Common Infrastructure for National Cohorts in Europe, Canada, and Africa
- CINECA -

Deliverable D6.3 Final Report and Impact Analysis

Work Package:	WP6 – Outreach, training and dissemination
Lead Beneficiary:	European Molecular Biology Laboratory
WP Leader(s):	Kim Gurwitz (EMBL-EBI) Saskia Hiltemann (EMC)
Contributing Partner(s):	UCT, EMC, EMBL-EBI
Contractual Delivery Date:	April 31st 2023
Actual Delivery Date:	April 28th 2023
Authors of this Deliverable:	Nicola Mulder, Mamana Mbiyavanga, Saskia Hiltemann, Kim Gurwitz, Marta Lloret Llinares, Daniel Thomas Lopez
Contributors:	
Reviewed by:	Cath Brooksbank
Approved by:	Thomas Keane
Dissemination Level:	Public
Type of Deliverable:	Report
Grant agreement:	No. 825775 Horizon 2020 (H2020-SC1-BHC-2018-2020)
Type of action:	RIA
Start Date:	1 Jan 2019
Duration:	54 months

Table of contents:

1. Executive Summary 3
2. Project objectives 3
3. Detailed report on the deliverable 3
3.1 Background 3
3.2 Description of Work and Impact 4
3.2.1. Staff Visits 4
3.2.2. Webinar Series 5
3.2.3. Workshops and training events 7
3.2.4. Learning Pathway 14
3.2.5. Short training videos 15
3.2.6. Short technical reports 18
3.2.7. Dissemination activities 21
Website 23
Social media 26
3.2.8. Publications 27
3.3 Conclusions and Sustainability 27
4. Abbreviations 28

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

1. Executive Summary

In this deliverable report, we highlight outreach and training achievements in the period of January 2022 to February 2023 and also summarise overall achievements for the entire CINECA project. In total, we have delivered 59 learning interventions (including staff visits, webinars, workshops and training events, short training videos, and an overarching self-paced learning pathway) and we have reached a large and diverse audience, through both training and dissemination activities and through our website and social media channels. Below we provide further detail in terms of the diverse activities that we have developed and delivered since the beginning of the project, the wide demographic we have engaged, our reach through dissemination and training activities, and the impact that our training has had in the longer term.

2. Project objectives

This deliverable has contributed to objectives of WP6, which are:
a) To raise awareness of CINECA and the opportunities and challenges of sharing cohort data on a global scale. Build relationships with related projects to facilitate the dissemination of information of mutual interest through all appropriate channels to the broadest potential audience.
b) To identify and address training needs, both within the consortium itself and more broadly among key stakeholder groups.
c) To develop and deliver training (or other learning interventions, e.g. secondments; communities of practice; hackathons) in areas of highest training needs, integrating the outputs of the technical work packages to deliver the project's training goals.
d) To assess the impact of the project's outreach, training and dissemination activities.

3. Detailed report on the deliverable

3.1 Background

The CINECA project has developed new standards and technical solutions to address existing challenges in responsible human data sharing and analysis across continents. The project comprises inter-related work packages addressing different needs. One of the needs within the consortium is for the training of staff within work packages and for knowledge dissemination across work packages.

With the CINECA products in the final stages of development, we have focused on training that targets external audiences in this final stage of the project, which also serves to disseminate the project results. This has been achieved through diverse activities including workshops, webinars, and a self-paced open access learning pathway on end-to-end federated data analysis. While the COVID-19 pandemic has affected the format of our training events, pushing the majority of events to online, we nevertheless have provided a full training agenda, which we outline in more detail below.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

3.2 Description of Work and Impact

As mentioned above, the training programme has several parallel activities, including staff visits, webinars, short videos, short technical reports and training events, which aim to meet identified needs and are tailored to the stakeholders identified in the stakeholder analysis outlined in D6.1 *Outreach and dissemination plan*. An overview of the project activities are summarised below (Table 1) and described more fully in the following subsections.

Table 1: *Overview of learning interventions since the start of the project.*

Activity Type	Number of activities
Staff visits	3
Webinars	18
Training Courses & Workshops	17 (16 past, 1 upcoming)
Short training videos	20
Learning Pathway	1
Total	**59**

3.2.1. Staff Visits

The procedure around staff visits was set up early on in the project, and an application form with guidance text has been made available for CINECA partners to request staff visits[1]. Due to the COVID-19 pandemic, the staff visit programme has not been utilised as extensively as envisioned at the start of the project, and we have focused on other modes of knowledge exchange. However, during this reporting period one staff visit has occurred, between HMGU and UTARTU for a knowledge transfer around the development of Nextflow workflows with the purpose of building a common framework for the Polygenic Risk Score use case pipeline. Another staff exchange has been requested for the near future to explore the feasibility of deploying Secure MultiParty Computation services on top of EGA/FEGA (between HEG Geneva and EMBL-EBI). This brings the total number of staff visits for the project to three, which includes one staff visit that occurred before the start of the pandemic.

In order to assess the effectiveness of our staff visit program, we've sent out impact surveys to the visiting party at 6 months post event (currently this survey has been sent to the first two participants). Both participants indicated they have since implemented the skills they learned during the staff visit in their own research, and both were still collaborating with their host institution. Both participants had also taught the obtained skills to 1-5 others since the time of the staff visit and both

[1] https://www.cineca-project.eu/blog-all/cineca-staff-visit-program

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

indicated they would recommend the staff visit programme to others. In addition to the obtained skills, the networking aspect of the staff visit was seen as immensely valuable by both participants.

3.2.2. Webinar Series

We set up a CINECA webinar series to deliver training and disseminate our activities within CINECA and to a broader audience of stakeholders. The CINECA webinar series aims to discuss common challenges, share best practices regarding cohort data analysis, and distribute CINECA project outputs. All CINECA webinars are open to everyone and include a live audience Q&A session during which attendees can ask questions and make suggestions. All webinars are recorded, made available via the CINECA project YouTube channel[2], and are embedded on the event webpage on the CINECA website[3]. From the beginning of the project in January 2019 until February 2023, we delivered 18 webinars, which reached 566 live attendees and received 1634 views on YouTube. We reported in detail on the webinars delivered until the end of 2021 in previous deliverables D6.2 - *Training programme*[4] (M12), D6.4 - *Training Programme, Detailed*[5] (M24) and D6.5 - *Training Programme, Annual Report 2021*[6] (M36). Here we will provide detail about the 4 webinars delivered between January 2022 and February 2023.

In March 2022, the webinar Bringing it all together: human cohort standards, tools and applications[7] was about how to integrate and harmonise diverse, large human cohort data at scale. It included an introduction to data standards and an explanation on how to integrate and harmonise diverse, large human cohort data at scale. After the summer 2022, we started a webinar series to show the final outputs of the CINECA project, including the project's use cases. This series started in November 2022 with Modular and reproducible workflows for federated molecular QTL analysis[8], which provided an overview of the CINECA workflows for genotype imputation, gene expression and splicing quantification, data normalisation and association testing, and demonstrated how these workflows can be used in a federated setting without transferring identifiable personal data between partners. In January 2023, the series continued with Federated analysis for polygenic risk score calculations[9], an overview of the development of a demonstrator of federated genetic analyses utilising a computational pipeline for polygenic risk score analysis. It demonstrated this pipeline utilising the CINECA UK1 synthetic dataset, derived from the 1000 genomes project, as a demonstrator. In February 2023, we ran a webinar on ELSI topics, The case of data reuse: ethical, legal, and societal issues in international genomic data access and sharing[10], which focused on the

[2] https://www.youtube.com/channel/UCfEWFBY8z-TLJi1ije1oWug
[3] https://www.cineca-project.eu/webinars
[4] https://zenodo.org/record/3952621#.ZBSj8HbP1D8
[5] https://zenodo.org/record/6223125#.ZBSkC3bP1D8
[6] https://zenodo.org/record/5795482#.ZBSji3bP1D8
[7] https://www.cineca-project.eu/news-events-all/bringing-it-all-together-human-cohort-standards-tools-and-applications
[8] https://www.cineca-project.eu/news-events-all/modular-reproducible-workflows-federated-molecular-qtl-analysi
[9] https://www.cineca-project.eu/news-events-all/federated-analysis-polygenic-risk-score
[10] https://www.cineca-project.eu/news-events-all/ethical-legal-societal-issues-internation-genomic-data

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

questions raised by the use of cohort data for research projects beyond the initial purpose. Table 2 summarises the webinars delivered between January 2022 and March 2023.

