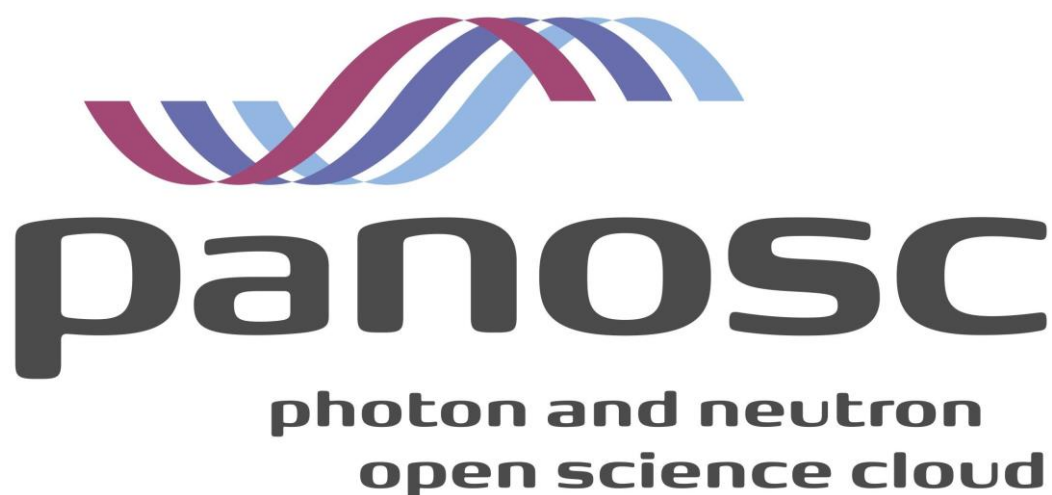


**PaNOSC**  
**Photon and Neutron Open Science Cloud**  
**H2020-INFRAEOSC-04-2018**  
**Grant Agreement Number: 823852**



**Deliverable:**  
**Publicly Accessible Demonstrator (4.4)**



This work is licensed under a Creative Commons Attribution 4.0 International License  
(<http://creativecommons.org/licenses/by/4.0/>)

# Project Deliverable Information Sheet

Project Reference No.	823852
Project acronym:	PaNOSC
Project full name:	Photon and Neutron Open Science Cloud
H2020 Call:	INFRAEOSC-04-2018
Project Coordinator	Andy Götz (andy.gotz@esrf.fr)
Coordinating Organization:	ESRF
Project Website:	<a href="http://www.panosc.eu">www.panosc.eu</a>
Deliverable No:	D4.4
Deliverable Type:	Demonstrator
Dissemination Level	Open
Contractual Delivery Date:	30/11/2022
Actual Delivery Date:	22/12/2022
EC project Officer:	Flavius Alexandru Pana

## Document Control Sheet

<b>Document</b>	Title: Publicly Accessible Demonstrator
	Version: 22 December 2022
	Available at: <a href="https://github.com/panosc-eu/panosc/tree/master/Submitted%20Deliverables">https://github.com/panosc - eu/panosc/tree/master/Submitted Deliverables</a>
	Files: 1
<b>Authorship</b>	Written by: Fabio Dall'Antonia
	Contributors: Jiri Majer, Axel Bocciarelli, Loic Huder, Thomas Vincent, Emiliano Coghetto, Stuart Cuant, Teodor Ivanoaica, Michael Schuh
	Reviewed by: Andy Götz
	Approved: Andy Götz

## List of participants

Participant No.	Participant organisation name	Country
1	European Synchrotron Radiation Facility (ESRF)	France
2	Institut Laue-Langevin (ILL)	France
3	European XFEL (XFEL.EU)	Germany
4	The European Spallation Source (ESS)	Sweden
5	ELI European Research Infrastructure Consortium (ELI-ERIC)	Belgium
6	Central European Research Infrastructure Consortium (CERIC-ERIC)	Italy
7	EGL Foundation (EGL.eu)	The Netherlands

# Table of Content

Project Deliverable Information Sheet	2
Document Control Sheet	2
List of participants	2
Table of Content	3
Introduction	4
Demonstrator scenarios	5
Open data access at ESRF	6
Open data access at CERIC-ERIC	9
Open data access at European XFEL	11
Conclusion and Outlook	15

# Introduction

The aspects of the FAIR principle, that is: Findability, Accessibility, Interoperability and Reusability (respectively Reproducibility) of scientific data, constitute not only a desirable, but meanwhile to a large extent even expected and required concept of good scientific practice. This holds for facility/user-exclusive (embargoed) data and open data alike.

The two major goals of the PaNOSC project are: first, making FAIR data a reality at the contributing research institutions (RIs), respectively facilities and, second, adding value to the European Open Science Cloud (EOSC) by providing means to expose and re-use open data on a federated and, to an extent, public level, including training material and services.

The PaNOSC work package four (WP4) has contributed to the aims of the project by means of software tool and software service development for scientific data analysis, that use and re-use raw scientific data with the aims of processing it and extracting meaning from the derived data, suited to produce publishable results.

Outcomes of WP4 include

- Infrastructure and services to utilise the Jupyter notebook ecosystem on distributed compute clusters
- Frameworks to inspect HDF5 format data (the de facto standard at all partner RIs) in the Jupyter environment, in local applications or embedded in web-services like data catalogues
- VISA: a cloud-based web platform for data analysis, offering both a remote desktop environment for GUI software and JupyterLab service in - optionally team-shared - virtual machine instances

The goal to align the outcomes horizontally has been successful as most of the products have been adopted, or are in the process of being adopted, by all of the partners. Moreover, the software developed is open-source, that is, source code of the respective software is freely accessible on GitHub and follows the best practices for software development. Deployed software services, as far as already used in production at the facilities, have been registered or are being registered on the EOSC marketplace.

The present demonstrator addresses a missing aspect, namely the connection of the aforementioned outcomes to open data. Demonstrator scenarios, as explained in the main part, start from a public and completely open (i.e. authentication-free) search portal. While such openness is desirable for encouraging re-usability, in reality both security and resource sustainability reasons justify the control over authentication to services and authorization of access in case of compute hardware use – as opposed to anonymous usage – even if considering a demonstrator use case. In the scope of this deliverable, open-access is understood in the following manner:

- URLs are public, i.e. not restricted to a local network and not requiring proxies or VPNs
- While authentication is required at certain stages, which may imply registration on first usage, there are no prerequisites to being registered (e. g. having been awarded beam-time at one of the facilities), meaning that there is no exclusivity of usage.

## Demonstrator scenarios

All use cases demonstrated in this section have a common starting point: the PaNOSC open data search portal at <https://data.panosc.eu>. The public portal (no authentication required) leads to the landing page of the web-site, serving a web front-end with a text query field to initiate the search. At the backend level, the portal employs a federated search API developed in PaNOSC WP3 (<https://github.com/SciCatProject/panosc-search-api>), connecting to open data search endpoints at the RI data catalogues. Upon retrieval of query matches, a result page is shown. Each data catalogue endpoint is connected to a facility-side scoring service based on a common algorithm (developed in WP3), in order to provide a relevance metric associated with a returned result. Upon aggregation of results, the back-end sorts the results by their scores. The result page features the optional use of search filters, such as experiment technique following the PaNET ontology (<https://expands-eu.github.io/ExPaNDS-experimental-techniques-ontology>), and important physical parameters associated with the experiment (probe wavelength, sample temperature, etc.)

