An Open Science strategy for ENAC

Katrin Beyer[1], Gilles Dubochet[2]

[1]*katrin.beyer@epfl.ch*, Associate Dean ENAC for Digitalization and Open Science, ENAC, EPFL
[2]*gilles.dubochet@epfl.ch*, Scientific Advisor on Open Science, EPFL

April 29, 2021[*]

1   Purpose of this document

The purpose of this document is to outline a strategy for the promotion of Open Science (OS) at ENAC. It serves as input for the direction of ENAC and as basis for ENAC services to design and implement measures. The document

- defines the term "Open Science" and describes the relevance of OS to ENAC's strategic goals;

- sets the scene for OS at ENAC by describing briefly the OS landscape at EPFL and in Switzerland;

- states why the adoption of OS principles is important for ENAC;

- outlines five OS goals for ENAC and propose first steps for their implementation.

2   Definition of Open Science

Open Science (OS) is used as an umbrella term that covers a wide range of initiatives that tap into the potential of digital technologies to improve scholarship. OS is defined by the European Commission as follows (EC 20): "Open science is about the way research is carried out, disseminated, deployed and transformed by digital tools and networks. It relies on the combined effects of technological development and cultural change towards collaboration and openness in research. By providing unlimited, barrier free, open access to research outputs, Open science makes scientific processes more efficient, transparent and

---

[*]At this date, the status of this document is the following: This proposal was discussed and approved in DirENAC.

responsive to societal challenges. It offers new tools for scientific collaboration, experiments and analysis and makes scientific knowledge more easily accessible."

The term "Open" is used here as defined by the Open Knowledge Foundation "Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)" (OKF15). This has implications on how research results are shared and which research results are shared. Six types of research outputs are differentiated here:

- Open access to publications (OA);

- Open research data (ORD);

- Open software;

- An open and reproducible description of the research pipeline;

- Open hardware;

- Open educational resources (OER).

In the following we refer to the ensemble of these OS outputs as OS artefacts.

3   Setting the scene

This section gives an overview on OS initiatives at EPFL, within the ETH domain and Switzerland. OS at the European and beyond are not discussed. Many of these initiatives are evolving very quickly (e.g. the ORD program of the ETH domain and ORD and OA measures by swissuniversities). The status reported here is of January 2021.

3.1   OS at EPFL

The Direction of EPFL supports OS as the "future of research", which drives scientific excellence and increases research impact. The degree and quality of open sharing of research outputs is expected to be an important measure of academic performance in the future. It is also a response to scientific funders and publishers, who are increasingly requiring adherence with best practices in research and considerations with regard to scientific integrity and social responsibility.

The Open Science Strategic Committee is a faculty-led committee, mandated by the Direction to make recommendations for Open Science policy and propose a strategic roadmap. The OS strategy of EPFL aims to find a balance between two approaches:

- A demand-driven approach, which supports researchers with an outstanding OS vision and research agenda. The EPFL Open Science Fund (see below) supported OS projects of EPFL researchers, who set new OS standards in their field.

- A supply-driven approach, which aims to develop services, infrastructure and practices that lower the barriers to engage in OS practices.

EPFL-wide actions to support the adoption of OS principles include the following:

- EPFL Open Science Fund: a 3 million program with two calls in 2018 and 2019, which supported the project of 18 "Open Science Champions" at EPFL (EPF20d).

- EPFL Open Access policy (EPF20b): it provides non-binding guidance (recommendations) for OA, notably for the use of Infoscience.

- Gold Open Access financial support (EPF20c): it supports the publication of Gold OA publications in OA journals.

- OER at EPFL: The MOOC initiative at EPFL (EPF20a) is a program for online courses to which already many ENAC researchers contribute. Switchtube (SWI20b) allows sharing recordings of lectures. Furthermore, a call on open education resources is currently in the early stages of discussion at the EPFL-level. Such call would very likely not focus on classical MOOCs but on sharing various components of a course (lecture recordings, slides, exercises, ...) under CC-BY licenses.