***Table 2**: Summary of the webinars delivered in the CINECA webinar series in 2022 and 2023*

Date	Title	Speaker
31/03/2022	Bringing it all together: human cohort standards, tools and applications	Melanie Courtot (OICR)
10/11/2022	Modular and reproducible workflows for federated molecular QTL analysis	Kaur Alasoo (University of Tartu)
31/01/2023	Federated analysis for polygenic risk score calculations	Will Rayner, Anshika Chowdhary (Helmholtz Munich)
09/02/2023	The case of data reuse: Ethical, legal, and societal issues in international genomic data access and sharing	Melanie Goisauf (BBMRI-ERIC), Emmanuelle Rial-Sebbag (CERPOP, Inserm/University of Toulouse)

There were between 16 and 62 attendees at each webinar with an average of 55% conversion rate of registrations. Webinars have a feedback form that is automatically shared with the attendees at the end of the webinar and is sent to participants in a follow-up email. Survey responses ranged from 6% to 13%, with a 9.75% average. The survey responses suggest that the pace of the webinars was very good (4.17 on average, with 1 being poor and 5 excellent). The attendees also found the webinar content to be good (3.63 on average). Table 3 illustrates the high-level statistics and feedback collected for the CINECA webinars run between January 2022 and March 2023.

***Table 3:** High-level statistics and feedback collected from the CINECA webinar series between January 2022 and March 2023*

Webinars	Nmb attendees	Conversion	Survey response	Average pace**	Average content **	Youtube views (on 10 March 2023)
Bringing it all together: human cohort standards, tools and applications	41	57.75%	*	*	*	93
Modular and reproducible workflows for federated molecular QTL analysis	16	34.78%	6.25%	4	4	42
Federated analysis for polygenic risk score calculations	62	65.96%	9.68%	4.17	3.83	97

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

The case of data reuse: Ethical, legal, and societal issues in international genomic data access and sharing	60	61.86%	13.33%	4.25	3.63	54
Average statistics	44.75	55.01%	9.75%	4.14	3.82	71.5

*Due to a technical issue, we did not collect feedback for this webinar.
**Range between 1 and 5 for Average Pace and Average Content.

3.2.3. Workshops and training events

Workshops and training events are crucial components of the CINECA project and provide a platform for stakeholders to share knowledge, promote collaboration and exchange of ideas, and build technical and non-technical skills. These events also help to disseminate project goals and outcomes and increase the project's impact and relevance. This report covers the workshop and training events that occurred in 2022 and 2023, while events organised in 2020 and 2021 are discussed in detail in deliverable D6.5 Training Programme, Annual Report 2021.

Between January 2022 and February 2023, we continued our workshop series on Beacons and organised the second annual Galaxy Smörgåsbord training event, a data harmonisation workshop, a federated data analysis workshop, and a workshop on deploying and querying a Beacon, as summarised in Table 4 and detailed in the following paragraphs.

Table 4: *CINECA workshop and training events delivered between January 2022 and February 2023.*

Date	Topic	Organisers/Trainers
Feb 15, 2022	Lighting a Beacon: training for (future) implementers[11]	Lauren Fromont, Roberto Ariosa, Sabela De La Torre
Feb 17, 2022	Beacon v2 Reference Implementation: An Overview[12]	Lauren Fromont, Manuel Rueda
Mar 14-18, 2022	Smörgåsbord 2022: Tapas Edition[13]	Saskia Hiltemann, Helena Rasche, et al.
May 18-19, 2022	Data harmonisation workshop[14]	Nicola Mulder, Mamana Mbiyavanga
Feb 26, 2023	Federated data analysis workshop[15]	Kim Gurwitz, Mamana Mbiyavanga

[11] https://www.cineca-project.eu/news-events-all/lighting-a-beacon-training-for-future-implementers
[12] https://www.cineca-project.eu/news-events-all/beacon-v2-reference-implementation-an-overview
[13] https://gallantries.github.io/video-library/events/smorgasbord2/tapas.html
[14] https://www.cineca-project.eu/news-events-all/data-harmonization-workshop
[15] https://www.cineca-project.eu/news-events-all/federated-data-analysis-workshop

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

Feb 27, 2023	Deploy and query a Beacon workshop at the University of Cape Town	Mauricio Moldes, Mamana Mbiyavanga, Coline Thomas, Sumr Panji, Nicky Mulder

Beacons training: In February 2022, CINECA hosted two online training events, targeted at the current and potential implementers of the GA4GH Beacon, which were attended by seven participants each. These training sessions were open to anyone after the delivery of two internal sessions for members of the consortium in November 2021, which were reported in deliverable D6.5 Training Programme, Annual Report 2021 (M36). The two online training sessions comprised talks and demonstrations. The first online training session was on Lighting a Beacon: training for (future) implementers. It provided the participants with the tools to get started with the implementation autonomously and identify the resource persons to answer their questions during the implementation process. The second online training in this series was on Beacon v2 reference implementation. It aimed to familiarise participants with Beacon queries using the UI and the API.

Galaxy Smörgåsbord training event: The second edition of the Galaxy Smörgåsbord event was held in March of 2022, and retained the same format as the first edition; it was a 5-day, online, self-paced, training workshop around a large set of training videos covering a wide range of topics. New topics in 2022 included SARS-CoV-2 data analysis, machine learning, ecology, climate analysis, plant genomics, microbial analysis and computational chemistry. Since this second edition of the event featured such greatly expanded content, it was organised as a choose-your-own-adventure style workshop, where participants could create their own training programme for the week from the available content (100+ hours of video tutorials in total), based on their own interests and background. We had over 3000 registrations with great gender balance (50.5% women, 47.6% men, 1.8% rather not say) (Figure 1). Over 100 instructors from the Galaxy community were available to provide 24/7 support via Slack during the week. We gathered post-event feedback from 262 participants, 96% evaluated the course as 4 or 5 out of 5 in terms of usefulness, with a score of 5 being the most useful, and 96.6% scored the organisation of the course as a 4 or 5 out of 5. Furthermore, 96.9% of respondents indicated they would recommend the course to others.

All materials from the event remain online and are openly available to anybody for self-study, and are conveniently linked to the main Galaxy training website[16].

[16] https://training.galaxyproject.org/

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

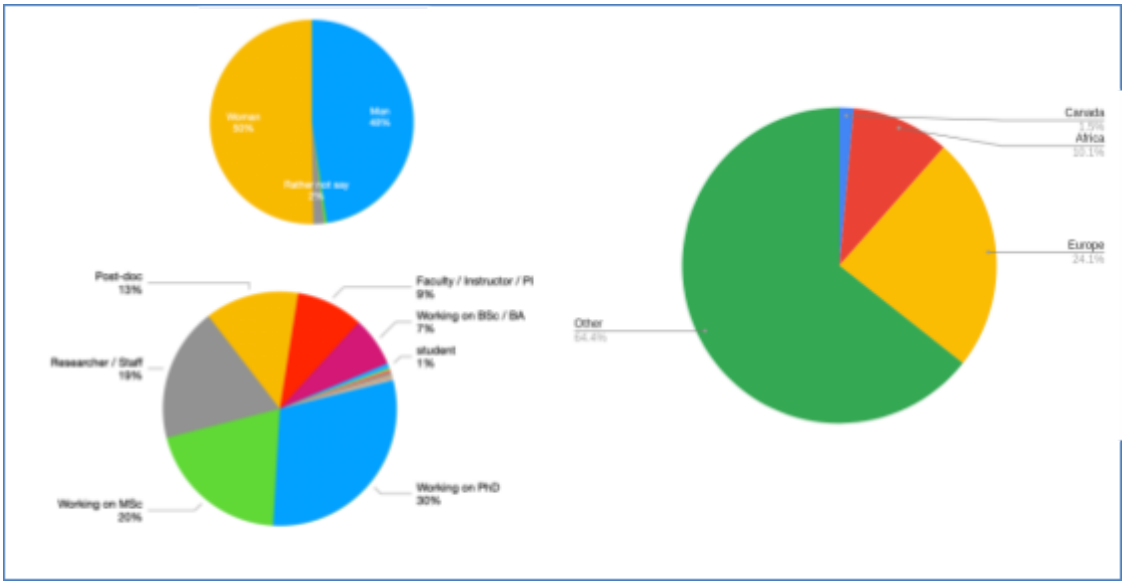

Figure 1: *Demographic data for the online Galaxy workshop showing the gender, career level and top 10 countries of residence of the participants.*

In order to assess the long-term impact of the annual Galaxy training courses, we send out an impact survey 6 months post-event. For this Galaxy course we had 63 responses to our long-term impact survey. This survey revealed that

- 88.9% of respondents indicated that they feel more confident in their data analysis since attending the course.
- 85.2% of respondents continue to apply the skills they learned on a regular basis.
- 90% of respondents had continued their learning journey after the course either by attending other Galaxy courses or using the Galaxy training materials for self-study.
- 70% of the respondents had taught one or more of their colleagues the skills they obtained during the course.