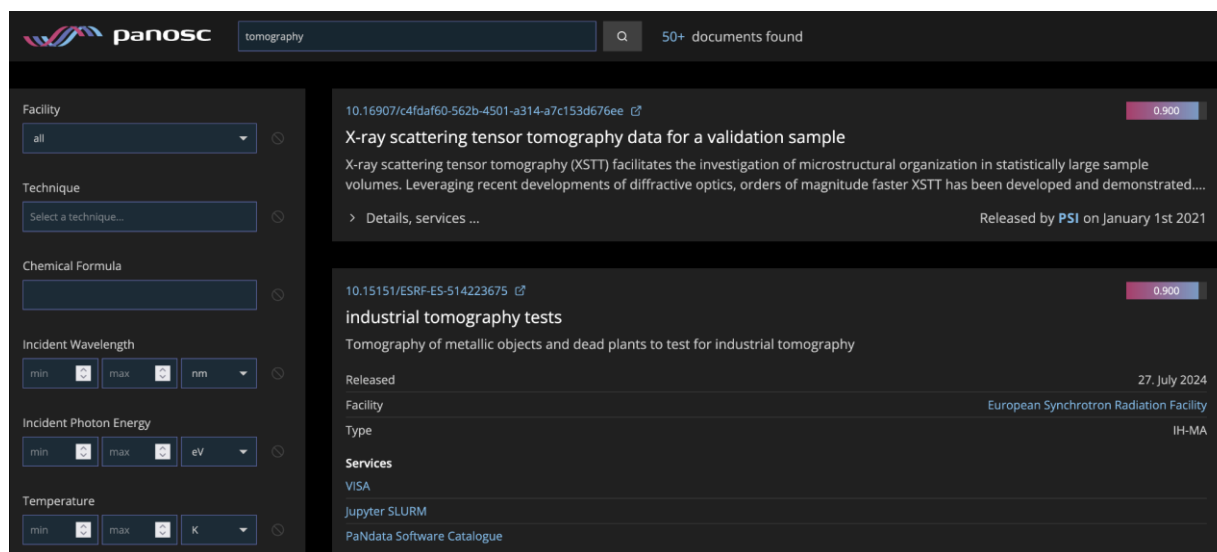


Fig. 1 Example of text query result representation in the PaNOSC search portal

Inspection and re-use of open data can start from the search portal following different routes. First, the result items on the portal redirect back to the DOI landing pages of the respective data catalogue, and some of the catalogue services have embedded tools to visualise and pre-analyse the raw data sets linked to by the DOI page. Second, search results can be expanded in the portal user interface and reveal URL links to existing services outside the data catalogues. Currently VISA is a demonstrator service accessible from the search portal in the case of European XFEL, ESRF and CERIC-ERIC.

Partner	Data catalogue (DOI target)	Result Granularity	Connected data analysis services	Service access level
ESRF	<a href="https://data.esrf.fr/doi/">https://data.esrf.fr/doi/</a>	Proposals or Experiments	<a href="https://visa.esrf.fr">https://visa.esrf.fr</a> <a href="https://jupyter-slurm.esrf.fr/">https://jupyter-slurm.esrf.fr/</a>	Authentication
ILL	<a href="https://doi.ill.fr/">https://doi.ill.fr/</a>	Proposals		
European XFEL	<a href="https://in.xfel.eu/metadata/doi/">https://in.xfel.eu/metadata/doi/</a>	Proposals	<a href="https://visa.xfel.eu">https://visa.xfel.eu</a>	Authentication [1]
ESS	<a href="https://doi.ess.eu/detail">https://doi.ess.eu/detail</a>			
ELI-ERIC				
CERIC-ERIC	<a href="https://vuo.elettra.eu/pls/vuo/open_access_data_portal.show_search">https://vuo.elettra.eu/pls/vuo/open_access_data_portal.show_search</a>	Proposals or Experiments	h5nuvola (from within catalogue)	Authentication [2]

[1] Helmholtz AAI, accepts also non-RI identity providers like GitHub, UmbrellaID

[2] Requires registration to Elettra's User Office (VUO), but no pre-conditions for registration

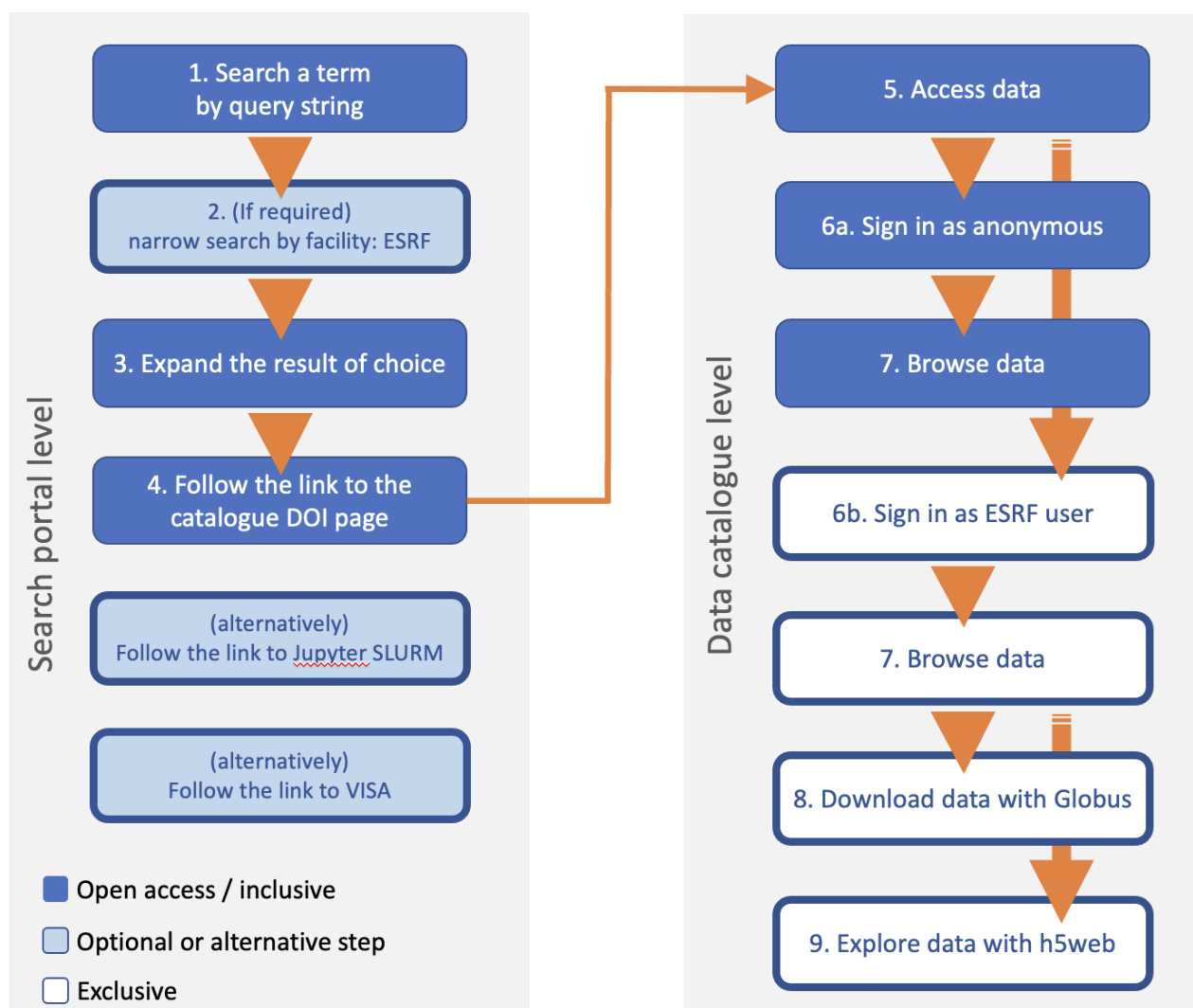
All search result entries provide a common link to the PaNdata Software Catalogue at <https://software.pan-data.eu>. This resource includes software services and tools developed for processing data from PaN facilities.

