

Grant Proposal

Author-formatted document posted on 05/12/2023

Published in a RIO article collection by decision of the collection editors.

DOI: <https://doi.org/10.3897/arphapreprints.e116672>

Open Imaging Data Sharing in EOSC: COVID-19 as Demonstrator

 Reagon Karki,  Andrea Zaliani, Philip Gribbon,  Carolina Simon,  Jose Ramon Macias, 
Bugra Özdemir,  Aastha Mathur

Title: Open Imaging Data Sharing in EOSC / COVID-19 as Demonstrator

Authors: Karki, R., Zaliani, A., Gribbon, P., Simon, C., Macias, J.R., Oezdemir, B., Mathur, A.

Abstract

This Science Project (SP) brings together three different domains of life sciences with the aim to create reproducible workflows, tools and web-services for data visualization. This SP focuses in building resources for handling data from bioimaging, structural and bio-chemical studies. The Euro-Bioimaging will implement a community standard cloud compatible open image data format and data submission workflow for high-throughput screening data. Whereas, Instruct-ERIC will develop a user-friendly web-service to access to multi-dimensional structural and imaging data. Lastly, EU-OpenScreen/Fraunhofer ITMP will create reproducible workflow for generating Knowledge Graphs that represent phenotype-chemotype of diseases. While these resources are being developed, the collaborators will also simultaneously harmonize the resources right from the beginning to enable FAIR data principles. This SP uses COVID-19 as a demonstrator, however the resources will be generalized for any disease of interest.

Keywords: Imaging data, Knowledge Graphs, 3Dbionotes, workflow, covid19, infectious diseases

Description

i. **Existing situation:** A plethora of formats for imaging data and inability of most of them to capture proper metadata has been an issue for a while. This situation demands a community accepted standard file format in the imaging community. Moreover, a cloud-based service to handle these files and a user-friendly service to publish the data in public repositories is also the need of the hour. Likewise, in structural biology, an intuitive and interactive service to browse and visualize protein structures with associated metadata is required to facilitate researchers. Lastly, a reproducible Knowledge Graph (KG) workflow to integrate chemotype-phenotype of disease of interests to provide new insights to the scientific community is lacking. Therefore, there is a need to develop resources to facilitate the wider community.