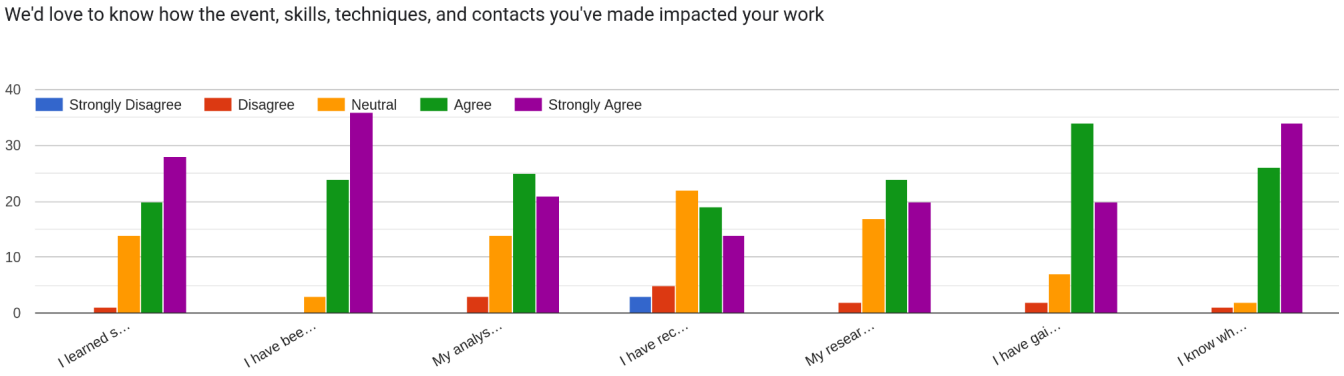

Figure 2: *Responses to the following statements (left to right) 1) I learned skills which I used to advance my career 2) I have been motivated to seek more knowledge about concepts and tools*

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

covered during the workshop 3) My analyses have become more reproducible as a result 4) I have received professional recognition for my work as a result of the tools and skills learned 5) My research productivity has improved as a result of attending 6) I have gained confidence in data analysis as a result 7) I know where I can go to learn more about the topics or skills covered.

Data harmonisation workshop: In collaboration with the H3ABioNet project, CINECA organised an online data harmonisation workshop. The workshop aimed to address common challenges in cohort data harmonisation, share best practices, and work towards practical steps to address these challenges. The workshop took place over two days in May 2022, with a four-hour session on each day. The event featured introductory presentations with example use cases as well as practical or demo sessions. It covered several topics, including clinical and metadata standards, data cleaning, review of existing efforts on cohort data harmonisation, and examples of data harmonisation algorithms. The workshop also included use cases from various cohort projects such as DPUK, PRIMED consortium, H3Africa, DSI-Africa and IHCC. Only 8% of participants were CINECA partners. The overall feedback from the workshop was very positive. Seventy-seven registrations were received from 17 countries, with a good balance of positions, including researchers (39%), data managers (20%), bioinformaticians (17%), data scientists (11%), and others (13%). Gender balance was excellent, with 55% Male and 45% Female (Figure 3). The average attendance across the two days was 52 participants, with more than half having some experience with ontologies and data harmonisation. The post-event feedback survey showed that 85.7% of respondents rated the usefulness and organisation of the event as 4 or 5 out of 5 (very good/excellent), with 100% of respondents indicating they would recommend the course to others and would want to see a follow-up event or engagement. Overall, the workshop was a success and helped to strengthen the collaboration among stakeholders working towards responsible human data sharing and analysis.

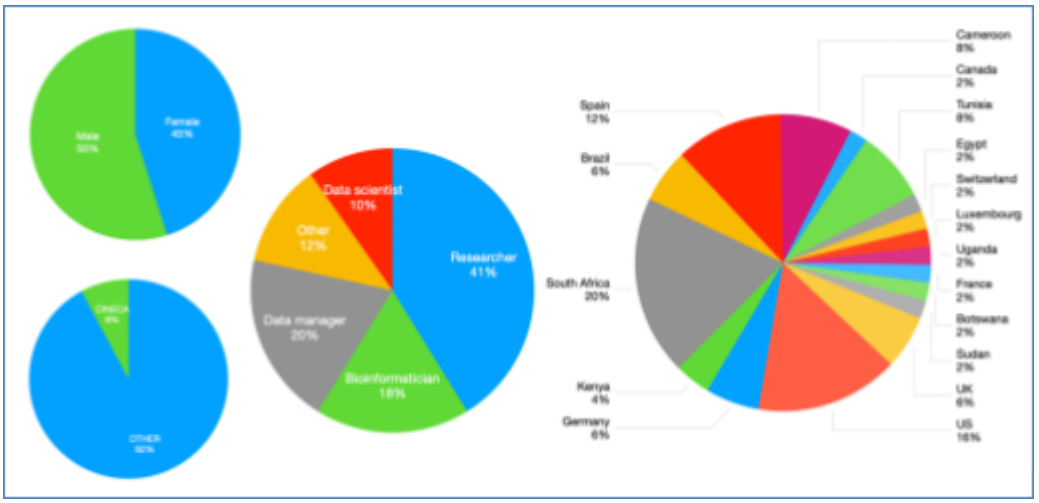

Figure 3: *Demographic data for the online data harmonisation workshop showing the gender, affiliation to CINECA, position and country of residence of the participants.*

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

Federated-data analysis: In the second half of 2022, our primary focus shifted toward the development of the learning pathway[17], a self-paced training pathway following the researcher's journey through an end-to-end federated data analysis. This learning pathway is described in more detail in the next section. This learning pathway was piloted during the face-to-face federated data analysis workshop organised as a satellite meeting of the International conference for human genetics (ICHG) in Cape Town, South Africa in February of 2023. The participants of this workshop were introduced to the learning pathway by instructors from all work packages in a series of lectures, demos and hands-on sessions.

The federated data analysis workshop, organised by CINECA, was designed to give participants a basic understanding of the components involved in federated data analysis. An end-to-end use case was used to showcase all aspects of federated data analysis, including federated data discovery and access, bioinformatics analysis, and ELSI considerations. The face-to-face workshop was structured as a series of short lectures that introduced key concepts for federated data sharing and analysis, such as the GA4GH Beacon API, GA4GH Passports and Visas, and the Data Use Ontology (DUO). The lectures were followed by practical examples and demos of real-world research and clinical applications that highlighted the benefits and challenges of federated data analysis. Out of 40 received applications, a diverse group of 30 applicants was selected from various geographical locations (13 countries from 3 continents), organisations, and positions, including researchers, data analysts, policymakers, and other stakeholders involved in responsible human data sharing and analysis. The attendees were equally balanced in gender, areas of expertise, and prior experience with federated data analysis. This diversity enriched the workshop and provided a platform for collaboration and for attendees to learn from each other's experiences and perspectives (Figure 4). CINECA's first in-person workshop received favourable feedback on organisation and content. The post-event feedback survey indicated that 67% of respondents rated the event as 4 or 5 and 33% as 3 out of 5 (poor/satisfactory/good/very good/excellent) in terms of usefulness and organisation. Attendees recommended more hands-on training to improve the practical experience with federated data analysis tools.