VISA has been designed to work with an authentication service such as Keycloak, that accepts OpenIdConnect tokens through an authentication and authorisation infrastructure (AAI). By accepting information from multiple remote identity providers, a PaN community (and beyond) level of openness is achieved. In the scope of PaNOSC (WP6), the UmbrellaID was envisaged as a common way to authenticate at respective facilities.

## Open data access at ESRF

The options to follow from a given search result entry are, besides the DOI landing page of the experimental data at the ESRF catalogue, VISA and Jupyter-SLURM as data analysis services. Currently these two service links directly to the respective ESRF service home page without specifying data to be pre-selected.

The ESRF data catalogue exposes all experiment proposal entries with a DOI to the search API, whether these are still under embargo (public release data in the future) or have become public already. On the DOI landing page, embargoed datasets are flagged as such and access is available only to the experimental team. Open data are accessible without restriction - upon following the respective access button, a login as user or anonymous login are available. After the login process, which requires no special registration for open data, the actual data catalogue page with a listing of data sets is opened, ready for browsing.



**Figure 2.** Schematic workflow of open data accessibility at ESRF. The essential chain of unrestricted user action steps starting from the initial search for data of interest by keyword terms and ending with the inspection of data sets is highlighted by blue boxes

Open data is not yet downloadable or visualisable by h5web at the moment because open proposals have been archived and they need to be restored from tape to disk-drive. Currently the path of the restored datasets is not automatically accessible by the download service or the h5web service. The path has to be specified manually. The data use options are added to the above work-flow to showcase the principal concept and desired behaviour in the future, whereas currently the relevant services work only with recent proposal data at regular disk storage paths, which applies mainly to data sets still under embargo. In practice, this makes the download and visualisation services exclusive to ESRF users, and more precisely: the subset of experiment team members. This will evolve in the future to provide access to viewing of open data for unauthenticated users too. More sophisticated viewers are being tested for large datasets e.g. <https://human-organ-atlas.esrf.fr/reconstructions/>

1.-3.

Facility: European Synchrotron Radiation Facility

Technique: Select a technique...

Chemical Formula:

Incident Wavelength: min max nm

Incident Photon Energy: min max eV

Temperature: min max K

10.15151/ESRF-ES-103850884  
**ihls3102**  
beamline alignment, fluorescence on cancer cells declared under proposal Is2767  
Released by ESRF on June 12th 2021

10.15151/ESRF-ES-105804379  
**ihls3112**  
Fluorescence Tomography and holotomography of cancer cell  
Released: 26. June 2021  
Facility: European Synchrotron Radiation Facility  
Type: IH-LS  
Services: VISA, Jupyter SLURM, PanData Software Catalogue

**ihls3112**  
Fluorescence Tomography and holotomography of cancer cell

Experimental Data: Access data

References: Researchers must acknowledge the source of the data and cite its unique identifier as well as any publications linked to the same raw data. Refer to the recommended format for citing this work to increase visibility.

Cited by: No citations were found.

Related DOIs: No related DOIs were found.

Data Portal: Find, browse and download your data

7.

Data Portal: My Data, Open Data, Closed Data

ihls3112 20/06/2018 24/06/2018 ihls3112

Dataset Link: Login

For users that want to download a large volume of experimental data (ESRF users can access the Globus service, please read the documentation for downloading: https://panosc.esrf.eu/ibiblio/Globus-Guide)