ii. **Objectives:**

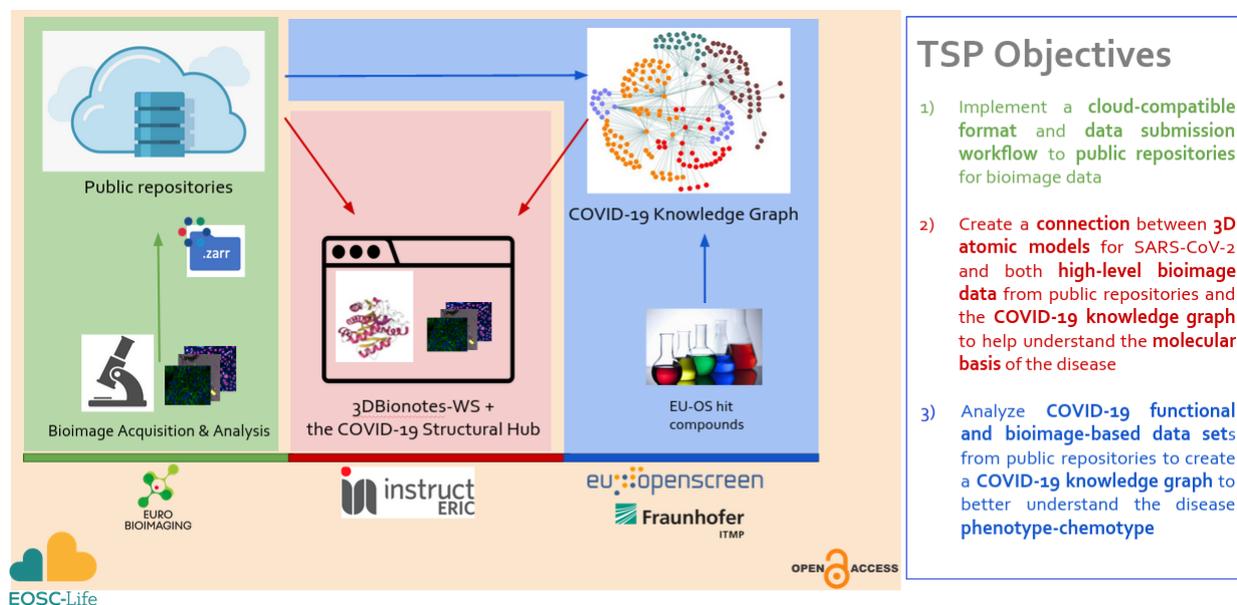


Figure 1: The scientific case in a nutshell for Open Imaging Data Sharing in EOSC

In the Test Science Project “Open Imaging Data Sharing in EOSC / COVID-19 as Demonstrator” Euro-BioImaging ERIC, INSTRUCT ERIC and EU-OPENSSCREEN ERIC will not only create useful resources in their domain but also aim to align these resources together. These resources are developed with FAIR principles and are cost and time effective.

Compliance to criteria developed by EOSC Future:

- **Eligibility:** This TSP comprises of 3 experts from 3 different domains of life sciences. Following the contributions made in the past, the experts will create useful resources to facilitate the wider scientific community.
- **Contribution to EOSC:** The resources built will be provided as horizontal services or softwares within EOSC catalogue of services. Wherever feasible, the aspects of integration, monitoring, accounting and helpdesk will be made available.
- **Quality:** The resources will be developed within the realm of FAIR data principles to ensure quality of the services/end products. Since the resources are being developed to tackle existing issues and fill in the existing caveats, the impact of the resources will be significant.
- **Relevance:** This TSP, although is based on COVID-19, will be relevant to researchers who are focusing in the areas of infectious diseases and virology/pathogens because of the fact that the resources are general workflows and should be applicable to any disease of interest.

TSP-C1) (Lead: Euro-BioImaging ERIC Bio-Hub at EMBL): Implement a community standard cloud-compatible open image data format and data submission workflow for high-throughput screening data

In COVID-19 related research projects - such as large-scale drug screens or SARS-CoV-2 X-ray and electron microscopy (EM) structural studies, efficient and open image data sharing among researchers and their communities around the globe is essential for fighting this pandemic. For example, our EU-OPENSOURCE partners have screened 5632 compounds for inhibition of viral induced cytotoxicity, using human cells and a SARS-CoV-2 isolate, to identify possible candidates for progression towards clinical studies against SARS-CoV-2 [1]. This EU-OPENSOURCE image data is now accessible on the Covid-19 Data Portal and the Image Data Resource (IDR), allowing access and visualization of the drug screens on SARS-CoV-2 infected human cells [2]. It is also available via the Covid-19 Data Portal via BioImage Archive [3].

However, the bioimaging field is plagued with myriad file formats with most of these not including enough metadata for the images to be reusable. High-throughput image-based screening is a case in point where complex metadata (e.g., spatial positions of images relative to each other) and intermediate analysis results often have to be preserved in sets of separate data files accompanying the images themselves. Re-use of such data in a cloud context is also limited by the requirement to move large amounts of data over the internet because the majority of current formats do not support remote extraction of fine-grained data subsets. Deposition of sets of image data screens by researchers in public repositories for re-use by others will, hence, tremendously benefit from the use of a community standard cloud-compatible open image data format. Development of such a standard, ome.zarr, file format has been initiated by the community itself with dedicated and continuing efforts by the Open Microscopy Environment (OME) team [4]. As part of this TSP, we further develop this standard for high-throughput screening data and use it as part of a workflow for automatic submission of COVID-19 screening data to the open and FAIR BioImage Archive [5].