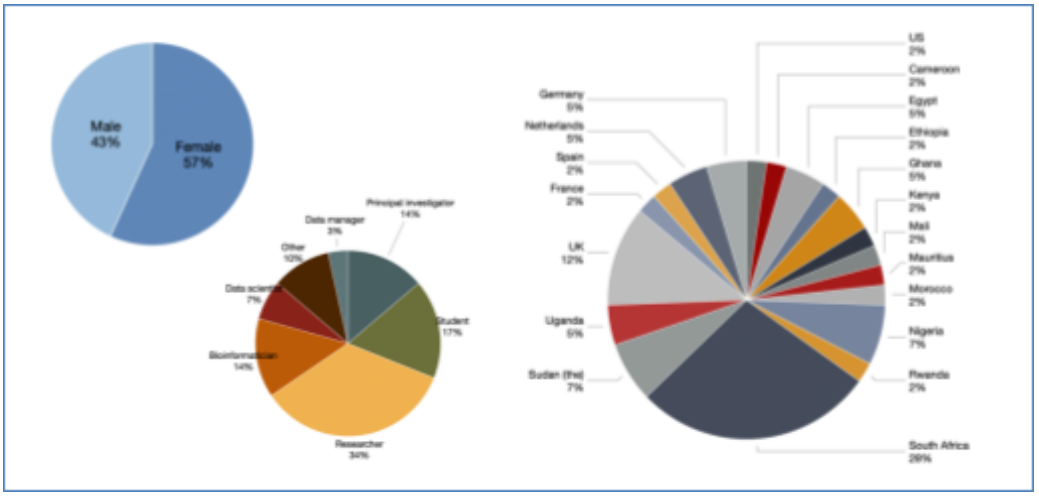

Figure 4: *Demographic data for the face-to-face federated data analysis workshop showing the gender, position and country of residence of the participants.*

[17] https://www.ebi.ac.uk/training/materials/cineca-federated-data-analysis/

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

Deploy and query a Beacon: Finally, on February 27, 2023, a workshop hosted at the University of Cape Town was organised by CINECA and the H3ABioNet project. The workshop aimed to provide training on deploying and querying a Beacon using the EGA reference implementation. The workshop targeted bioinformatics software developers, data managers, and PIs interested in implementing the GA4GH Beacon API within their projects. Ten participants from various bioinformatics groups in Africa attended the workshop, including the South African National Bioinformatics Institute (SANBI), the University of Cape Town Computational Biology group (CBIO), the African Population Cohort Consortium (APCC), the H3Africa TrypanoGEN project, ILIFU, and other research institutions in Africa. The workshop focused on the technical aspects of deploying a Beacon, including tools and techniques developed by the Beacon team at the Centre for Genomic Regulation (CRG) in Spain. The hands-on training and demos on the EGA Beacon tool reference implementation helped participants learn how to make their sensitive data discoverable without jeopardising the privacy of the data. Overall, the workshop was a valuable opportunity for participants to learn about the latest advancements of the GA4GH Beacon API, how to implement it in their own projects, and how to leverage the technology to advance bioinformatics research in Africa further.

Training workshops and events since the beginning of the project:

Since the inception of the CINECA project, sixteen training workshops/events have been organised (as reported on the CINECA news-events site[18]), covering the key challenges that the project addresses. These challenges include: federated data discovery (FDD – 29%), access and authorisation infrastructure (AAI – 19%), harmonised metadata (HM – 14%), federated data analysis (FDA – 24%), and the framework for ethical, legal, and societal issues (ELSI – 14%) (Figure 5). These topics have been covered evenly by the training interventions delivered. For instance, ELSI training was an essential component of the CINECA programme. By always including an ELSI component in our training, CINECA ensured all participants were well-versed in the best practices for ethical data sharing and could promote a culture of responsible and transparent data sharing across borders. Only three events happened in person, while the remaining 13 were conducted virtually, mostly due to the COVID-19 pandemic.

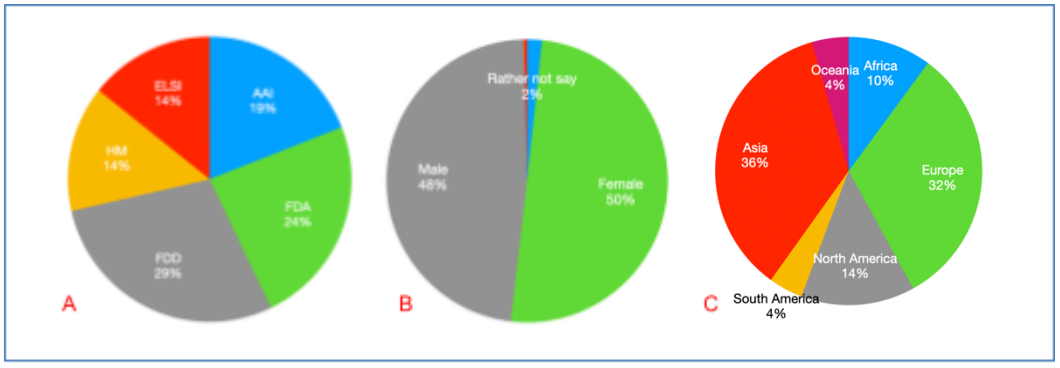

Figure 5: *Overall statistics on workshop and training interventions delivered by CINECA. (A) CINECA challenges covered by CINECA training and workshop interventions. (B) Gender balance achieved in workshop and training events. (C) Continents covered by workshop and training events*

[18] https://www.cineca-project.eu/news-events

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

These events have reached over 5000 participants from around the world, including researchers, policymakers, and data managers. The impact of these workshops and training events has been significant, with participants reporting increased knowledge and skills related to federated data discovery and analysis. For a full list of training events organised by the project, please refer to Table 5.

Table 5: *All CINECA workshop and training events with the location and challenges addressed.*

Year	Topic	Location	CINECA challenge
2020	CINECA AAI Training for Relying Services[19]	Online	AAI
	ELIXIR AAI online advanced training for Relying Services[20]	Online	AAI
2021	GTN Smörgåsbord: A Global Galaxy Course[21]	Online	FDA
	Online training on GA4GH Passports[22]	Online	AAI
	"How FAIR are you" hackathon[23]	Online	FDD, HM, ELSI
	Using Twitter to promote your project - beginner	Online	
	Using Twitter to promote your project - intermediate	Online	
	Lighting a Beacon: training for (future) implementers	Online	FDD
	Browsing genomes: a Beacon training for users	Online	FDD
	Project at ELIXIR BioHackathon[24] (FAIR training with Galaxy)	Barcelona, Spain	FDA
2022	Lighting a Beacon: training for (future) implementers[25]	Online	FDD
	Beacon v2 Reference Implementation: An Overview[26]	Online	FDD
	Smörgåsbord 2022: Tapas Edition[27]	Online	FDA
	Data harmonisation workshop[28]	Online	HM, ELSI

[19] https://docs.google.com/document/d/19nFaar-Aylv8llEnBHWIk3SOAAZZMT18WqogXlUKWw8/edit#heading=h.gulcg6zenne

[20] https://www.cineca-project.eu/news-events-all/elixir-aai-online-advanced-training-for-relying-services

[21] https://www.cineca-project.eu/news-events-all/gtn-smrgsbord-a-global-galaxy-course

[22] https://www.cineca-project.eu/news-events-all/on-line-training-on-ga4gh-passports

[23] https://www.cineca-project.eu/news-events-all/how-fair-are-you-hackathon

[24] https://biohackathon-europe.org/

[25] https://www.cineca-project.eu/news-events-all/lighting-a-beacon-training-for-future-implementers

[26] https://www.cineca-project.eu/news-events-all/beacon-v2-reference-implementation-an-overview

[27] https://gallantries.github.io/video-library/events/smorgasbord2/tapas.html

[28] https://www.cineca-project.eu/news-events-all/data-harmonization-workshop

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

2023	Federated data analysis workshop[29]	Cape Town, South Africa	FDD, AAI, HM, FDA, ELSI
	Deploy and query a Beacon workshop at the University of Cape Town	Cape Town, South Africa	FDA
	Galaxy Smörgåsbord 3 (upcoming in May)	Online	FDA

3.2.4. Learning Pathway

CINECA deliverable D6.4 - *Training Programme, Detailed* (M24) mentioned that the project would develop "an end-to-end course covering all CINECA products including federated data analysis, incorporating tools, standards and use cases resulting from the technical work packages and incorporating training on the ELSI standard developed by work package 7". As updated in deliverable D6.5 - *Training Programme, Annual Report 2021* (M36), this course has been "implemented as a learning pathway that will cover all the necessary steps for performing federated data access and analysis. The choice of a learning pathway format is informed both by the pandemic and to aid the sustainability of the training materials."
The learning pathway on federated data analysis is now available on the EMBL-EBI Training website and is aimed at end-users. This open-access, self-paced learning resource can be accessed at any time and learners can move through the content at their own pace. It includes different types of training material, such as: explanations, videos, links to other resources and small exercises. The average time to read through the main body of the pathway is 1 to 3 hours, with around 0.5 days needed for addressing exercises and external links, depending on the learner's prior knowledge and how they choose to work through the course.

The pathway includes the following sections, which were developed in collaboration with experts in the various CINECA work packages:

- Introduction
- Ethical, legal and societal issues
- Find data:
    - Zenodo
    - EGA and ENA
    - Beacon
- Access data
- Federated data analysis:
    - Performing end-to-end molecular QTLs analysis with the eQTL Catalogue workflows
    - Using Galaxy for a Trio Analysis using Synthetic Datasets
- More information and resources

Each section of the pathway includes two User stories (Figure 6) to show how the content is applicable to two types of end-users: researchers using human data and bioinformatics software developers.