3.2   OS in the ETH domain

The ETH Board's interest in OS is focused on Open Research Data (ORD). In May 2020, it adopted a position in support of ORD (ETH20). A series of support measures for ORD to be implemented in the ETH domain during the period 2021 to 2024 is currently planned. The measures will likely include the following:

- Coordination between ETH institutions and funding for the deployment of ORD infrastructure and services;

- Grants to support researchers in developing ORD practices in their communities;

- Development of training material and legal guidance;

- Reflections on improving career options for ORD professionals.

3.3   OS at the national level

Swissuniversities is mandated by the SBFI to coordinate the development of OS at the national level. It is treated separately for OA and ORD. OA is more advanced, and a strategy (SS17) as well as an action plan (Swi18) and funding program (Swi20a) are defined. For ORD, only a draft strategy has been circulated for consultation.

4 Why fostering OS at ENAC?

At ENAC we support the adoption of OS principles in research, teaching and innovation and outline in the following the objectives behind this choice.

Research: We aim at making research openly accessible and reproducible in order to

- increase ENAC's visibility and impact on the research community;

- accelerate research processes and facilitate inter-, trans- and multi-disciplinary and collaborative research, which is a strategic goal of ENAC (Bin20);

- make our young researchers competitive: OS will be the new norm and therefore it is essential for our young researchers to engage in it.

With the exception of OA, for which an institutional policy exist, the adoption of OS principles is a decision of the PI. OS principles are not imposed top-down but need to be endorsed by the PIs on a project-by-project basis. ENAC's role is to create an environment in which the adoption of OS principles is facilitated and valued.

Teaching: We aim at educating students of all levels on OS principles and opportunities:

- so that ENAC-educated professionals in the private or public sector have the know-how to accelerate open knowledge transfer from academia to practice;

- in order to make good use of OER during their studies and for life long learning;

- to be citizens who are aware about good principles for research funded with public money.

Innovation: ENAC researchers transfer technology to the society in many ways (Bar19). The following tech transfer paths build on OS artefacts:

- open software and data platforms developed at ENAC, which are also used in the public and private sector (for examples see Section 6.6 in (Bar19));

- knowledge transfer in the form of policy advice to governmental bodies;

- knowledge transfer through contributions to civil engineering standards.

Until today, there are no funding instruments that are dedicated to developing start-ups based on OS artefacts. As a result, ENAC start-ups were so far not built on OS artefacts. We aim at strengthening the knowledge transfer to the private and public sector on the basis of OS artefacts:

- to increase the impact on ENAC's research on the private and public sector and to accelerate the tech transfer process;

- to take societal responsibility by sharing results openly that are of societal importance (e.g. sustainability challenges and climate change);

- explore new avenues of tech transfer on the basis of OS artefacts (e.g. within the association FUSTIC (ENA21)).

5 OS goals for ENAC

The following OS goals for ENAC are proposed:

- to make visible and communicate actively OS artefacts by ENAC researchers,

- to make ENAC compliant with EPFL's OA policy;

- to include OS in Bachelor, Master and PhD education;

- to provide ENAC specific services and infrastructure that support the adoption of ORD principles;

- to support tech transfer on the basis of OS artefacts.

These goals and their implementation steps are outlined in more detail in the following.

5.1 Goal 1: Make visible and communicate actively OS open outputs by ENAC researchers

Engaging in OS means sharing research artefacts beyond the publication as described in Chapter 2 (Eur20a). It is ENAC's goal to acknowledge and showcase important OS contributions. ENAC will therefore collect information on all open outputs by ENAC researchers and help ENAC researchers to make these outputs visible. In addition, ENAC will evaluate the need for services to support OS artefacts (e.g. ORD, open software) over a defined lifecycle. ENAC will adopt the following implementation strategy to reach this goal:

- Include in the activity reports OS artefacts of all kind (OA, ORD, open software, OER, ...);

- Showcase OS artefacts of the ENAC community and their impact

  - Develop together with ENAC communication a way to display the overall ENAC OS output information through the ENAC webpage and highlight outstanding OS contributions (including MOOCs developed by ENAC researchers);

– Develop solutions to showcase specific ORD artifacts simply and effectively through web applications.

