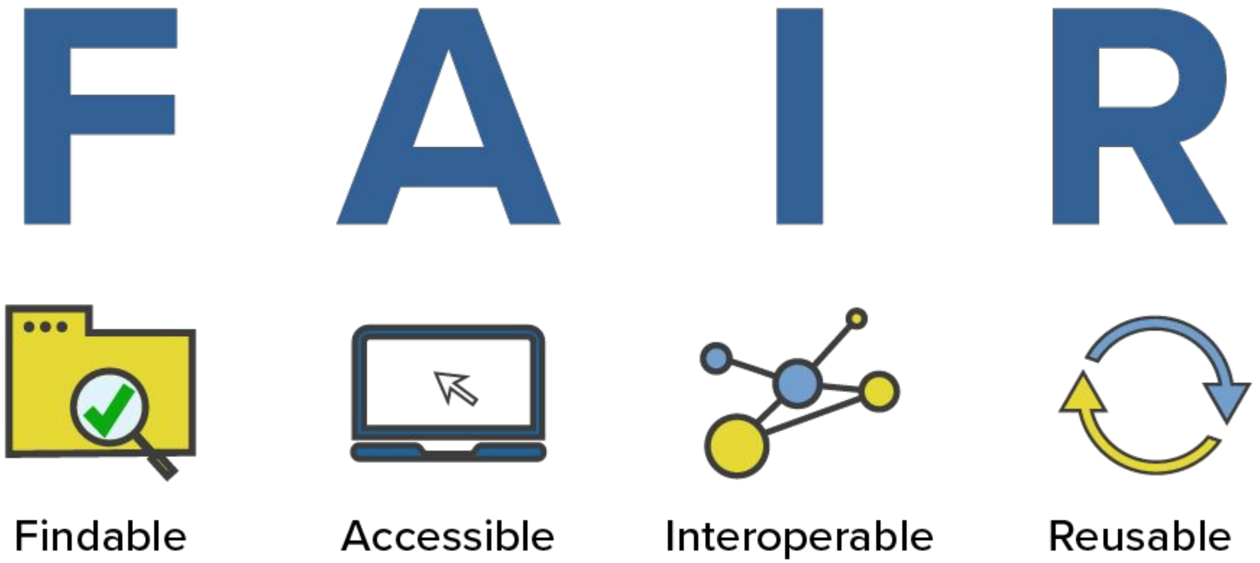
Must-haves

- A good research question — clear, meaningful, and suitable for public participation
- A strong, diverse team — scientists, communicators, developers, educators
- Sustainable resources — time, funding, infrastructure, and long-term planning

Design steps:

1. Define a Clear and Feasible Research Question
2. Build a Strong, Diverse Team
3. Secure Funding and Infrastructure
4. Design the Project Workflow
5. Test, Iterate, and Launch
6. After Launch: Keep Participants Engaged and Data Reliable
7. During and at the end of the project: **FAIR data publication** (open) of the outcomes

FAIR data management



[4] *FAIR Framework Wilkinson M. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3:160018.*

Making your data...

Findable:

- Publish datasets, code, workflows, and documentation in a stable repository (e.g., Zenodo, OSF, institutional storage, GitHub + archival)
- Assign a GUID (Globally Unique Identifier) such as DOI, Handle, ARK, or w3id to ensure persistent referencing.

Accessible:

- Use open data formats (CSV, FITS, HDF5, NetCDF) instead of proprietary ones
- Clearly specify a license (e.g., CC-BY, CC0, MIT)
- Provide comprehensive documentation

Making your data...

Interoperable:

- Use established metadata schemas (DataCite, Dublin Core) and domain standards (e.g., IVOA / VO in astronomy)
- Define input/output formats and units explicitly
- Where possible, rely on controlled vocabularies rather than free text

Reusable:

- Document data quality checks and preprocessing pipelines in a reproducible way
- Provide versioned datasets (v1.0, v1.1, ...)
- Include scientific context and limitations, so others understand how to use — and not misuse the data.

Take home

- It’s **real research**
- Many projects fail from lack of planning, not lack of ideas
- Start small: join, observe, collect their experience
- Success needs a clear question, strong team, and solid funding
- Good DMP and FAIR publishing of the results are great practices to adapt
- Physics has unique challenges — and unique opportunities we should use