[29] https://www.cineca-project.eu/news-events-all/federated-data-analysis-workshop

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

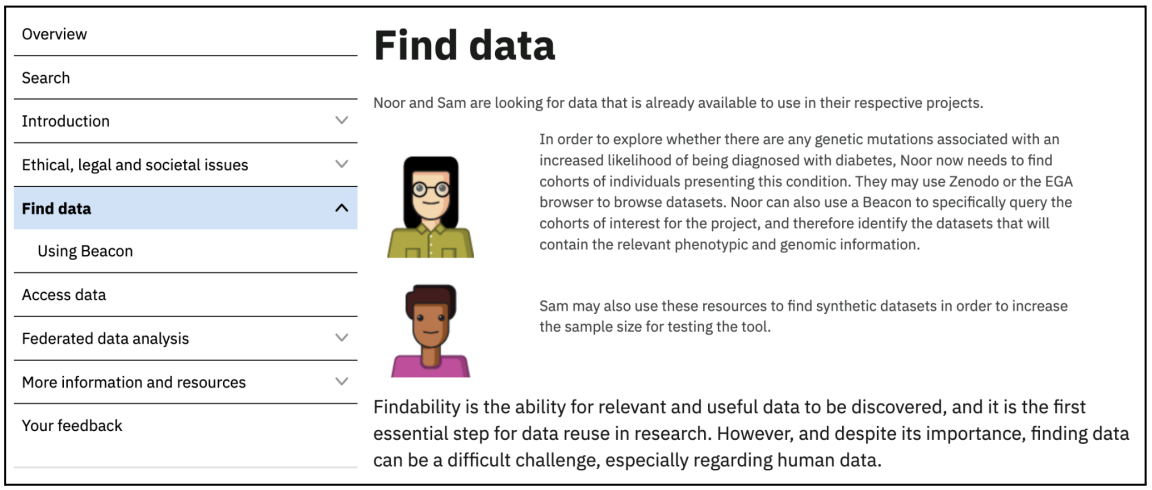

***Figure 6**: Screenshot showing the sections of the learning pathway as well as the two user stories.*

This pathway can be expanded in the future if new content becomes available. It is designed in such a way that a learner can follow all sections in order, or visit each section separately. At the end of the learning pathway, there is a section where learners can share feedback, so that the pathway can be improved.

At the workshop held in Cape Town on federated data analysis (26 February 2023) the learning pathway was shared with participants as part of piloting the dissemination of the pathway. We are now also focusing on advertising the pathway more widely through social media and the project partners.

3.2.5. Short training videos

A series of short training videos was created to facilitate the uptake of CINECA outputs, in addition to a handful of selected recordings from our training events. These videos can be used as stand-alone training resources by learners and can be sent out to participants attending future training courses to support them in developing prerequisite knowledge or in combination with other training resources that are covered during courses.

Between January 2022 and February 2023, four short videos were produced on query expansion in collaboration with work package 1 (Table 6). These videos have been added to a query expansion playlist[30] in our YouTube channel[31] together with the two videos produced by WP1 in December 2020 as part of D1.2 - *Query expansion service*[32]. The first video is an introduction to the query expansion services[33] and is followed by videos about specific types of query expansion services: horizontal

[30] https://www.youtube.com/playlist?list=PLD6bAdANKoMW_tO3A-VC2ILT93BaBFNBS
[31] https://www.youtube.com/channel/UCfEWFBY8z-TLJi1ije1oWug
[32] https://zenodo.org/record/4609335#.ZBSrUnbP1D8
[33] https://www.youtube.com/watch?v=Rly69Hrm9sw

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

expansion[34], a service that retrieves synonyms based on semantic similarities; vertical expansion[35], a service that retrieves hypernyms and hyponyms, a.k.a. the 'parents' and the 'children' of a concept in a specific ontology; and data driven expansion[36], a QE service based on a word2vec algorithm trained on 2 million PubMed abstracts. The description of these videos contains links and information on how to test the web services.

***Table 6**: CINECA short training videos produced between January 2022 and February 2023.*

Title	Length	Date published
Query expansion - introduction	3:20	January 2022
Query expansion - horizontal expansion	5:17	January 2022
Query expansion - vertical expansion	4:59	January 2022
Query expansion - data driven expansion	2:46	January 2022

Since the initiation of the CINECA project, 20 short videos have been created and are listed in Table 7. These videos can be accessed from the project's YouTube channel and in the dissemination section of the CINECA website[37]. To increase user engagement, the videos were also submitted to TeSS[38], ELIXIR's training portal, which provides a centralised location for users to access training materials. Furthermore, the videos are disseminated extensively through CINECA's various communication channels, including social media.

***Table 7**: All short training videos produced by CINECA (accessible through the CINECA Youtube[39] and CINECA blog[40] sites).*

Year	Title	Views on 16/03/23	Transcription available	CINECA Challenge
2020	Data harmonisation and enrichment using EMBL-EBI Ontology Tools	755	English	HM
	Solutions for overcoming cohort data integration challenges using ontology: an introduction	237	English	HM
	Annotating data using ontologies	1100	English	HM
	Annotating data using next generation biobanking ontology (NGBO)	82	English	HM

[34] https://www.youtube.com/watch?v=-QXueP2wTnU
[35] https://www.youtube.com/watch?v=xgXa_v6OOkU
[36] https://www.youtube.com/watch?v=h_SFEdGHCLo
[37] https://www.cineca-project.eu/short-videos
[38] https://tess.elixir-europe.org/
[39] https://www.youtube.com/@cinecaproject1265
[40] https://www.cineca-project.eu/blog-all/

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

	CanDIG and ELIXIR AAI interoperability demonstration - ELIXIR user accessing CanDIG service	56	Planned	AAI
	CanDIG and ELIXIR AAI interoperability demonstration - CanDIG user accessing ELIXIR service	37		AAI
	An introduction to CINECA's Service Catalog - Deliverable 1.1	81	English	FDD
	CINECA Discovery Service Catalog - Technical Walkthrough - Standalone Service	47	English	FDD
	CINECA Discovery Service Catalog - Technical Walkthrough - Beacon Integration	92	English	FDD
	Query Expansion Ontology Expansion	118	English	FDD
	Query Expansion Variant Expansion	66	English	FDD
2021	Applying data standards to the harmonisation of COVID 19 datasets from different sources	123	English	HM
	Useful ontologies for harmonising cohort data	115	English	HM
	A common framework for designing portable federated pipelines	104	English	FDA
	FAIRplus FAIRification wizard	51	Planned	HM
	Overview of the FAIRsharing.org	64		HM
2022	Query expansion - introduction	288		FDD
	Query expansion - horizontal expansion	65		FDD
	Query expansion - vertical expansion	49		FDD
	Query expansion - data driven expansion	114		FDD

The short demo videos were primarily targeted towards technical WPs, and as a result, no videos were produced for the ELSI framework. The first WPs to produce short training videos were WP2 (Authentication & Authorisation Infrastructure – AAI) and WP3 (Harmonised metadata – HM). Approximately 80% of the short demo videos were created by WP1 (Federated data discovery – FDD) and WP3 (Harmonised metadata – HM) (Figure 7). These two WPs were able to produce outputs right from the beginning of the project without waiting for input from other WPs, unlike WP4. The observation that WP3's short videos on YouTube are receiving more views than videos from other

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

WPs is valuable information that can be used to enhance the engagement and promotion of other videos, particularly those in the project's learning pathway.