To bring image data sharing to the next level and facilitate not only efficient but collaborative Covid-19 related research, we foresee the development of a modular, shareable, open, and FAIR workflow, based in Galaxy [6] for submitting data to the BioImage Archive in the cloud compatible, ome.zarr, image data format to be further developed in close collaboration with the OME team [7,8]. The Galaxy implementation will also enable us to incorporate existing image analysis workflows [9] and include intermediate analysis results before submission to the BioImage Archive. Additionally, the ome.zarr data format could be extended to all types of image data acquired by structural biology (including e.g. phase contrast imaging, x-ray imaging, EM) and light microscopy technologies (including high throughput screening such as the ones conducted by EU-OPENSOURCE). This will enormously facilitate integration and sharing of large-scale FAIR image data sets via EOSC across disciplines within the life sciences and beyond, thereby enabling and stimulating

collaborative image-based science. The image data standards developed in the project will be made interoperable with resources established as part of previous work by the EU-OPENSSCREEN team which established an ontology for working with COVID-19 data [10]. The ontology covers the roles of molecular and cellular entities in virus-host interactions and in the virus life cycle, as well as a wide spectrum of medical and epidemiological concepts [11]. Additionally, we have aligned with BY-COVID metadata standards to ensure FAIRness of the data [12]. Inclusion of additional image-based standards will further enhance the utility of this resource for widespread reuse.

TSP-C2) (Lead: Instruct ERIC): User-friendly access to multidimensional structural image data and quality-related metadata in EOSC space

TSP-C2 will concentrate on multidimensional image data in structural biology, with a special focus on cryoEM data together with fragment-based screening (currently primarily done with X-rays). There are three major motivations behind this selection; the first one is the very fast development of cryoEM and the high expectations in coming years for further technological advances and new applications in biomedical systems; the second one is the availability of map data as multidimensional images (besides 3D structural models) in public repositories like EMPIAR, and the third one, regarding fragment-based screening, is its direct connection with drug discovery. In all cases we will concentrate first on Covid19 data as development examples, although the goal is to be as general as possible.

With respect to cryoEM, and with well over half a thousand maps and models of different SARS-Cov2 and related specimens which are still rapidly growing, we are in a situation in which new quality-related metadata have to be derived so as to complement already available data. Indeed, not all data have the same quality, but this is not easily discernible for the users. In this TSP we will access map repositories such as EMDB, raw data in EMPIAR, structural model data in PDB and a whole range of biomedical annotations (in collaboration with ELIXIR) to extend our existing 3DBionotes application and provide the user with a very much needed “quality analysis interactive environment”. Note that former versions of 3DBionotes without this environment are already available in the EOSC MarketPlace [13] and the Covid19 data portal [14], demonstrating our good integration in EOSC space that support expectations for future widespread use of these new developments.

Regarding fragment-based screening, and in collaboration with EOSC-Life, EU-OPENSSCREEN, PaNOSC and Euro-Bioimaging, our work involves developing workflows to categorize the results from the different large-scale experiments publicly available in terms of target, compound and interaction region. We will then proceed by analyzing how best richer experimental structural information, such as event maps, could be presented. This is supported by developing user-friendly ways to access IDR and BioImage Archive so that information on the structure of the drug and the target could be presented together with key functional information building on the EU-OPENSSCREEN-IDR work presented in TSP-C1. Automated cheminformatics annotations for hit fragments and proposed elaborated compounds will be provided by EU-OPENSSCREEN via linking to the European Chemical Biology database resources [15, 16].