Date	Sample	Dataset
2022 24 Jun 2018	C3_H3_pmc3_24h	ihls3112
1601 24 Jun 2018	C3_H3_pmc3_24h	ihls3111
1559 24 Jun 2018	C3_H3_pmc3_24h	ihls3110
1546 24 Jun 2018	C3_H3_pmc3_24h	ihls3109
1542 24 Jun 2018	C3_H3_pmc3_24h	ihls3108
1537 24 Jun 2018	C3_H3_pmc3_24h	ihls3107
1536 24 Jun 2018	C3_H3_pmc3_24h	ihls3106
1530 24 Jun 2018	C3_H3_pmc3_24h	ihls3105
1527 24 Jun 2018	C3_H3_pmc3_24h	ihls3104
1525 24 Jun 2018	C3_H3_pmc3_24h	ihls3103
1524 24 Jun 2018	C3_H3_pmc3_24h	ihls3102
1518 24 Jun 2018	C3_H3_pmc3_24h	ihls3101
1516 24 Jun 2018	C3_H3_pmc3_24h	ihls3100
1512 24 Jun 2018	C3_H3_pmc3_24h	ihls3099
1509 24 Jun 2018	C3_H3_pmc3_24h	ihls3098
1504 24 Jun 2018	C3_H3_pmc3_24h	ihls3097
1503 24 Jun 2018	C3_H3_pmc3_24h	ihls3096
1502 24 Jun 2018	C3_H3_pmc3_24h	ihls3095
1501 24 Jun 2018	C3_H3_pmc3_24h	ihls3094
1500 24 Jun 2018	C3_H3_pmc3_24h	ihls3093
1499 24 Jun 2018	C3_H3_pmc3_24h	ihls3092
1498 24 Jun 2018	C3_H3_pmc3_24h	ihls3091
1497 24 Jun 2018	C3_H3_pmc3_24h	ihls3090
1496 24 Jun 2018	C3_H3_pmc3_24h	ihls3089
1495 24 Jun 2018	C3_H3_pmc3_24h	ihls3088
1494 24 Jun 2018	C3_H3_pmc3_24h	ihls3087
1493 24 Jun 2018	C3_H3_pmc3_24h	ihls3086
1492 24 Jun 2018	C3_H3_pmc3_24h	ihls3085
1491 24 Jun 2018	C3_H3_pmc3_24h	ihls3084
1490 24 Jun 2018	C3_H3_pmc3_24h	ihls3083
1489 24 Jun 2018	C3_H3_pmc3_24h	ihls3082
1488 24 Jun 2018	C3_H3_pmc3_24h	ihls3081
1487 24 Jun 2018	C3_H3_pmc3_24h	ihls3080
1486 24 Jun 2018	C3_H3_pmc3_24h	ihls3079
1485 24 Jun 2018	C3_H3_pmc3_24h	ihls3078
1484 24 Jun 2018	C3_H3_pmc3_24h	ihls3077
1483 24 Jun 2018	C3_H3_pmc3_24h	ihls3076
1482 24 Jun 2018	C3_H3_pmc3_24h	ihls3075
1481 24 Jun 2018	C3_H3_pmc3_24h	ihls3074
1480 24 Jun 2018	C3_H3_pmc3_24h	ihls3073
1479 24 Jun 2018	C3_H3_pmc3_24h	ihls3072
1478 24 Jun 2018	C3_H3_pmc3_24h	ihls3071
1477 24 Jun 2018	C3_H3_pmc3_24h	ihls3070
1476 24 Jun 2018	C3_H3_pmc3_24h	ihls3069
1475 24 Jun 2018	C3_H3_pmc3_24h	ihls3068
1474 24 Jun 2018	C3_H3_pmc3_24h	ihls3067
1473 24 Jun 2018	C3_H3_pmc3_24h	ihls3066
1472 24 Jun 2018	C3_H3_pmc3_24h	ihls3065
1471 24 Jun 2018	C3_H3_pmc3_24h	ihls3064
1470 24 Jun 2018	C3_H3_pmc3_24h	ihls3063
1469 24 Jun 2018	C3_H3_pmc3_24h	ihls3062
1468 24 Jun 2018	C3_H3_pmc3_24h	ihls3061
1467 24 Jun 2018	C3_H3_pmc3_24h	ihls3060
1466 24 Jun 2018	C3_H3_pmc3_24h	ihls3059
1465 24 Jun 2018	C3_H3_pmc3_24h	ihls3058
1464 24 Jun 2018	C3_H3_pmc3_24h	ihls3057
1463 24 Jun 2018	C3_H3_pmc3_24h	ihls3056
1462 24 Jun 2018	C3_H3_pmc3_24h	ihls3055
1461 24 Jun 2018	C3_H3_pmc3_24h	ihls3054
1460 24 Jun 2018	C3_H3_pmc3_24h	ihls3053
1459 24 Jun 2018	C3_H3_pmc3_24h	ihls3052
1458 24 Jun 2018	C3_H3_pmc3_24h	ihls3051
1457 24 Jun 2018	C3_H3_pmc3_24h	ihls3050
1456 24 Jun 2018	C3_H3_pmc3_24h	ihls3049
1455 24 Jun 2018	C3_H3_pmc3_24h	ihls3048
1454 24 Jun 2018	C3_H3_pmc3_24h	ihls3047
1453 24 Jun 2018	C3_H3_pmc3_24h	ihls3046
1452 24 Jun 2018	C3_H3_pmc3_24h	ihls3045
1451 24 Jun 2018	C3_H3_pmc3_24h	ihls3044
1450 24 Jun 2018	C3_H3_pmc3_24h	ihls3043
1449 24 Jun 2018	C3_H3_pmc3_24h	ihls3042
1448 24 Jun 2018	C3_H3_pmc3_24h	ihls3041
1447 24 Jun 2018	C3_H3_pmc3_24h	ihls3040
1446 24 Jun 2018	C3_H3_pmc3_24h	ihls3039
1445 24 Jun 2018	C3_H3_pmc3_24h	ihls3038
1444 24 Jun 2018	C3_H3_pmc3_24h	ihls3037
1443 24 Jun 2018	C3_H3_pmc3_24h	ihls3036
1442 24 Jun 2018	C3_H3_pmc3_24h	ihls3035
1441 24 Jun 2018	C3_H3_pmc3_24h	ihls3034
1440 24 Jun 2018	C3_H3_pmc3_24h	ihls3033
1439 24 Jun 2018	C3_H3_pmc3_24h	ihls3032
1438 24 Jun 2018	C3_H3_pmc3_24h	ihls3031
1437 24 Jun 2018	C3_H3_pmc3_24h	ihls3030
1436 24 Jun 2018	C3_H3_pmc3_24h	ihls3029
1435 24 Jun 2018	C3_H3_pmc3_24h	ihls3028
1434 24 Jun 2018	C3_H3_pmc3_24h	ihls3027
1433 24 Jun 2018	C3_H3_pmc3_24h	ihls3026
1432 24 Jun 2018	C3_H3_pmc3_24h	ihls3025
1431 24 Jun 2018	C3_H3_pmc3_24h	ihls3024
1430 24 Jun 2018	C3_H3_pmc3_24h	ihls3023
1429 24 Jun 2018	C3_H3_pmc3_24h	ihls3022
1428 24 Jun 2018	C3_H3_pmc3_24h	ihls3021
1427 24 Jun 2018	C3_H3_pmc3_24h	ihls3020
1426 24 Jun 2018	C3_H3_pmc3_24h	ihls3019
1425 24 Jun 2018	C3_H3_pmc3_24h	ihls3018
1424 24 Jun 2018	C3_H3_pmc3_24h	ihls3017
1423 24 Jun 2018	C3_H3_pmc3_24h	ihls3016
1422 24 Jun 2018	C3_H3_pmc3_24h	ihls3015
1421 24 Jun 2018	C3_H3_pmc3_24h	ihls3014
1420 24 Jun 2018	C3_H3_pmc3_24h	ihls3013
1419 24 Jun 2018	C3_H3_pmc3_24h	ihls3012
1418 24 Jun 2018	C3_H3_pmc3_24h	ihls3011
1417 24 Jun 2018	C3_H3_pmc3_24h	ihls3010
1416 24 Jun 2018	C3_H3_pmc3_24h	ihls3009
1415 24 Jun 2018	C3_H3_pmc3_24h	ihls3008
1414 24 Jun 2018	C3_H3_pmc3_24h	ihls3007
1413 24 Jun 2018	C3_H3_pmc3_24h	ihls3006
1412 24 Jun 2018	C3_H3_pmc3_24h	ihls3005
1411 24 Jun 2018	C3_H3_pmc3_24h	ihls3004
1410 24 Jun 2018	C3_H3_pmc3_24h	ihls3003
1409 24 Jun 2018	C3_H3_pmc3_24h	ihls3002
1408 24 Jun 2018	C3_H3_pmc3_24h	ihls3001
1407 24 Jun 2018	C3_H3_pmc3_24h	ihls3000
1406 24 Jun 2018	C3_H3_pmc3_24h	ihls2999
1405 24 Jun 2018	C3_H3_pmc3_24h	ihls2998
1404 24 Jun 2018	C3_H3_pmc3_24h	ihls2997