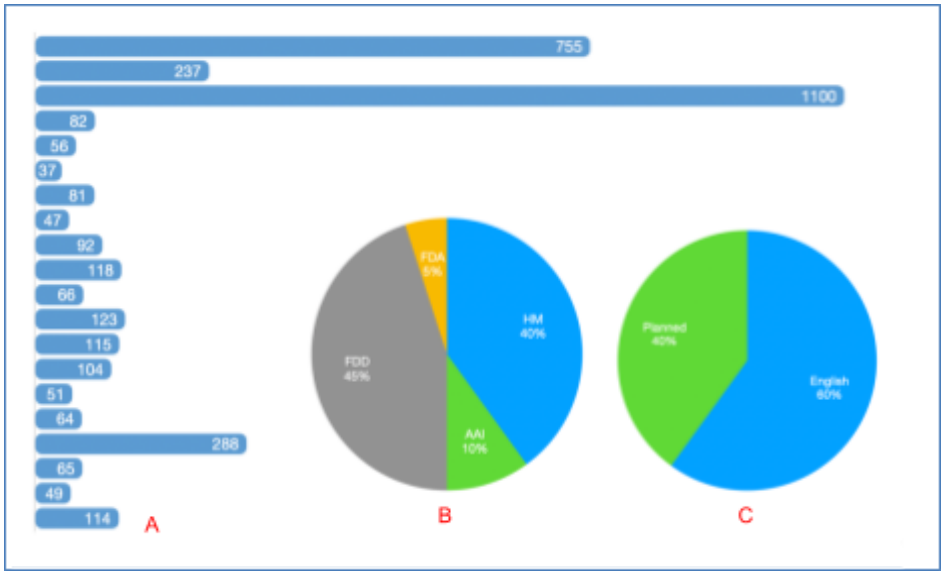

Figure 7: *Overall statistics on short technical videos produced by CINECA. (A) Short technical video views on the project CINECA Youtube channel as of 16/03/23. (B) CINECA challenges covered by the short technical videos. (C) Percentage of the short technical videos with English transcription.*

We took steps to enhance the accessibility of CINECA short videos and training materials to non-English speakers by creating subtitles. We have made significant progress by transcribing over 60% of the CINECA short demo videos into English, which will then be translated into other languages, such as French and Arabic. This initiative has been particularly effective in reaching out to our diverse stakeholder groups, as short demo videos have proven to be one of our most successful strategies. Our goal is to ensure that our resources are accessible to as many people as possible, regardless of language barriers.

3.2.6. Short technical reports

To ensure effective communication of the project's goals and outcomes to stakeholders, the CINECA project developed technical reports that provide detailed information about the project's activities and outputs. The reports were derived from WP deliverables and were specifically tailored to speak to technical stakeholders. The technical reports developed by the project team play a critical role in providing a detailed understanding of the project's work and products to technical stakeholders. These reports offer comprehensive insights into the methodologies and tools used to produce outputs, which can be highly technical in nature. These training materials can serve as a valuable resource for future projects and inform the ongoing discussion on the federated approach of data discovery and analyses.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

Between January 2022 and February 2023, one blog post was published in collaboration with WP2 on "A CINECA deliverable describing the ELIXIR Passport Broker"[41]. This blog post, which draws from WP2 deliverable D2.2 published in December 2021, provides an accessible introduction to GA4GH Passports 1.0. The post explains the concept of the Passport and Visas through practical examples and a demo video on real protocol flows and setups, making it an easy-to-read resource for anyone interested in learning more about this technology (Table 8).

***Table 8**: Summary of short technical reports published in 2022.*

Report title	Date	Author/Editor
A CINECA deliverable describing the ELIXIR Passport Broker	Feb 2022	M. Linden (CSC)

Throughout the CINECA project, we have produced 23 technical reports documenting the project's progress and outputs. To organise these reports clearly and systematically, they were grouped into thematic series covering topics relevant to the project's goals and objectives. These thematic series included the Connect with CINECA series (4 reports), CINECA Guest series (1 report), Text-mining series (4 reports), CINECA GA4GH standards series (4 reports), CINECA Use case series (1 report), and 9 CINECA short reports for any uncategorised reports[42]. Table 9 lists all the technical reports produced by the project with the corresponding CINECA challenge it addresses. Figure 8 summarises the CINECA challenges the reports address and the percentage for each challenge.

***Table 9**: Complete list of all CINECA technical reports (accessible through the CINECA blog site[43]).*

Year	Report title	CINECA challenge	Series
2019	CINECA stakeholder engagement session at IHCC	ELSI	Short Reports
	CINECA Staff Visit Program		
2020	Integration of new cohort infrastructures to the ELIXIR AAI	AAI	
	Connect with CINECA - Dr. Éloïse Gennet		Connect with CINECA
	Connect with CINECA - Amanjeev Sethi		
	Connect with CINECA - Saskia Hiltemann		
	Connect with CINECA - Emma Garlock		
	Sharing patient genomic and biomolecular data across continents	FDD, ELSI	Short Reports
	Sustainability and SMEs	FDA	

[41] https://www.cineca-project.eu/blog-all/a-cineca-deliverable-describes-the-elixir-passport-broker
[42] https://www.cineca-project.eu/blog
[43] https://www.cineca-project.eu/blog-all/

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

	Gary Saunders ELIXIR CINECA and the ELIXIR Federated Human Data Community.	FDD, FDA, ELSI	CINECA Guest
	Uncovering metadata from semi-structured cohort data	HM	Text mining pipeline
	LexMapr - A rule-based text-mining tool for ontology term mapping and classification	HM	
	Assigning standard descriptors to free text	HM	
	Biomedical Named entity recognition - Pros and cons of rule-based and deep learning methods	HM	
2021	Implementation of GA4GH standards in CINECA	FDD, FDA, HM, ELSI	GA4GH standards
	Beacon cohorts: A model for cohort discovery in CINECA and beyond	FDD	
	Passport is the glue between the researcher, data and computing	AAI	
	Powering up data discovery and access using the Data Use Ontology	AAI	
	Joint Variant genotyping use case	FDA	Use case
	“How FAIR are You” webinar series and hackathon	HM, ELSI	Short Reports
	CINECA project poster	FDD, FDA, HM, ELSI	
	EUCAN Projects General Factsheets		
2022	A CINECA deliverable describing the ELIXIR Passport Broker	AAI	

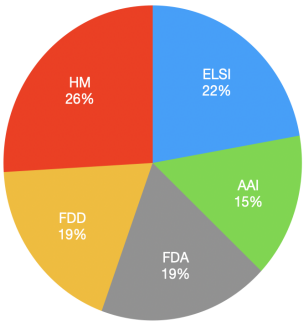

Figure 8: *Overall statistics on technical reports produced by CINECA.*

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

3.2.7. Dissemination activities

To ensure the dissemination of CINECA's outputs, our partners regularly attend conferences, workshops, and seminars, where they give talks or present posters. We maintain a comprehensive record of all the events that our partners have attended or presented at. This allows us to stay up-to-date with the latest research in our field and share our findings with the wider scientific community.

Despite the challenges posed by the pandemic, CINECA maintained its engagement and outreach efforts between January 2022 and February 2023. While face-to-face meetings have resumed, our engagement and outreach primarily occurred through online meetings and events. Table 10 provides a summary of these events, split into events that we have co-organised and events that we have attended as participants.

The CINECA partners have been actively engaging with their audience through various events, both organised by CINECA and externally, with a majority of them taking place online. The available data indicates that about 70% of the registered events were attended virtually, while approximately 20% occurred in Europe and 10% in Africa. Through this level of engagement, the CINECA project has been able to indirectly reach over 4000 individuals, both virtually and in person. The CINECA partners' ability to leverage technology and host events online has been a significant factor in the project's ability to connect with a broader audience, regardless of geographical barriers. The 20% of events that took place in person in Europe and the 10% in Africa highlight the importance of establishing a physical presence and building relationships with stakeholders in different regions.

***Table 10**: Selected events attended by CINECA partners between January 2022 and February 2023.*

Title of Event (External)	Location	People reached
Helmholtz Computational Health seminar	Munich	unknown
GA4GH DURI call: Self-Sovereign identity PoC and demo	online	15
GA4GH Connect Meeting	Online	>500
19th H3Africa Consortium Meeting	Nigeria	>500
ELIXIR Beacon, ELIXIR All Hands	Amsterdam	unknown
ELIXIR AAI Engagement call: Self-Sovereign identity PoC and demo	Online	10
Galaxy Community Conference 2022	Online	300
JSDI_Hyve Presentation_White Rabbit	Online	25
FAWG_Generating reliable evidence through a federated approach with OHDSI'	Online	81

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

BY-COVID Federated Analysis Solutions Workshop	Online	unknown
10th GA4GH Plenary 2022	Barcelona	unknown
Federated EGA Celebration Symposium	Barcelona	40
FASP Hackathon	Online	>30
BioIT Europe	Berlin	
International congress of human genetics 2023: Incorporating ELSI as a core support for international genomic data access and sharing	Cape Town	>500
Title of Event (CINECA-organised)	**Location**	**People reached**
Galaxy Smorgasbord 2022: Tapas Edition	online	>3000
Beacon v2 Reference Implementation: An Overview[44]	online	10
Lighting a Beacon: training for (future) implementers[45]	online	10
CINECA/H3ABioNet Data Harmonization workshop	online	40
Webinar: Bringing it all together: human cohort standards, tools and applications	Online	41
Webinar: Modular and reproducible workflows for federated molecular QTL analysis	Online	16
Tools and techniques to make sensitive data discoverable (Use-cases, hands-on session of Beacon implementation)	Online	20
Webinar: Federated analysis for Polygenic Risk Score Calculations	Online	62
Webinar: The case of data reuse: ethical, legal, and societal issues in international genomic data access and sharing	Online	60
A CINECA deliverable describes the ELIXIR Passport Broker	online	20
CINECA Federated data analysis	Cape Town	35

Benefit sharing and knowledge exchange

In addition to attending conferences and giving presentations, CINECA partners collaborate with various stakeholders on topics of mutual interest, as benefit sharing and knowledge exchange is a key goal for the CINECA project. Stakeholders include GA4GH, IHCC, ELIXIR, H3Africa, EUCAN Dissemination and Communication, EUCAN Data Catalogue, Joint Synthetic Data Initiative (JSDI), and

[44] https://www.cineca-project.eu/news-events-all/beacon-v2-reference-implementation-an-overview
[45] https://www.cineca-project.eu/news-events-all/lighting-a-beacon-training-for-future-implementers

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

EUCAN Harmonisation & Federated Analysis Working Groups. For example, CINECA representatives participate in the EUCAN Dissemination and Communication project, which aims to share dissemination efforts among all participating projects, including CINECA, euCanShare, EUCAN-Connect, iReceptor Plus, EUCANCan, and ReCoDID. This project releases regular newsletters that feature dissemination and outreach communications from all participating projects.