- Evaluate the need for ENAC-IT4Research services for supporting OS artefacts over their lifecycle (in particular ORD and open software); Develop these services based on a sustainable cost model.

5.2 Goal 2: Make all ENAC publications OA by 2024

The Swiss OA Strategy (SS17) has the objective that by 2024 the entire[1] scholarly publication activity in Switzerland should be OA. As a member of swissuniversities, which adopted the strategy, EPFL and in turn ENAC should work towards this objective.

As with the adoption of OS principles in general, the choice of the publication approach (and therefore the journal) is the PI's choice and this list only for orientation among the various OA approaches. The freedom in choice is important as in many fields the most prestigious journals are unfortunately not journals that shine with regard to "openness". The various publication approaches can be ranked according to "openness" as follows:

- The best approach is to publish in such a way that the article can be read and reused freely (CC-BY licence or equivalent) and immediately. This corresponds to Gold OA, Green OA with no embargo period, or Platinum OA.

- The second best approach corresponds to Green OA as above, but with a (short) embargo period.

- The third approach is a publication that can be read freely but reuse is associated with charges. Some publishers claiming to do Gold OA do that, but this is not considered OA and is not compliant with SNSF or EU rules.

For ENAC to reach this goal, the following should take place:

- Building on Goal 1, evaluate in the ENAC activity report what proportion of ENAC publications can be made OA by relying on the above mechanisms.

- Calculate the additional financial cost (e.g. APC not covered by other measure, support to Diamond OA platforms) that would be required to make all possible publications OA.

- Inform and support researchers in using OA measures available to individual researchers and communities (see Appendix B).

- Support individual researchers to improve their own OA score by making their most important publications OA retrospectively. Use for example SNSF's OA checking tool (checks publications from 2015 onwards).

---

[1]In the implementation of the strategy, this is expected to be around 80% of all publication.

5.3 Goal 3: OS in Education

We expect that OS will influence many aspects of our students' education as well as profession, be it inside or outside academia. ENAC's education should therefore prepare students to contribute to and take advantage of OS during the various stages of their career:

- ENAC students of all levels should be made aware of the potential impact of OS to society at large, and the promises that are related to OS (e.g. effective use of public money, open access to knowledge created through public money, ...)

- ENAC students of all levels should know about the large range of open educational resources (OER). This knowledge will also be useful in life-long learning, be it a formal continuing education program or informal learning by taking single online courses.

- ENAC students who carry out research work should know about the mechanisms of OS and master key tools that foster OS practices. This will help them conduct excellent research and advance their career using the tools and practices of open science.

- ENAC students who will work in the private, public, or non-profit sector, should be prepared to work with open data, software, publications and other open research outputs as part of their professional practice. This is for three reasons:

  1. For OS to be an accelerator of knowledge transfer from academia to practice, OS practices need to be understood on both sides, i.e., also on the professional practice side.

  2. As AI will increasingly be used to solve standard problems, it is likely that applied research, in a way that integrates OS artefacts, will become more and more important in professional practice.

  3. Some professionals, e.g., in federal offices, might be involved in allocating resources to research projects. Knowledge on OS principles are necessary to design research contracts in a way that best serves society.

To achieve these goals, OS aspects should be included in the Bachelor, Master and PhD curricula. A first list of topics to be covered is proposed in Appendix C. To discuss within ENAC how OS can be anchored within the various curricula, the following steps are proposed:

- Discuss with the Associate Dean of Education and the section directors

  - which OS principles should be taught within ENAC's curricula,
  - in which course modules they could be best integrated or design courses that address such topics,

7

- how the use of OER (including OS literature and MOOCs developed by ENAC faculty) can be promoted within ENAC courses.