1403 24 Jun 2018	C3_H3_pmc3_24h	ihls2996
1402 24 Jun 2018	C3_H3_pmc3_24h	ihls2995
1401 24 Jun 2018	C3_H3_pmc3_24h	ihls2994
1400 24 Jun 2018	C3_H3_pmc3_24h	ihls2993
1399 24 Jun 2018	C3_H3_pmc3_24h	ihls2992
1398 24 Jun 2018	C3_H3_pmc3_24h	ihls2991
1397 24 Jun 2018	C3_H3_pmc3_24h	ihls2990
1396 24 Jun 2018	C3_H3_pmc3_24h	ihls2989
1395 24 Jun 2018	C3_H3_pmc3_24h	ihls2988
1394 24 Jun 2018	C3_H3_pmc3_24h	ihls2987
1393 24 Jun 2018	C3_H3_pmc3_24h	ihls2986
1392 24 Jun 2018	C3_H3_pmc3_24h	ihls2985
1391 24 Jun 2018	C3_H3_pmc3_24h	ihls2984
1390 24 Jun 2018	C3_H3_pmc3_24h	ihls2983
1389 24 Jun 2018	C3_H3_pmc3_24h	ihls2982
1388 24 Jun 2018	C3_H3_pmc3_24h	ihls2981
1387 24 Jun 2018	C3_H3_pmc3_24h	ihls2980
1386 24 Jun 2018	C3_H3_pmc3_24h	ihls2979
1385 24 Jun 2018	C3_H3_pmc3_24h	ihls2978
1384 24 Jun 2018	C3_H3_pmc3_24h	ihls2977
1383 24 Jun 2018	C3_H3_pmc3_24h	ihls2976
1382 24 Jun 2018	C3_H3_pmc3_24h	ihls2975
1381 24 Jun 2018	C3_H3_pmc3_24h	ihls2974
1380 24 Jun 2018	C3_H3_pmc3_24h	ihls2973
1379 24 Jun 2018	C3_H3_pmc3_24h	ihls2972
1378 24 Jun 2018	C3_H3_pmc3_24h	ihls2971
1377 24 Jun 2018	C3_H3_pmc3_24h	ihls2970
1376 24 Jun 2018	C3_H3_pmc3_24h	ihls2969
1375 24 Jun 2018	C3_H3_pmc3_24h	ihls2968
1374 24 Jun 2018	C3_H3_pmc3_24h	ihls2967
1373 24 Jun 2018	C3_H3_pmc3_24h	ihls2966
1372 24 Jun 2018	C3_H3_pmc3_24h	ihls2965
1371 24 Jun 2018	C3_H3_pmc3_24h	ihls2964
1370 24 Jun 2018	C3_H3_pmc3_24h	ihls2963
1369 24 Jun 2018	C3_H3_pmc3_24h	ihls2962
1368 24 Jun 2018	C3_H3_pmc3_24h	ihls2961
1367 24 Jun 2018	C3_H3_pmc3_24h	ihls2960
1366 24 Jun 2018	C3_H3_pmc3_24h	ihls2959
1365 24 Jun 2018	C3_H3_pmc3_24h	ihls2958
1364 24 Jun 2018	C3_H3_pmc3_24h	ihls2957
1363 24 Jun 2018	C3_H3_pmc3_24h	ihls2956
1362 24 Jun 2018	C3_H3_pmc3_24h	ihls2955
1361 24 Jun 2018	C3_H3_pmc3_24h	ihls2954
1360 24 Jun 2018	C3_H3_pmc3_24h	ihls2953
1359 24 Jun 2018	C3_H3_pmc3_24h	ihls2952
1358 24 Jun 2018	C3_H3_pmc3_24h	ihls2951
1357 24 Jun 2018	C3_H3_pmc3_24h	ihls2950
1356 24 Jun 2018	C3_H3_pmc3_24h	ihls2949
1355 24 Jun 2018	C3_H3_pmc3_24h	ihls2948
1354 24 Jun 2018	C3_H3_pmc3_24h	ihls2947
1353 24 Jun 2018	C3_H3_pmc3_24h	ihls2946
1352 24 Jun 2018	C3_H3_pmc3_24h	ihls2945
1351 24 Jun 2018	C3_H3_pmc3_24h	ihls2944
1350 24 Jun 2018	C3_H3_pmc3_24h	ihls2943
1349 24 Jun 2018	C3_H3_pmc3_24h	ihls2942
1348 24 Jun 2018	C3_H3_pmc3_24h	ihls2941
1347 24 Jun 2018	C3_H3_pmc3_24h	ihls2940
1346 24 Jun 2018	C3_H3_pmc3_24h	ihls2939
1345 24 Jun 2018	C3_H3_pmc3_24h	ihls2938
1344 24 Jun 2018	C3_H3_pmc3_24h	ihls2937
1343 24 Jun 2018	C3_H3_pmc3_24h	ihls2936
1342 24 Jun 2018	C3_H3_pmc3_24h	ihls2935
1341 24 Jun 2018	C3_H3_pmc3_24h	ihls2934
1340 24 Jun 2018	C3_H3_pmc3_24h	ihls2933
1339 24 Jun 2018	C3_H3_pmc3_24h	ihls2932
1338 24 Jun 2018	C3_H3_pmc3_24h	ihls2931
1337 24 Jun 2018	C3_H3_pmc3_24h	ihls2930
1336 24 Jun 2018	C3_H3_pmc3_24h	ihls2929
1335 24 Jun 2018	C3_H3_pmc3_24h	ihls2928
1334 24 Jun 2018	C3_H3_pmc3_24h	ihls2927
1333 24 Jun 2018	C3_H3_pmc3_24h	ihls2926
1332 24 Jun 2018	C3_H3_pmc3_24h	ihls2925
1331 24 Jun 2018	C3_H3_pmc3_24h	ihls2924
1330 24 Jun 2018	C3_H3_pmc3_24h	ihls2923
1329 24 Jun 2018	C3_H3_pmc3_24h	ihls2922
1328 24 Jun 2018	C3_H3_pmc3_24h	ihls2921
1327 24 Jun 2018	C3_H3_pmc3_24h	ihls2920
1326 24 Jun 2018	C3_H3_pmc3_24h	ihls2919
1325 24 Jun 2018	C3_H3_pmc3_24h	ihls2918
1324 24 Jun 2018	C3_H3_pmc3_24h	ihls2917
1323 24 Jun 2018	C3_H3_pmc3_24h	ihls2916
1322 24 Jun 2018	C3_H3_pmc3_24h	ihls2915
1321 24 Jun 2018	C3_H3_pmc3_24h	ihls2914
1320 24 Jun 2018	C3_H3_pmc3_24h	ihls2913
1319 24 Jun 2018	C3_H3_pmc3_24h	ihls2912
1318 24 Jun 2018	C3_H3_pmc3_24h	ihls2911
1317 24 Jun 2018	C3_H3_pmc3_24h	ihls2910
1316 24 Jun 2018	C3_H3_pmc3_24h	ihls2909
1315 24 Jun 2018	C3_H3_pmc3_24h	ihls2908
1314 24 Jun 2018	C3_H3_pmc3_24h	ihls2907
1313 24 Jun 2018	C3_H3_pmc3_24h	ihls2906
1312 24 Jun 2018	C3_H3_pmc3_24h	ihls2905
1311 24 Jun 2018	C3_H3_pmc3_24h	ihls2904
1310 24 Jun 2018	C3_H3_pmc3_24h	ihls2903
1309 24 Jun 2018	C3_H3_pmc3_24h	ihls2902
1308 24 Jun 2018	C3_H3_pmc3_24h	ihls2901
1307 24 Jun 2018	C3_H3_pmc3_24h	ihls2900
1306 24 Jun 2018	C3_H3_pmc3_24h	ihls2899
1305 24 Jun 2018	C3_H3_pmc3_24h	ihls2898
1304 24 Jun 2018	C3_H3_pmc3_24h	ihls2897
1303 24 Jun 2018	C3_H3_pmc3_24h	ihls2896
1302 24 Jun 2018	C3_H3_pmc3_24h	ihls2895
1301 24 Jun 2018	C3_H3_pmc3_24h	ihls2894
1300 24 Jun 2018	C3_H3_pmc3_24h	ihls2893
1299 24 Jun 2018	C3_H3_pmc3_24h	ihls2892
1298 24 Jun 2018	C3_H3_pmc3_24h	ihls2891
1297 24 Jun 2018	C3_H3_pmc3_24h	ihls2890
1296 24 Jun 2018	C3_H3_pmc3_24h	ihls2889
1295 24 Jun 2018	C3_H3_pmc3_24h	ihls2888
1294 24 Jun 2018	C3_H3_pmc3_24h	ihls2887
1293 24 Jun 2018	C3_H3_pmc3_24h	ihls2886
1292 24 Jun 2018	C3_H3_pmc3_24h	ihls2885
1291 24 Jun 2018	C3_H3_pmc3_24h	ihls2884
1290 24 Jun 2018	C3_H3_pmc3_24h	ihls2883
1289 24 Jun 2018	C3_H3_pmc3_24h	ihls2882
1288 24 Jun 2018	C3_H3_pmc3_24h	ihls2881
1287 24 Jun 2018	C3_H3_pmc3_24h	ihls2880
12		