Another example is the Joint Synthetic Data Initiative (JSDI), which is a collaboration between a subset of EUCAN projects, B1MG, and INTERVENE. The JSDI aims to address the lack of available realistic test datasets, a significant bottleneck in developing federated analysis applications and technical demonstrators. The JSDI focuses on developing algorithms and generating synthetic genetic datasets for the next generation of synthetic datasets.

CINECA's collaboration with H3Africa has enabled the exchange of valuable knowledge and expertise with researchers working on microbiome projects in H3Africa and at the University of Cape Town. As part of this collaboration, CINECA's WP5 has demonstrated the clinical applications being developed and how Galaxy can be utilised for microbiome research. Through this partnership, CINECA has expanded its reach and engagement with experts in the field of microbiome research in Africa.

Website

The CINECA website has been an excellent dissemination tool for effectively communicating the project's goals and progress to a wide range of stakeholders worldwide in an accessible and engaging way. Through the website, the project has been able to easily reach a global audience and share updates and information about the project's progress. The website's ability to attract unique visitors and page views demonstrates its effectiveness in disseminating information about the CINECA project. Additionally, the website's diverse sources of traffic, including direct, search engine, and referral traffic, suggest that it is well-known and well-regarded within the research community.

Between January 2022 and February 2023, the CINECA website saw a significant increase in unique visitors, with a total of 9.9K unique visitors and 12K total visits, representing an increase of +13% and +1%, respectively, compared to the same period the previous year. However, the website experienced a slight decrease in page views, with over 20K page views recorded, representing a -12% decrease compared to the same period the previous year. Interestingly, the website received more direct traffic than search engine traffic, with direct traffic accounting for 50% of the visits and search engines accounting for 43%. Referrals from other websites accounted for 4%, while social media accounted for 3% (Figure 9). The top referring websites to the CINECA website during this period were elixir-europe.org, ichg2023.com, eosc-portal.eu, bbmri-eric.eu, b2ri-documentation.readthedocs.io, ga4gh.org, gallantries.github.io, competency.ebi.ac.uk, tess.elixir-europe.org, helmholtz-munich.de, and h3africa.org. These websites likely contributed to the increase in direct traffic to the CINECA website.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

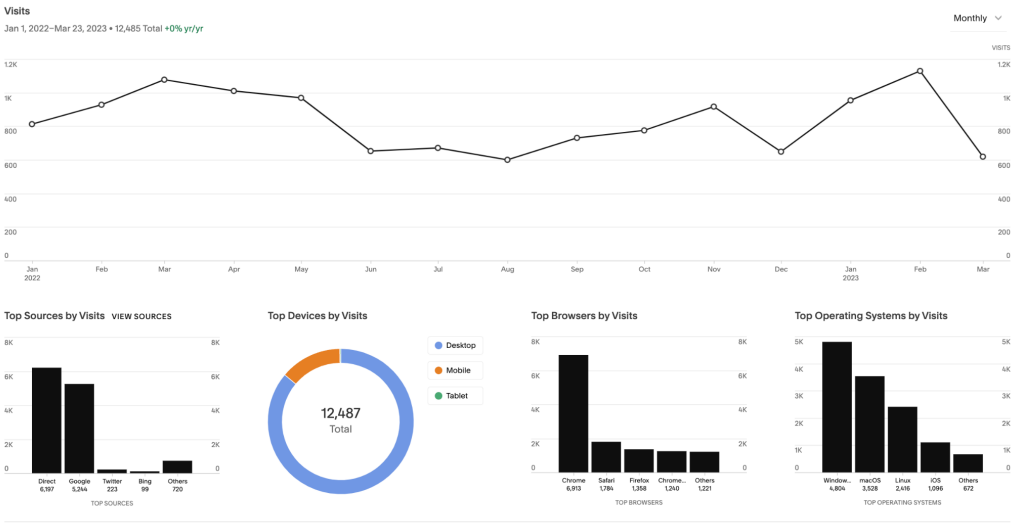

Figure 9: *Monthly CINECA website visits from January 2022 and February 2023, stratified by source, device, browser and operating system.*

During the period from January 2022 to February 2023, the CINECA website had several pages that received the most traffic. The homepage of the website, located at https://www.cineca-project.eu, was the most visited page. In addition, the technical report titled "Biomedical Named entity recognition - Pros and cons of rule-based and deep learning methods"[46], the History[47] page, the webinar titled "Bringing it all together: human cohort standards, tools and applications"[48], the News & Events[49] page, and the CINECA Synthetic Datasets[50] impact page were also among the most visited pages.

Furthermore, the CINECA website saw traffic coming from a variety of countries. The United States was the leading traffic source, accounting for 18% of the total traffic, followed by the United Kingdom with 12%. Canada, Germany, Spain, and the Netherlands each accounted for between 4-9% of the traffic, while South Africa accounted for 4%. These countries likely have a significant interest in the research and tools provided by the CINECA project (Figure 10).

46 https://www.cineca-project.eu/blog-all/biomedical-named-entity-recognition-pros-and-cons-of-rule-based-and-deep-learning-methods

47 https://www.cineca-project.eu/history

48 https://www.cineca-project.eu/news-events-all/bringing-it-all-together-human-cohort-standards-tools-and-applications

49 https://www.cineca-project.eu/news-events

50 https://www.cineca-project.eu/cineca-synthetic-datasets

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

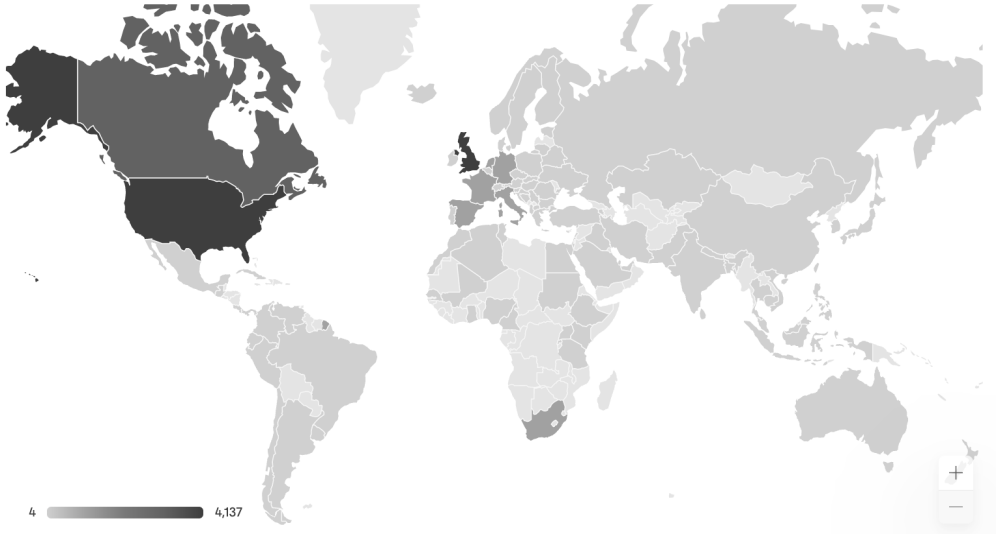

Figure 10: *CINECA website traffic geographic origin from January 2022 to February 2023.*

Overall, since the beginning of the project, the CINECA project website has experienced considerable traffic since its launch in May 2019, with interest from project partners and external stakeholders. Specifically, the website has received 29,000 visits, including 21,000 unique visitors with more than 54,000 pageviews. The CINECA website has experienced significant growth in traffic over the past few years, with an average of 7,000 unique visitors and 15,000 pageviews per year. In 2019, the website had about 2,000 visitors, but by 2020, this number had doubled to over 4,000. In 2021, the website experienced even more growth, with 10,000 visitors accessing the site. This growth trend continued in 2022, with 10,000 visitors accessing the website (Figure 11). The most visited pages all-time included the homepage, history, news & events, pages, and webinar pages "Biomedical Named entity recognition - Pros and cons of rule-based and deep learning methods"[51] and "Bringing it all together: human cohort standards, tools and applications" (Figure 12). A substantial number of individuals visit the CINECA website from various parts of the world, including but not limited to the United States, the United Kingdom, Canada, Germany, South Africa, and Spain. The website has also attracted visitors from other regions, such as Egypt, Uganda, and Nigeria in Africa, Japan and India in Asia, Brazil and Mexico in South America, France and Finland in Europe, and Australia. We have developed a comprehensive website that includes 49 pages categorised into four sections, namely "About CINECA," "How We Work," "Dissemination" and "Impact," as well as 23 engaging blog posts and 20 events.