- Encourage ENAC's PhD students to take courses on OS that are offered by EPFL services;

5.4 Goal 4: Provide ENAC specific services that support the adoption of ORD principles

ENAC researchers who are looking to improve their research data management (RDM) practices and to open their data, should be able to build on domain-relevant, easy-to-use yet powerful services and infrastructure for their entire data lifecycle (see Appendix A). ENAC's goal is to lower the technical barriers to good RDM and ORD practices, by supporting researchers in using existing services and working towards making them interoperable. To implement this goal, the following steps are proposed:

- Contribute to the development of OS infrastructure and services within EPFL:

  - Accompagny from the beginning the discussions around data infrastructure and software needs within EPFL to encourage the deployment of solutions that serve ENAC researchers well;

  - Wherever possible, do not operate infrastructure ourselves but ensure access to shared facilities and give input to the design of new EPFL infrastructure so that they are useful to ENAC researchers.

- Develop ENAC-IT4Research services that

  - develop services that make complex data sets visable, discoverable and exploitable (see also Goal 1).

  - support the adoption of open software, open research data and reproducible workflows;

  - support data valorisation and exploitation through ML techniques;

  - explore a strategic partnership with SDSC to provide ENAC researcher the infrastructure and data management tools needed for ORD and get access to machine learning competences;

- Goals with regard to clusters/centers:

  - accompany from the beginning the discussions around data infrastructure and software needs;

  - ensure that data management structures are adopted that are suitable for trans-disciplinary projects and collaborative projects;

  - develop approaches to link various data sets and make these complex data sets discoverable (see also Goal 1).

5.5 Goal 5: Support tech transfer on the basis of OS artefacts

It is expected that in particular the clusters and CLIMACT generate significant OS artefacts of high societal relevance. It is ENAC's goal to increase and accelerate the tech transfer on the basis of these OS artefacts and to communicate actively on this tech transfer. The following implementation strategy is envisaged:

- Communicate actively on ENAC's OS artefacts that are used in the private and public sector;

- Develop with the clusters an approach on how to promote OS in collaborations with the private and public sector;

- Investigate founding associations to ensure the long-term support of open-source software.

6 Strategic foci per year and goal

Table 1: Proposed implementation timeline for OS at ENAC strategy

	2020	2021	2022	2023
Focus	Focus on development of OS strategy and strategy for ENAC-IT4Research	Focus on making ENAC's OS artefacts visible and offering services through ENAC-IT4Research	Focus on data management of clusters and creating visibility for ENAC's data	Focus on data valorisation of clusters & CLIMACT and OS in education
Goal 1: Recognise all open outputs	• Housekeeping OA (explore the impact of transformative agreements on OA to ENAC's publication).	• Collect information on OS artefacts in activity report of laboratories; • Work in ENAC communications on setting up a webpage that is dedicated to OS artefacts at ENAC; develop a pipeline how this webpage can be continuously fed (e.g. input through activity reports of laboratories; during the year: form to be completed by ENAC researchers); • Evaluate the need for ENAC-IT4Research services for supporting OS artefacts over their lifecycle (in particular ORD and open software).	• Showcasing: Improve pipeline that is continuously fed through automated pipelines; write a news article on OS outputs of ENAC in general / a particular OS output; • Activity report: include automated data collection, by interfacing with relevant databases and platforms such as Zenodo, Github, Renku, ... should be considered; • Develop (based on analysis of needs from previous year) a sustainable cost model ENAC-IT4Research services for supporting OS artefacts over their lifecycle and implement it.	• Evaluate development of OS outputs over last three years.
Goal 2: Adopt OA strategy		• Evaluate OA situation.	• Design measures to improve OA policy compliancy;	• Design measures to improve OA policy compliancy if necessary and evaluate ENAC's performance wrt OA;