TSP-C3) (Lead: EU-Openscreen ERIC): Integrated phenotype:chemotype analysis of multi-omics open Covid-19 data

Clinical and drug discovery scientists are seeking to identify underlying therapeutic targets of phenotypic hits from COVID-19 repurposing. When integrated, image analysis (Cell Profiler), cheminformatics tools (KNIME) and Knowledge Graph (KG) approaches will allow generation of chemotypic-phenotypic relationships and aid in selection of individual compounds and combinations of compounds for clinical trials. Work recently reported by the EU-Openscreen team demonstrates the power of this approach by making use of public data sets that were available in the initial stages of the pandemic [17]. We extend and further validate this approach with additional data sets established within the reported project. Initially, a deeper integration of the data with complementary data sets in ChEMBL covering primary and secondary targets of the hit compounds is performed. Advanced work would involve integration with public proteomic and transcriptomic data sets associated with virus: host interactomes held in the COVID-19 data portal. This approach will facilitate systematic reverse-genomic analyses and aid identification of putative targets and druggable pathways. In addition to phenotypic data sets, EU-OPENSSCREEN generated target based data sets from the European Chemical Biology database (ECBD) and included in the phenotype:chemotype analyses. These reference data support the global analyses and cover key viral and host related target-mechanisms such as i) Inhibitors targeting the Spike: ACE2 interactions, ii) inhibitors of the host TMPRSS2 enzyme involved in processing of the Spike protein and iii) inhibitors of the virus Main Protease which is key to viral polypeptide processing and iv) inhibitors of the papain like protease (PL-Pro of NSP3) which is involved in the

evasion of the virus cellular immune responses. Data integration across target, phenotypic, and 'omics' data sets using knowledge graphs will facilitate the selection of appropriate synergistic combinations of compounds.

The work in the TSP "Open Image Data Sharing - Covid-19" is conducted in close collaboration with part "COVID-19 metadata findability and interoperability in EOSC", to guarantee the inclusion of minimal metadata standards as integral part of the new tools for image data sharing. Furthermore, Euro-Biolmaging, INSTRUMENT and EU-OPENSOURCE will reach out to and work together with the PaNOSC cluster project and XFEL to ensure integration with formats in use for dynamic biological data processes.

- **Implementation**

- i. Plan of work: (Tasks, Deliverables, Milestones, Responsible entities/persons, etc.)

- T1.1 - Implementation of a community standard cloud compatible open high-throughput screening image data format**

- The target format is ome.zarr [18] whose community-based development is coordinated by the OME team as part of the development of the next generation file formats for imaging [19]. T1.1 works with the community to finalize the currently incomplete specifications for high throughput screening data and metadata [20]. This will pave the way for the integration of developed standards into updated ontologies.

- Deliverable D1.1:** Report on the implementation of a standardized community-based cloud efficient file format for high throughput screening image data

- T1.2 – A workflow for submitting high-throughput screening data to the BiImage Archive in a cloud-efficient standard format.**

- The aim of this workflow is to facilitate the release of FAIR COVID-19 screening data to public repositories. The current workflow components include:

- Collection of primary image data and metadata in ome.zarr format implementing the community standard from T1.1.
 - Submission to the BiImage Archive and COVID-19 data portal.

- In the future versions the workflow will integrate intermediate analysis results, such as from CellProfiler or deep learning outputs, in the corresponding ome.zarr files.

- To support its development, the workflow is focused on three data resources: the already mentioned EU-OPENSOURCE drug screens on SARS-CoV-2 infected human cells (BioStudies accession #S-BIAD29) and a data set on SARS-CoV2 detection in cells [21], also available in BioStudies (accession # S-BIAD24). The third image data set covers profiling of drugs which inhibit PL-Pro mediated translocation of P-65 in epithelial cells Caco2. This SARS-CoV2 PL-Proviral protease antagonizes cellular ubiquitination and ISGylation which play important roles in the regulation of innate immune responses to viral infection.

- Given the large amount of data to process in screening applications, the workflow is implemented for the Galaxy platform [6] which ensures the workflow itself is FAIR and cloud compatible. In particular, we expect this workflow to be of use to screening data producers such as EU-OPENSOURCE and Euro-Biolmaging but equally to image data repositories for converting large amounts of existing data to support cloud-based re-use of public image data.

