



HAL
open science

A unified path from data to publications: Three French infrastructures in the COMMONS project

Nicolas Larrousse, Olivier Baude, Dominique Roux, Marie Pellen

► To cite this version:

Nicolas Larrousse, Olivier Baude, Dominique Roux, Marie Pellen. A unified path from data to publications: Three French infrastructures in the COMMONS project. DH2022, The University of Tokyo, Japan; Alliance of Digital Humanities Organizations (ADHO), Jul 2022, Tokyo (Japan), Japan. <https://dh2022.dhii.asia/dh2022bookofabsts.pdf>. halshs-03881307v1

HAL Id: halshs-03881307

<https://shs.hal.science/halshs-03881307v1>

Submitted on 9 Dec 2022 (v1), last revised 23 Dec 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A unified path from data to publications:

Three French infrastructures in the COMMONS project

Authors:

Nicolas Larrousse (Huma-Num, CNRS, France) - ORCID 0000-0002-4968-797X

Marie Pellen (OpenEdition, CNRS, France)

Dominique Roux (METOPES, CNRS & Université de Caen, France)

Olivier Baude (Huma-Num, CNRS, France) - ORCID 0000-0002-8627-5229

In order to support research projects in SSH (Social Sciences and Humanities), France has developed centralized infrastructures which are listed on the French national roadmap for infrastructures: this means that they are financed over the long term.

Three of these infrastructures are especially complementary:

Huma-Numⁱ provides a set of platforms, tools and support for the processing, conservation, dissemination and long-term preservation of digital research. OpenEditionⁱⁱ provides a set of four scholarly communication platforms and Metopesⁱⁱⁱ provides a set of tools and methods, built according to single source publishing model, enabling the creation of natively structured editorialized contents.

Even though these infrastructures provide all the services required throughout the life cycle of research, their different environments are separate, making it necessary to switch from one platform to another, for instance to carry out the process of linking data and publications.

In the context of the increasing development of Open Science policies, both at national and European levels, these three main infrastructures joined their forces to build a project to address these challenges. The main focus of the COMMONS^{iv} project (COnsortium of Mutualised Means for Open data & Services for SSH) is of course to develop bridges between the three infrastructures articulated around the link between data and publications but it also covers various other related aspects.

A key step to achieve the project goals is to develop communication modules between the different platforms which represent the foundation on which the other achievements will be based. For instance, one of the most important components to build this environment is to use a common authentication system which allows users to connect to any platform with the same credentials. This requires technical adaptations but also substantial preliminary work to ensure coherence between our lists of respective users.

Similarly, to link data and publications, it will be necessary to establish bi-directional links between the data repository provided by Huma-Num and the scholarly communication platforms from OpenEdition associated with tools from Metopes. Interoperability modules will be developed to implement these links associated with a specific workflow aiming to maintain consistency of information between the different

platforms. For example, one of the key challenges will be to take into account the synchronization aspects when establishing the link between the data and the publication.

But of course, it is not sufficient to implement these technical requirements. It is also necessary to accompany the various categories of users in the use of this new system. Therefore, an important part of the project is to model and implement a seamless user experience associated with extensive training sessions adapted to all the stages of the research lifecycle.

Alongside these three major axes which structure the project, there will be a strong focus on data, particularly what are called “sensitive data” which are frequently used for research in SSH. In this case, the requirements between the need for strong security on the one hand, and the need to provide research projects with easy access to data on the other hand, are contradictory: the possibility of providing access to protected or sensitive resources is one of the important issues of the project. Finally, as connection and interoperability with other infrastructures at national and international levels is a key part of the project, APIs respecting general standards will be developed to be able to disseminate at large the information generated by the project.

The main expected outcome, besides making the publication process more fluid for research projects, is to foster the visibility of French research. This will be achieved mainly by improving the quality of its contents by following the FAIR^v principles and the development of research project skills. Another outcome is to have available all the bricks necessary for the construction of new types of complex publications in the context of rapidly evolving new forms of scientific writings, in particular data papers or actionable papers.

This paper will develop both the technical and organizational aspects of this ambitious project, which will require highly specialized professional profiles to support both producers (researchers, engineers, editors) and users (researchers, students, companies, public authorities, media) of content, and will last 8 years for a total cost around 30 M€.

Bibliography

1. Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) The citation advantage of linking publications to research data. PLoS ONE 15(4): e0230416. <https://doi.org/10.1371/journal.pone.0230416>
2. Burton, Adrian, & Koers, Hylke. (2016). ICSU-WDS & RDA Publishing Data Services WG Interoperability Framework Recommendations (1.0). <https://doi.org/10.15497/RDA00002>

ⁱ See <https://www.huma-num.fr> and <https://documentation.huma-num.fr/humanum-en/>

-
- ii See <https://www.openedition.org>
 - iii See <https://www.metopes.fr>
 - iv See <https://www.cnrs.fr/en/node/5607>
 - v See <https://www.go-fair.org/fair-principles>