Table 1: Proposed implementation timeline for OS at ENAC strategy

	2020	2021	2022	2023
Goal 3: OS in education		• Launch discussion on OS of education with Assoc. Dean for Education and Section directors.	• If there is an agreement that OS should be included in ENAC education: Implement first modules on Bachelor and Master level.	• If there is an agreement that OS should be included in ENAC education: Continue implementing modules on Bachelor and Master level.
Goal 4: Support ORD	• Develop a strategy for data scientist position and for offering ENAC IT4research; • Define the corner stones of a strategic partnership between ENAC and SDSC; • Develop ENAC case studies with Renku and investigate whether there is an interest by the ENAC community to use Renku.	• Implement ENAC IT4research strategy and carry out a first few projects; • Explore potential of a strategic partnership with SDSC and take a decision on whether this should be pursued further; • Accompagny post-doc cluster grants through the data cycle; collect lessons learnt for big ENAC cluster projects; • Advise on RDM plans for big ENAC cluster projects and CLIMACT; • Support clusters and CLIMACT in data aspects; • Encourage and coordinate participation in ORD calls of ETH domain.	• Obtain feedback on ENAC-IT4Researchservices and adjust offer if necessary; • Scale up ENAC-IT4Researchservices; • Develop (potentially with SDSC) ways to showcase, explore and exploit the complex datasets of ENAC (in particular datasets generated within clusters and CLIMACT); • Accompagny data management of big ENAC cluster projects; • Encourage and coordinate participation in ORD calls of swissuniversities;	• Support ENAC clusters and CLIMACT wrt data valorisation and visualisation; • Accompagny data management of big ENAC cluster projects; • Encourage and coordinate participation in ORD calls of swissuniversities.

Table 1: Proposed implementation timeline for OS at ENAC strategy

	2020	2021	2022	2023
Goal 5: Support tech transfer of OS artefacts	• Communicate actively on ENAC's OS artefacts that are used in the private and public sector.	• Investigate for a specific case whether the foundation of an association could be helpful to ensure the long-term support of open-source software.		• Develop with the clusters an approach on how to promote OS in collaborations with the private and public sector.

References

[Bar19] BARRY, A: *ENAC self-assessment report Audit 2019 – Volume 2019.* 2019

[Bin20] BINDER, Claudia: *ENAC Strategy 2020-2023.* `https://www.epfl.ch/schools/enac/wp-content/uploads/2020/12/ENAC_GA_PLENARY_2020.pdf`. Version: 2020

[EC 20] EC EUROPEAN COMMISSION: *Open Science - Shaping Europe's digital future.* `https://ec.europa.eu/digital-single-market/en/open-science`. Version: 2020

[ENA21] ENAC: *Association Future Sustainable Territories, Infrastructure and Cities (FUSTIC).* `https://www.epfl.ch/schools/enac/homepage-enac/sustainability-challenges/fustic_en/`. Version: 2021

[EPF20a] EPFL CENTER FOR DIGITAL EDUCATION: *MOOCs production.* `https://www.epfl.ch/education/educational-initiatives/cede/moocproduction/production/`. Version: 2020

[EPF20b] EPFL LEX 3.5.1: *EPFL OA policy.* `https://www.epfl.ch/about/overview/wp-content/uploads/2019/09/LEX-3.5.1_EN.pdf`. Version: 2020

[EPF20c] EPFL LIBRARY: *Gold Open Access financial support.* `https://www.epfl.ch/campus/library/services/services-researchers/gold-open-access-financiel-support/`. Version: 2020

[EPF20d] EPFL OPEN SCIENCE INITIATIVE: *EPFL Open Science Champions.* `https://www.epfl.ch/research/open-science/champions/`. Version: 2020

[ETH20] ETH BOARD: *Open Research Data Position of the ETH Domain.* `https://www.ethrat.ch/sites/default/files/ORD_Position_ETH_Domain.pdf`. Version: 2020

[Eur20a] EUROPEAN COMMISSION: *Next-generation metrics: Responsible metrics and evaluation for open science.* `https://ec.europa.eu/research/openscience/pdf/report.pdf`. Version: 2020

[Eur20b] EUROPEAN COMMISSION: *Open Research Europe.* `https://open-research-europe.ec.europa.eu`. Version: 2020

[OKF15] OKF OPEN KNOWLEDGE FOUNDATION: *The Open Definition.* `https://opendefinition.org/`. Version: 2015