- Deliverable D1.2:** A Galaxy based workflow to submit ome.zarr files to BiImage Archive.

- T2.1. Easy access to structural and biomedical annotations databases and quality analysis metadata**

Development of rich and user-friendly environments to analyze multidimensional image data in structural biology, focusing on quality analysis for cryoEM and on fragment-based screening data complemented with functional information and medicinal chemistry insight and annotations.

Deliverable D2.1: New 3DBionotes version incorporating these developments and available in Covid 19 Data Portal and EOSC MarketPlace.

T3.1. Analysis pipeline for large scale COVID-19 functional and image based data sets for integration with complementary resources

Establish analysis workflows based on a cloud-deployed COVID-19 knowledge supergraph which integrate semantically consistent data and metadata from the repositories and workflows described in T2 and T1. The KG analysis workflow will integrate further public and RI-originated data resources and convert all knowledge graphs into OpenBEL [22], to capture and represent “cause-and-effect” relationships in a form which is fully interoperable with major image, pathway, and compound databases. Analysis of the COVID-19 super-graph will be used to identify mechanisms pertaining to selected targets where biochemical bioactivity data are available and improve the biological rationale when prioritizing repurposing candidates and combinations of compounds.

Deliverable D3.1: Extended cloud deployed COVID-19 Knowledge Graph competent for characterisation of relevant viral:host interactome pathways and the corresponding influence of compound perturbations revealed as phenotypes in image based readouts, inhibitory effects in functional assays and as ligand:target interactions from structural studies.

ii. Use of resources (data, services, etc.):

The TPS will build on, benefit from and support the following existing public data and tools resources in EOSC via EOSC-Life:

- IDR [7]
- BioImage Archive [5]
- Covid-19 Data Portal [14]
- 3DBionotes [23]
- COVID-19 Structural Hub, a dedicated section in 3DBionotes to study all macromolecular structures related to SARS-CoV-2 virus and COVID-19 [24]
- Structural Databases (PDB, EMDB, EMPIAR) [25]
- Bioactivity Databases (ChEMBL [26], ECBD [27]) and others.

iii. What are the demands of the TSP from EOSC Future platform: (e.g. Vertical and Horizontal composability services, HPC power, storage volume, etc.)

The TSP aims to facilitate use and reuse of high-throughput image, structural, and chemical screening datasets from around the globe with added value information in a user-friendly manner. This by definition raises the need for open access large scale cloud-based storage platforms. For the proposed TSP, the partners can make use of existing resources for HPC and cloud storage (EMBL-EBI Embassy cloud, EMBL HD 3Dcloud, de.NBI cloud) and use Covid19 data as a proof-of-concept implementation of seamless workflows that will provide a glimpse into the endless future possibilities within the EOSC ecosystem. The partners envision that this timely demonstration using Covid-19 data will help attract the much needed funding resources towards cloud storage resources.

iv. What the TSP brings to EOSC Future platform

- A community standard cloud compatible open image data format for FAIR sharing of large image datasets in the EOSC
- FAIR and modular workflow for submission of high-throughput image screens to BioImage Archive which will promote submission of large datasets to the open repository.
- User-oriented interactive environment to access structural and biomedical annotations databases and quality analysis metadata.
- Analysis workflows for cloud deployed COVID-19 data for integrated analysis and inference of phenotype:chemotype relations that will promote industrial vaccine and drug development.

- Reproducible workflow for generating knowledge graphs for infectious diseases and possible future threats within no time

- **Impact**

- i. Strategic:

TSP-C1: Researchers will benefit from and participate in Open Science by using a standardized community grown cloud compatible file format for large scale outputs from heterogeneous imaging modalities. This will in turn fuel a FAIR and modular workflow for (among others) high throughput microscopy image data and metadata from accession to analysis to deposition into public data repositories as part of the EOSC platform. The TSP-C1 activity will bring image data sharing to the next level for facilitating not only efficient but collaborative Covid-19 related research. The development of a workflow for submitting data to the BioImage Archive using a community standard cloud compatible open image data format will benefit all types of image data acquired by structural biology and light microscopy technologies. This will enormously facilitate integration and sharing of large-scale FAIR image data sets via EOSC across disciplines in the life sciences and thereby enable and stimulate collaborative image-based science.