## Open data access at CERIC-ERIC

Currently the only service available for a given search result entry is the DOI landing page of the experimental data at the CERIC-ERIC data catalogue.

The CERIC-ERIC data catalogue exposes to the search API only the experiment proposals that have become public. All the public experiment proposals are exposed with their corresponding DOIs and every DOI takes the user to the corresponding experiment proposal landing page in the CERIC-ERIC data catalogue.

Open data is available without restriction to registered users of the data catalogue. Registration to the data catalogue is available to everyone: there are no prerequisites to becoming a registered user.

An intermediate login page is shown when the user asks for access to the data on the corresponding DOI landing page. After the login process the actual data catalogue page opens with a listing of experiments associated with the selected experiment proposal ready for browsing, analysis and downloading.

A tool called h5nuvola (<https://github.com/ElettraSciComp/h5nuvola>) is available for data browsing and analysis. H5nuvola is a web-based HDF files viewer with additional functionalities: it implements file browsing, data visualisation services and selective exporting of data in the web browser. Users are required to authenticate to use h5nuvola.

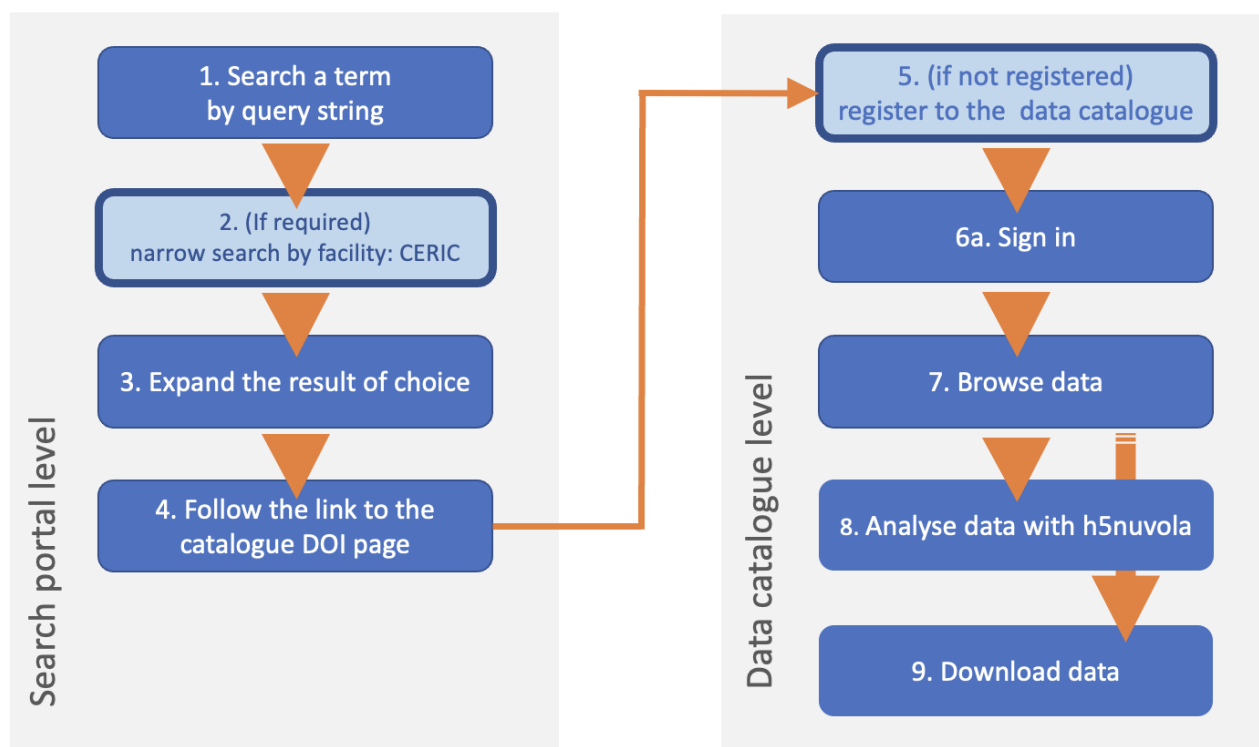


Figure 4: Schematic workflow of open data accessibility at CERIC-ERIC, same color scheme as for ESRF.