[SS17] SWISSUNIVERSITIES ; SWISS NATIONAL SCIENCE FOUNDA-
TION: *Swiss National Strategy on Open Access.* `https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/Open_Access/Open_Access_strategy_final_e.pdf`. Version: 2017

[Swi18] SWISSUNIVERSITIES: *Stratégie nationale suisse sur l'Open Access - Plan d'action.* `https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/Open_Access/Plan_d_action-f.pdf`. Version: 2018

[Swi20a] SWISSUNIVERSITIES: *Open Science Program I (2021-2024): Phase A - Open Access Implementation Plan.* `https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Organisation/SUK-P/SUK_P-2/PgB_OpenScience_-_Implementation_Phase_A_2021-2024_published.pdf`. Version: 2020

[SWI20b] SWITCHTUBE: *Academic video sharing.* `https://tube.switch.ch/`. Version: 2020

A   Data management lifecycle needs

- Data management platforms for relevant types of data, input standards and metadata (e.g. imaging data, source code, etc.);

- Connection of data, source code and computing environments to facilitate data reuse and reproducible (e.g. renku);

- Specialised computing infrastructure for big data and high performance computing, when needed;

- Long-term archival services;

- Open/sharing platforms, that make data visible, discoverable, and provide FAIR services (findable, accessible, indexable and reusable);

- Visualisation and advanced analytic services, including using ML.

- Laboratory information management systems and other data input services;

- Generation of data for publication / publisher standards;

B   Available support measures for OA publishing

To adopt OA publishing strategy, various measures are available to the individual researcher and communities:

- Publishing in OA journals that do not depend on commercial publishers, are community-managed and funded by grants or institutions (Eur20b) (sometimes called Diamond or Platinum OA);

- Open Science funding schemes can be used to launch or support the development of such community-driven OA publication venues;

- When publishing in traditional OA (or hybrid) journal, article processing charges (APCs) can be included in funding requests to SNSF, EU funds and many other schemes (Gold OA);

- A number of OA article publication rights are available through Swiss OA "big deals" with publishers (including Springer Nature, Elsevier), currently on a "first-serve" basis;

- If insufficient funding was requested through projects, APC costs may sometimes be reimbursed through EPFL library;

- In most other cases, full texts can be deposited on Infoscience with an appropriate embargo period (Green OA), in line with the EPFL OA policy (EPF20b).

C   First proposal for OS curricula for Bachelor, Master and PhD education

Bachelor education:

- OS as a driver for societal change and responsibility,

- know about open education resources and other OS outputs,

- know the various licenses that are used in academic publishing,

- know about plagiarism and the difference to reusing of content within OS,

- data ethics in particular with regard to AI.

Master education (for student without research orientation):

- ORD: FAIR principles, where to find ORD and exploit it,

- Metadata of ORD,

- Open software: how to find it and reuse it;

Table 2: Summary table of OS curricula for Bachelor, Master and PhD education; Notation of skill level: 1 = knowledge of, 2 = ability to apply/ design research project to embrace these principles

Topic	BSc	MSc	Research-MSc & PhD
OS and society			
OS as a driver for societal change and responsibility	1		
OS outputs	1	1	2
OER	1		
OA			
Licenses that are used in academic publishing, publishing landscape	1	1	2
Plagiarism and the difference to reusing of content within OS	2		
ORD			
FAIR principles		1	2
Metadata		1	2
Data ethics in particular with regard to AI	1		2
Open software & reproducible workflows			
Code carpentry to foster reuse	1,2	1,2	2
Fully reproducible workflows		1	2
Field-specific standards wrt OS			1,2

- Principles of good code carpentry,

- Reproducibility crisis and the importance of a reproducible workflow,

- Community science.

Master student carrying out research projects and PhD education:

- OA: Why and how, costs of academic publishing, overview on publishing landscape,

- ORD: good data management practices,

- Metadata of ORD: good standards for its implementation,

- Practice on good code carpentry,

- Implementation of a fully reproducible workflow,

- Field-specific standards with regard to OS.