TSP-C2: The development and publication of a new workflow for combining structural datasets in formats derived from various technology platforms will underpin annotation of 3D structural maps of biological molecules and thereby continuously increase their accuracy. Furthermore, the community of structural biologists will benefit from the publication of the new quality analysis interactive tools for cryoEM data, which increasingly plays a key role for biotech and pharmaceutical industries.

TSP-C3: The workflows established in TSP-C3 will provide cloud enabled resources for researchers to analyze orthogonal Covid-19 data and metadata sets in order to generate new hypotheses on therapeutic target selection. The KG workflows will allow for reuse by non-experts and external data holders and cover integration of public and proprietary data resources. Additionally, the KG workflow can be used to create KGs for other diseases of interest.

- ii. Scientific /User communities

Combined efforts of the three EOSC-life research infrastructures will foster interaction, collaboration and knowledge sharing among academic researchers in the life sciences and industrial vaccine and drug development scientists. Standardized workflows and access to open data will expedite current and future efforts for drug target identification crucial for handling the global pandemic.

- iii. Societal / Economic

The TSP “**Open Imaging Data Sharing in EOSC / COVID-19 as Demonstrator**” will foster collaborative and accelerated Open science which has proven to be an absolute requirement for tackling the Covid-19 pandemic. We will also find ourselves better prepared in the face of future challenges and be ready to work as a global community.

- iv. EU Policies

The implementation of our proposed TSP is deeply rooted in Open science and data sharing and hence will promote open innovation across the EOSC community and globally. It is also designed to improve our readiness towards future health crises by reaching out to all user communities with the promise of diminishing geographical biases.

- v. Other: None

- **Engagement Plan**

1. Target groups: Researchers/scientists in the EOSC community contributing actively to modern day life sciences.
2. SP key concept: The SP contributes in creating reproducible workflows and resources for data standardization, submission, integration, visualization and downstream analyses. All in all, the objective is to facilitate researchers by providing research products that can be readily used for research purposes.
The SP will be implemented by: 1) creating open-source tool for submission and standardization of bio-images, developing web-service for visualization of 3D structures of proteins and ligand-bound-protein

and generating a Knowledge Graph embedding phenotypic-chemotypic aspects of COVID-19. 2) Using workflows to analyse the data; 3) Integrating its resources with EOSC Core services and potentially Horizontal services available; 4) Engaging the relevant scientific communities.

Users will be able to: 1) Readily use aforementioned resources; 2) Provide managerial suggestions to relevant authorities; 3) Build on the existing infrastructure to address more complex scientific questions.

3. Dissemination measures

- a. EuroBiomaging resource
 - i. BatchConvert available on GitHub [28], Zenodo [29] and WorkflowHub [30]
 - ii. Registered in the EOSC marketplace
 - iii. Contributed to the recent NGFF publication titled 'OME-Zarr: a cloud-optimized bioimaging file format with international community support' [31]
 - iv. Training:
 - Defragmentation Training School - NEUBIAS Academy & EOSC-Life 2022 [32]
 - Defragmentation Training School 2 - NEUBIAS Academy & EOSC-Life 2023 [33]
 - v. Online seminars [34]
 - vi. Hackathons and workshops [35, 36]
 - vii. Demo and poster presentation at EOSC symposium 2022 ("Open Imaging Data Sharing in EOSC/COVID-19 as Demonstrator")