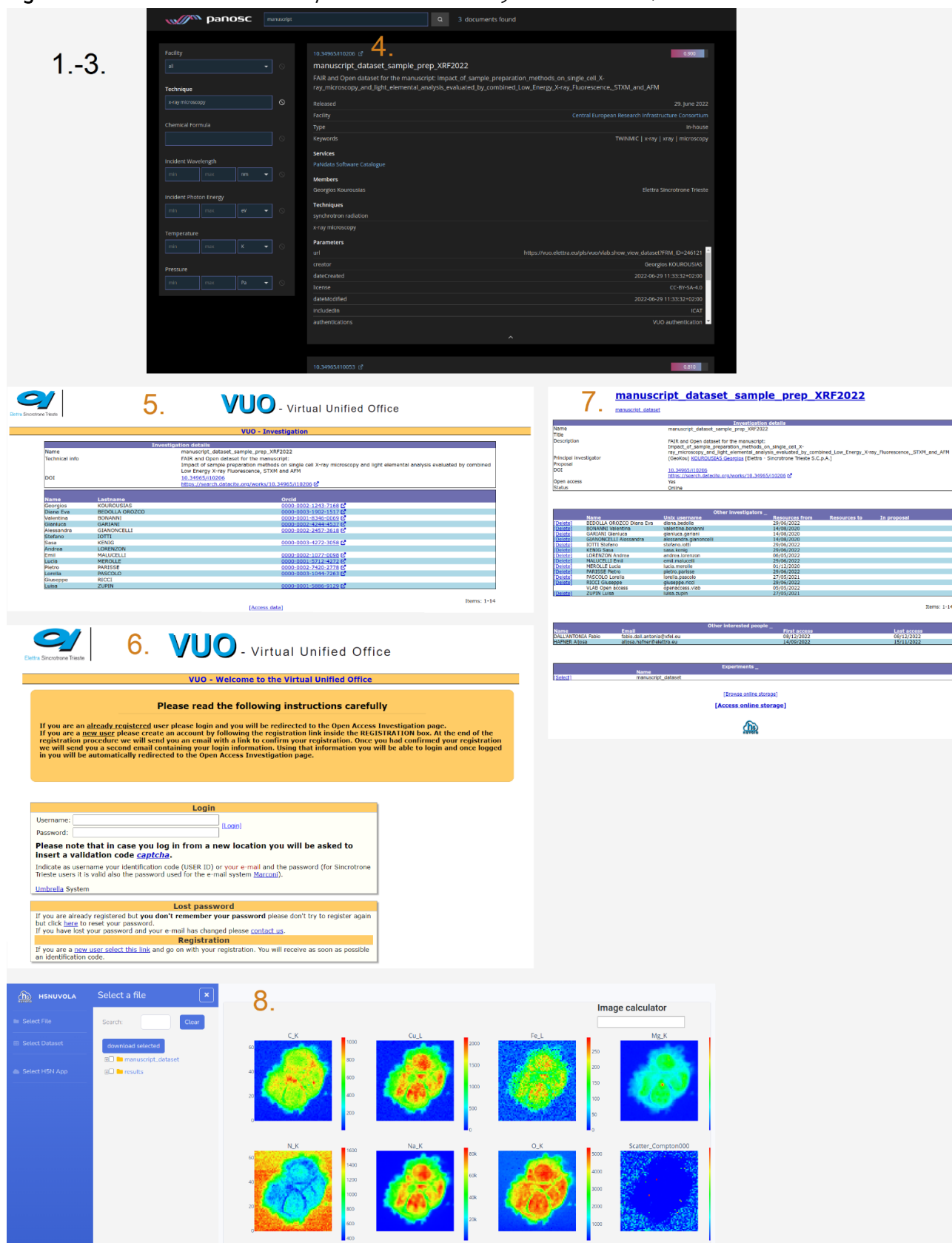
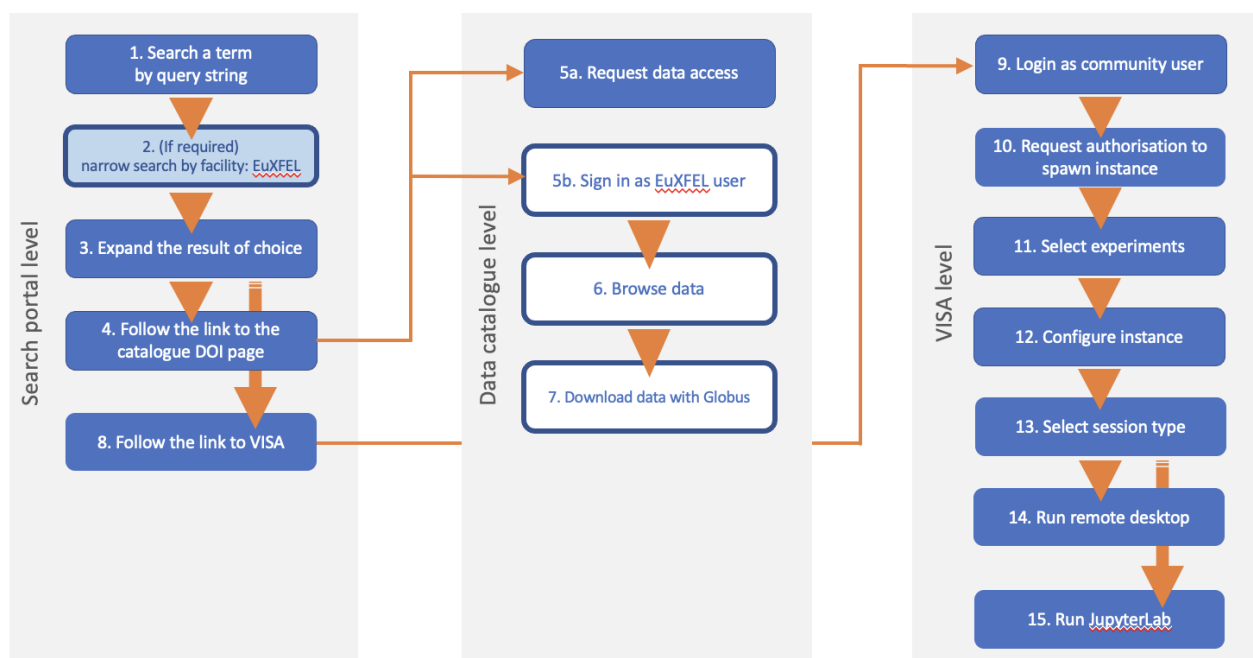


Figure 5: Demonstration of the CERIC-ERIC workflow by screenshots.

Figure 4 shows the corresponding demonstration of the CERIC-ERIC workflow by screenshots. In the depiction of the federated data search portal on top, steps 1 to 3 have already been performed: the results matching the “manuscript” term were filtered by the PaNET experiment technique “x-ray microscopy” and the entry of interest, “manuscript\_dataset\_sample\_prep\_XRF2022”, has been expanded. Clicking on the panel close to the DOI mention on top (4) will open a new browser tab with the DOI landing page on the catalogue site. On the following data portal login page the user has to log in (6) and once logged in he reaches the dedicated data portal page for the proposal of interest (7). From here the user can browse, download or analyse the data using the tool H5nuvola (7, 8).

## Open data access at European XFEL

Starting from the PaNOSC data search portal, currently two open proposals can be retrieved from European XFEL, as exposed to the search API by the metadata catalogue myMDC. These proposals represent collections of representative experiment data taken from various user proposals post embargo period, or commissioning proposals. These data sets cover a range of experimental techniques. The public DOI pages to which the search result entries link back mark the proposals as open, but do not yet provide the full functionality of the data catalogue interface for data set browsing or the Globus download service. One can get to this level only after signing in as EuXFEL user. Currently open data will be made available to any scientific guest user on request by e-mail. Coverage of the community re-use case with exploration and download on data catalogue level is being worked on; it will include a service for downloading via Globus.



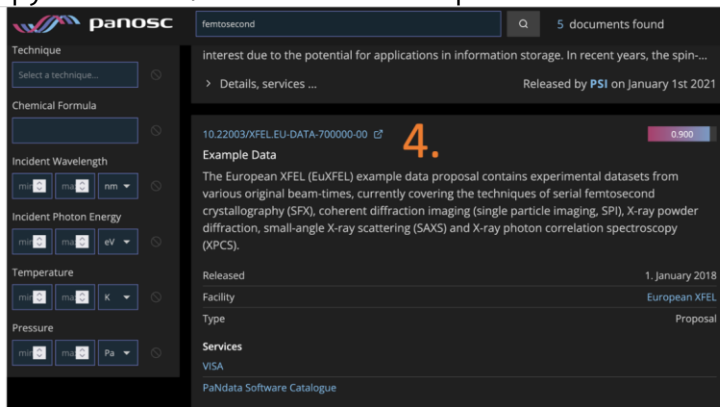
**Figure 6.** Schematic workflow of open data accessibility at European XFEL, same color scheme as for ESRF.

European XFEL results include service links to VISA. The deployed VISA service at <https://visa.xfel.eu> is accessible to all PaN community users after they have signed in to the Helmholtz AAI system with an authorised identity provider including GitHub and soon

Umbrellaid, and requested authorisation to build and use a VISA instance on the cloud cluster, hosted by DESY.

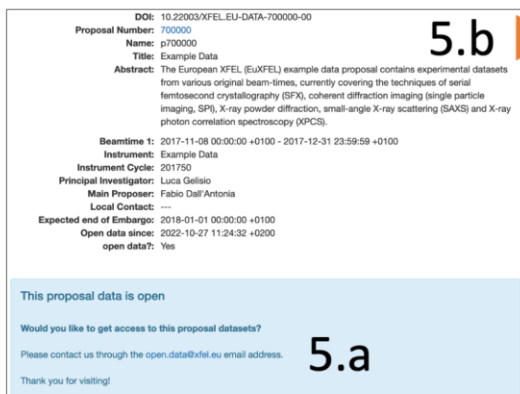
Since the portal link to VISA directs to a proposal-specific URL, the experiment data sets of the selected proposal are explorable in the VISA frontend (like they are in the data catalogue for XFEL users and will be, soon, for everyone). After selection of an experiment, the compute instance (virtual machine) can be configured and used in order to start a data analysis session based on JupyterLab and/or a remote desktop.