51 https://www.cineca-project.eu/blog-all/biomedical-named-entity-recognition-pros-and-cons-of-rule-based-and-deep-learning-methods

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

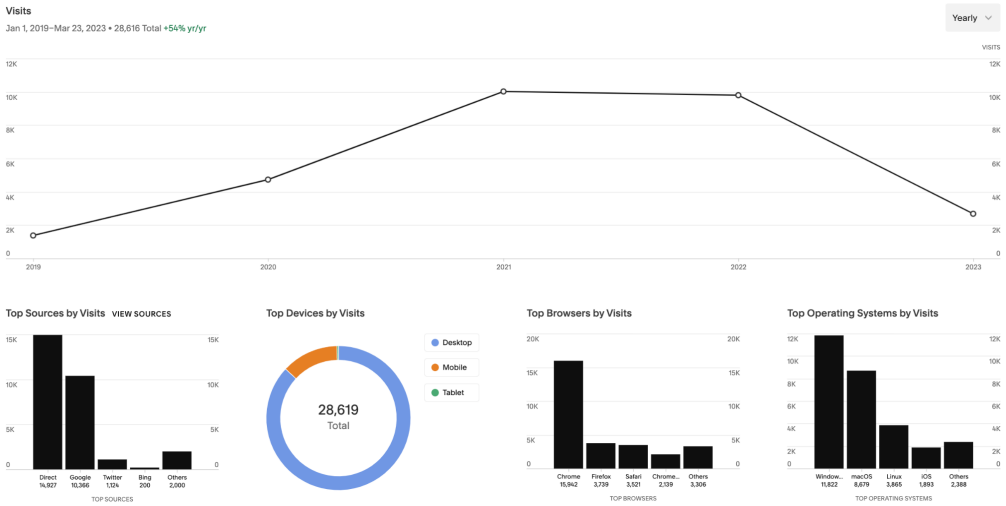

Figure 11: *CINECA website overall visits from* May 2019 to March 2023*.*

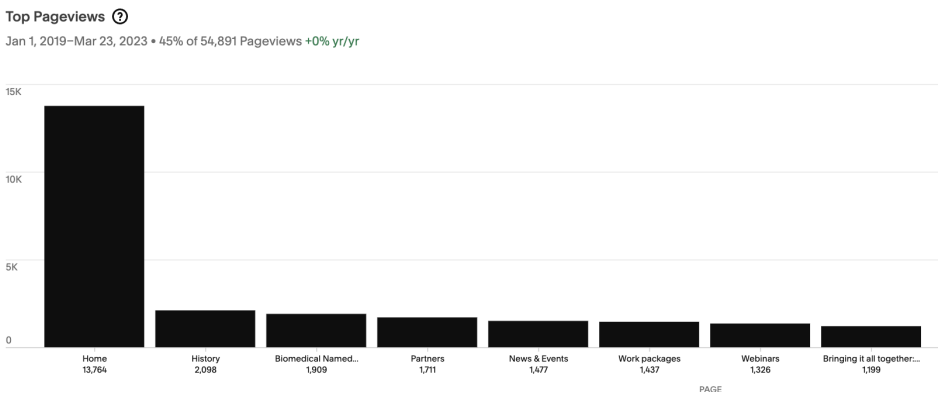

Figure 12: *CINECA website top viewed content from* May 2019 to March 2023*.*

Social media

Social media has also played a crucial role in promoting the CINECA project and driving traffic to its website. Through regular updates and posts on various platforms, such as Twitter and YouTube, the CINECA project has been able to increase its visibility and engagement with a broader audience. These efforts have contributed significantly to the project's success and helped it attract diverse visitors to its website.

CINECA maintains an active social media presence with accounts on popular platforms like Twitter and YouTube. The CINECA Twitter account[52] has grown significantly since its launch, with over 600 followers as of March 2023. The channel has tweeted over 700 times, an average of approximately one tweet every two days, highlighting the project's commitment to keeping its followers informed

[52] https://twitter.com/cinecaproject

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

about the latest news and developments related to CINECA and its partners. Similarly, the CINECA YouTube channel[53] features informative and engaging videos that provide insights into the project's progress and impact.

3.2.8. Publications

CINECA has compiled a list of peer-reviewed publications that feature work related to the CINECA project, which can be found on the project's website publication page[54]. The 14 peer-reviewed publications published in reputable journals highlight the impact and relevance of CINECA's contribution to the scientific community and the general public. This is a testament to the dedication and expertise of the CINECA project. The impact of these publications is likely to be far-reaching and may contribute to advancements in various fields of study. It is encouraging to see the impact that CINECA has made in this regard.

We have also published all deliverable reports for the project in Zenodo and have collated these on the CINECA website[55].

3.3 Conclusions and Sustainability

The activities developed and delivered in WP6 since the beginning of the project have contributed to our objectives and have ensured that we have had wide reach and wide impact within the community.

We have raised awareness of the CINECA project, as well as the opportunities and challenges of sharing cohort data on a global scale, by populating and maintaining the CINECA website and social media channels, collating publications, sharing CINECA successes at conferences and events, and publishing blog articles from consortium members. Furthermore, we have developed a diverse and engaging training programme based on the training needs and knowledge gaps that we identified both within the consortium as well as in the larger stakeholder community.

These initiatives have been immensely impactful, as evidenced by the wide reach of our dissemination and activities, as well as by the responses to the impact surveys for our larger training courses and staff exchanges. The results indicate that the obtained skills have been useful for participants in the long term, and that they have a better understanding of the community around the topic, and the participants have frequently passed on the skills to others around them.

[53] https://www.youtube.com/channel/UCfEWFBY8z-TLJi1ije1oWug
[54] https://www.cineca-project.eu/publications
[55] https://www.cineca-project.eu/deliverables

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.

To ensure that CINECA and the training programme continue to be a source of information and skills going forward beyond the project, the website will remain accessible online as a repository of all resources developed during the project including tools, publications, short training videos, recorded webinars, reports etc. Further, the self-paced learning pathway will continue to be accessible through the EMBL-EBI Training website and is an openly accessible resource that anyone can work through in their own time. In this way we ensure that the CINECA outputs remain useful to a wide audience even beyond the lifetime of the project.

4. Abbreviations

AAI - Access and Authorisation infrastructure
API - Application Programming Interface
B1MG - Beyond 1 Million Genomes
DUO - Data Use Ontology
EBI - European Bioinformatics Institute
EGA - European Genome-Phenome Archive
ELSI - Ethical, Legal, and Societal Issues
EMBL - European Molecular Biology Laboratory
ENA - European Nucleotide Archive
eQTL - Expression Quantitative Trait Locus
EUCAN - European-Canadian
FAIR - Findable, Accessible, Interoperable, Reusable
FASP - Federated Analysis Systems Project
FDD - Federated Data Discovery
FDA - Federated Data Analysis
GA4GH - Global Alliance for Genomics and Health
GTN - Galaxy Training Network
GCC - Galaxy Community Conference
HM - Harmonized Data
HMGU - Helmholtz Zentrum München - German Research Center for Environmental Health
IHCC - International HundredK+ Cohorts Consortium
JSDI - Joint Synthetic Data Initiative
PI - Principal Investigator
PoC - Proof of Concept
PRS - Polygenic Risk Score
Q&A - Question and Answer
QE - Query Expansion
QTL - Quantitative Trait Locus
SME - Small and Medium-sized Enterprise
WP - Work Package

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 825775.