- b. Instruct ERIC resource
 - i. Journal titled "3DBionotes COVID-19 edition" [37]
 - ii. Onboarded in EOSC marketplace
 - 3Dbionotes [13]
 - COVID19-Structural Hub [38]
 - iii. Poster presentation at ECCB2022, ELIXIR (3D-Bioinfo Community) and EOSC Symposium 2022 titled
 - "3DBionotes COVID-19 structural hub: a central resource for validation information and refined models" [39]
 - "Adding High-Content Screening annotations to 3DBionotes-WS" [40]
 - "Imaging Data: Adding a new type of annotations to 3DBionotes-WS" [41]
 - iv. Web service
 - 3Dbionotes [23]
 - COVID19-Structural Hub [24]
 - v. Link to other projects:
 - BY-COVID
 - IMPaCT-Data (Spanish Initiative)
 - vi. Demo and poster presentation at EOSC Symposium 2022 ("Open Imaging Data Sharing in EOSC/COVID-19 as Demonstrator")

- c. EU-Openscreen/Fraunhofer ITMP resources
 - i. GitHub repository:
 - Mpox [42]
 - COVID-19 [43]
 - ii. Research article titled 'Mpox Knowledge Graph: a comprehensive representation embedding chemical entities and associated biology of Mpox' published in *Bioinformatics Advances* [44]
 - iii. Similar KG applications in a journal paper titled 'Comprehensive Fragment Screening of the SARS-CoV-2 Proteome Explores Novel Chemical Space for Drug Development' [45]
 - iv. BioModels:
 - Mpox [46]

- COVID-19 [47]
- v. Poster/demo presentation at EOSC symposium 2022 (“Open Imaging Data Sharing in EOSC/COVID-19 as Demonstrator”)
- vi. Available at COVID-19 Data Portal
- vii. Available in EOSC catalogue of services
 - Mpox [48]
 - COVID-19 [49]
- viii. Presented a poster on a workshop organized by Leopoldina institute, Berlin on Infectious Diseases Beyond COVID-19 [50]
- ix. Demo and poster presentation at EOSC Symposium 2022 (“Open Imaging Data Sharing in EOSC / COVID-19 as Demonstrator”)

Funding

This SP is part of the EOSC Future Project WP6.3 and is funded by European Union Commission (Grant Number: 101017536).