1.-3.

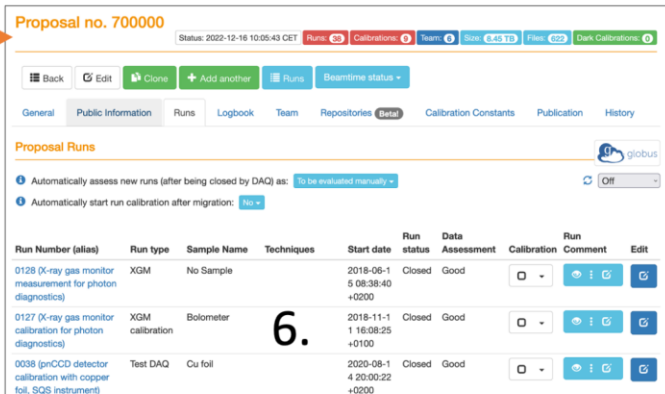


4.

5.a



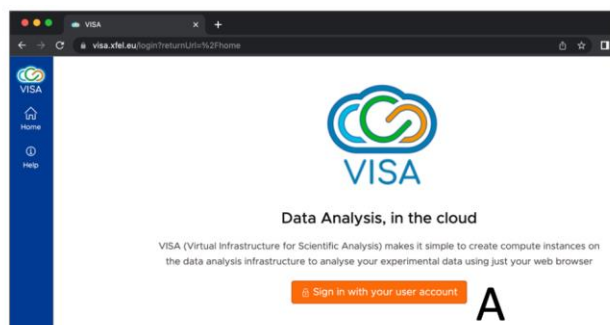
5.b



6.

7.

**Figure 7.** Demonstration of the EuXFEL workflow, on data search portal and catalogue (myMDC) level, by screenshots. For guest users the workflow currently ends at step 5a.



## Step 9

Deutsch

Bei Ihrem Konto anmelden

Benutzername

Passwort

☐ Angemeldet bleiben

Oder anmelden mit

Was email address verified?:

Name:

eduPerson Scoped Affiliation:

eduPersonUniqueId:

eduPerson Assurance (1):

eduPerson Assurance (2):

Givenname:

Familyname:

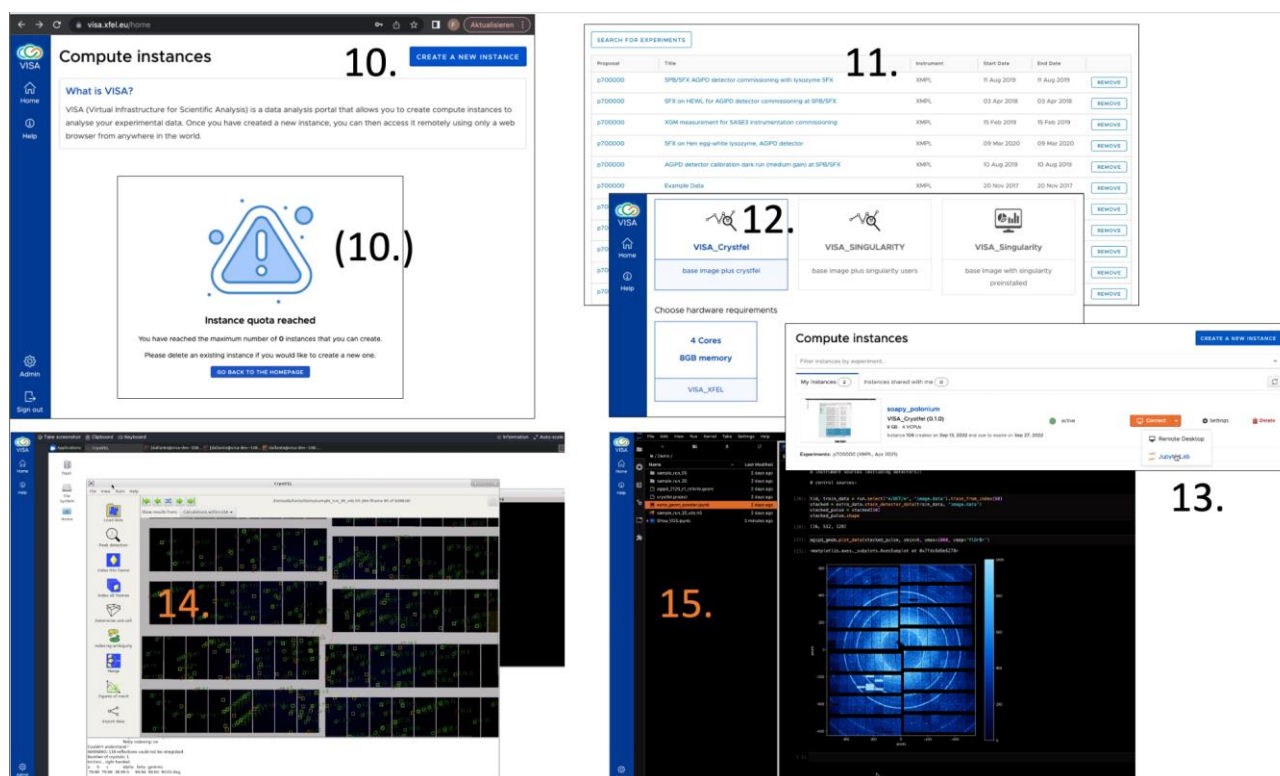
eduPersonEntitlement:

Email:

Note that in any case your credential (password, private key, ...) will NOT be exposed.

**Figure 8.** Login to VISA (step 9 of the workflow) via authentication to the Helmholtz AAI.

The login button (A) directs a user to the Keycloak authentication service, where the Helmholtz AAI is selected (B). Currently GitHub is an option for semi-anonymous identity provision (C). After external login to GitHub and confirmation by e-mail (D, E, not shown), the authentication information must be accepted ("Allow" button, F).



**Figure 9.** Use of VISA as service for remote data analysis, steps 10 to 15 of the workflow. The instance creation request (10.) will show a message for guest users which requires contacting the system administrator for quota increase, which means authorisation to use the cloud-cluster infrastructure for the subsequent steps.

## Conclusion and Outlook

The next step for the demonstrator scenarios lies in the implementation of the full concept into the respective production services. Such activities fall into the future after the end of the PaNOSC project and must be catered for by regular RI IT activities or projects funded by future community grants. Ideally all of the services developed in the PaNOSC scope should allow for access to open data, and usage by PaN community members, not necessarily belonging to the registered facility user groups. This poses challenges with respect to limitations of resources like hardware infrastructure and security aspects like prevention of misuse.

Given the current budgets facilities cannot easily sustain open access to compute resources at a scale that would be expected for community-wide open data re-use. Demonstrator setups as well as the catering of FAIR services to the own users of facilities can however spark the interest and even demand of such services on a broader basis, which will in turn create user numbers that convince funding agencies to support investments in more hardware and manpower to maintain the technology run on such hardware.

Finally, for a truly federated approach of data storage, exposition and compute resources, one could imagine centralised services hosted by a dedicated organisation which receives direct funding from commissions like EOSC. Of course, a flawless interplay between such a data hub and the data-producing facilities requires optimised standards and workflows. On the route to establish such, the PaNOSC project has delivered valuable insights and outcomes that can be built upon.