References

1. Ellinger B, Bojkova D, Zaliani A, et al. A SARS-CoV-2 cytopathicity dataset generated by high-content screening of a large drug repurposing collection. *Sci Data*. 2021;8(1):70. Published 2021 Feb 26. doi:10.1038/s41597-021-00848-4
2. <http://idr.openmicroscopy.org/mapr/compound/?value=Remdesivir>
3. <https://www.ebi.ac.uk/biostudies/studies/S-BIAD29>
4. <https://ngff.openmicroscopy.org/latest>
5. <https://www.ebi.ac.uk/bioimage-archive>
6. <https://usegalaxy.org>
7. <https://idr.openmicroscopy.org/>
8. <https://bit.ly/eoss-covid-zarr>
9. <https://training.galaxyproject.org/training-material/topics/imaging/tutorials/tutorial-CP/tutorial.html>
10. Sargsyan A, Kodamullil AT, Baksi S, et al. The COVID-19 Ontology. *Bioinformatics*. 2021;36(24):5703-5705. doi:10.1093/bioinformatics/btaa1057
11. <https://bioportal.bioontology.org/ontologies/COVID-19>
12. <https://zenodo.org/record/6885016#.Y3-fv98xmUk>
13. <https://marketplace.eosc-portal.eu/services/eosc.cnb-csic.3dbionotes-ws>
14. <https://www.covid19dataportal.org/related-resources>
15. <https://instruct-eric.eu/platform-type/electron-microscopy>
16. <https://instruct-eric.eu/platform/xchem-fragment-screening-diamond-light-source-harwell-uk/>
17. Schultz B, Zaliani A, Ebeling C, et al. A method for the rational selection of drug repurposing candidates from multimodal knowledge harmonization. *Sci Rep*. 2021;11(1):11049. Published 2021 May 26. doi:10.1038/s41598-021-90296-2
18. <https://ngff.openmicroscopy.org/0.1/#intro>
19. Moore J, Allan C, Besson S, et al. OME-NGFF: a next-generation file format for expanding bioimaging data-access strategies. *Nat Methods*. 2021;18(12):1496-1498. doi:10.1038/s41592-021-01326-w
20. <https://ngff.openmicroscopy.org/0.1/#hcs-layout>
21. Pape C, Remme R, Wolny A, et al. Microscopy-based assay for semi-quantitative detection of SARS-CoV-2 specific antibodies in human sera: A semi-quantitative, high throughput, microscopy-based assay expands existing approaches to measure SARS-CoV-2 specific antibody levels in human sera. *Bioessays*. 2021;43(3):e2000257. doi:10.1002/bies.202000257
22. <http://www.openbel.org/>
23. <http://3dbionotes-ws.cnb.csic.es/ws>
24. <https://3dbionotes.cnb.csic.es/ws/covid19>
25. <https://www.ebi.ac.uk/pdbe/emdb/empiar/>
26. Davies M, Nowotka M, Papadatos G, et al. ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Res*. 2015;43(W1):W612-W620. doi:10.1093/nar/gkv352
27. <https://ecbd.eu/>
28. <https://github.com/Euro-Biolmaging/BatchConvert>
29. <https://zenodo.org/record/7955974>
30. <https://workflowhub.eu/workflows/453>
31. Moore J, Basurto-Lozada D, Besson S, et al. OME-Zarr: a cloud-optimized bioimaging file format with international community support. *Histochem Cell Biol*. 2023;10.1007/s00418-023-02209-1. doi:10.1007/s00418-023-02209-1
32. <http://eubias.org/NEUBIAS/training-schools/neubias-academy-home/defragmentation-training-school-2022/>
33. <https://eubias.org/NEUBIAS/training-schools/neubias-academy-home/defragmentation-training-school-2023/>
34. <https://www.youtube.com/watch?v=DeCWV274I0c&list=PLW-oxncaXRqU4XqduJzwFHvWLF06PvdVm&index=13&t=2s>

35. <https://globalbioimaging.org/international-training-courses/bioimage-hackathon-01>
36. <https://galaxyproject.org/events/2022-10-egd>
37. Macias JR, Sanchez-Garcia R, Conesa P, et al. 3DBionotes COVID-19 edition. *Bioinformatics*. 2021;37(22):4258-4260. doi:10.1093/bioinformatics/btab397
38. https://marketplace.eosc-portal.eu/services/eosc.cnb-csic.covid-19_structural_hub
39. <https://doi.org/10.7490/f1000research.1119196.13>
40. <https://doi.org/10.7490/f1000research.1119240.1>
41. <https://doi.org/10.7490/f1000research.1119239.1>
42. <https://github.com/Fraunhofer-ITMP/mpox-kg>
43. <https://github.com/Fraunhofer-ITMP/BY-COVID-KG>
44. Karki R, Gadiya Y, Zaliani A, Gribbon P. MpoX Knowledge Graph: a comprehensive representation embedding chemical entities and associated biology of MpoX. *Bioinform Adv*. 2023;3(1):vbad045. Published 2023 Apr 3. doi:10.1093/bioadv/vbad045
45. Berg H, Wirtz Martin MA, Altincekic N, et al. Comprehensive Fragment Screening of the SARS-CoV-2 Proteome Explores Novel Chemical Space for Drug Development. *Angew Chem Int Ed Engl*. 2022;61(46):e202205858. doi:10.1002/anie.202205858
46. <https://www.ebi.ac.uk/biomodels/MODEL2208040001>
47. <https://www.ebi.ac.uk/biomodels/MODEL2209020001>
48. <https://explore.eosc-portal.eu/search/software?pid=10.5281%2Fzenodo.7351393>
49. <https://explore.eosc-portal.eu/search/software?pid=10.5281%2Fzenodo.7351221>
50. <https://doi.org/10.5281/zenodo.7990992>