A Gateway to Science: Fostering Access, Exchange, and Use of Social Science and Humanities Research Through a Digital Discovery Platform

Blotière, Emilie
emilie.blotiere@huma-num.fr
Huma-Num (CNRS), France

Disch, Leonie
ldisch@know-center.at
Know-Center GmbH, Austria

Breitfuss, Gert
gbreitfuss@know-center.at
Know-Center GmbH, Austria

Cupar, Drahomira
drahomira.cupar@gmail.com
University of Zadar, Croatia

Stojanovski, Jadranka
jadranka.stojanovski@irb.hr
University of Zadar, Croatia; Ruđer Bošković Institute, Croatia

Digital access to Social Sciences and Humanities (SSH) research resources is necessary to enable wide-scale use and reuse. However, regarding the access and usage of digital SSH research resources, we have identified - based on analyses and literature - the following challenges at the current state for SSH communities are:

1) No single access point allows SSH researchers to explore, find, access, and reuse materials such as literature, data, projects, and researcher profiles at the European level (competitor analyses; Breitfuss et al., 2022; Dumouchel et al., 2020b) → **Low visibility of SSH research results**.

2) SSH researchers have comparatively little experience and skills in using digital services/tools/platforms (user requirement analyses; Dumouchel et al., 2020a; Dumouchel et al., 2020b) → **Lack of digital literacy in the SSH community**.

3) Other potential users of SSH research (SMEs, journalists, citizens, policymakers) struggle to leverage these resources (Ross-Hellauer et al., 2022) → **Low usage of SSH research results out-side academia (SMEs, journalists, citizens, policymakers)**.

4) SSH research is not seen as "valuable" (Olmos-Peñuela et al., 2014) → **Low perceived impact on SSH research**.

Because of these challenges, a clear need for a digital open science platform that targets them is apparent. Such a platform should also be openly accessible so that anyone can access these resources. Based on the motivation to address these challenges, the European GoTriple discovery platform[1] dedicated to the SSH data including researchers' profiles and research projects, open to the whole scientific community and other socio-economic actors and citizens, was created. GoTriple responds to requests from SSH researchers to foster interdisciplinarity and enhance collaborations. The platform makes more than four million publications and datasets available from leading European aggregators, local repositories in less represented languages, more than three million researchers' profiles, and more than 20,000 research projects.

Five innovative services complement this approach by making available a trust building system, a recommender system, a visualisation tool, an open annotation tool and a crowdfunding service. Based on a co-design approach, the multicultural and multilingual GoTriple platform is indexing research data available in ten languages (Croatian, English, French, German, Greek, Italian, Polish, Portuguese, Spanish and Ukrainian) and in Slovenian by the end of the year.

To that extent, multilingualism is a core aspect of the platform. English as a lingua franca within scholarship is threatening the use of other languages which are especially relevant in the SSH disciplines dealing with locally relevant research. The widespread use of English in scholarly communication tends to marginalize non-English works and to make these works more difficult to access by most researchers. From the perspective of open science, GoTriple is promoting equity, diversity, inclusivity principles and closer connections to each language.

The TRIPLE consortium, consisting of 21 partners from 14 European countries[3], developed an interoperable multilingual discovery service for SSH research fields to enable scientists to make research results findable by searching in their own language and to make research results in different European languages available. A machine learning methodology was developed to train the search engine by documents in each language and a controlled vocabulary was created to deliver a TRIPLE thesaurus, currently composed of more than 4,000 concepts with labels in each of the 27 SSH disciplines (according to the MORESS categories[4]), in the eleven languages. The thesaurus was a prerequisite to align the vocabularies in the languages and for instance to find Greek documents with polish keywords. The GoTriple Vocabulary was initially based on the Library of Congress Subject Heading[5] (LCSH) and used primarily by the annotation service of GoTriple platform. The constitution of the textual training dataset of SSH sources can be reused for training machine learning and artificial intelligence algorithms. This corpus currently comprises more than 250,000 documents and has been tested in production for the implementation of the automatic classification service in GoTriple. For less represented languages in the scholarly literature such as Croatian, Polish, Ukrainian and Slovenian, the research community of these countries were solicited to contribute to finding relevant documents to train the system and appropriately translate the defined concepts. The solution is a publicly available SSH vocabulary, published in an open format (e.g., SKOS). The eleven included languages are expandable. It can be integrated into various types of platforms (via JSON/database file import or more dynamically by using the API that will be developed). The classification process is unique: algorithms for automatic classification and translation and human-curated classification and translation. Standardisation is necessary to offer efficient, useful, and updated vocabularies. This work is notably valuable to increase the impact and foster the reuse of research for the SSH communities worldwide.

To advance research, it is also necessary to improve the means of collaboration between the scientific community and society, which too often remains impervious to scientific work and, more

particularly, to SSH. One way to increase the impact of SSH research in society is to directly involve citizens in research projects to make them actors. The OPERAS[6] crowdfunding service is addressing this issue by introducing non-funded projects to a large audience. This service is integrated into the platform by giving direct and easy access to both researchers and contributors.

Users' requirements are evolving fast, and there is a strong need from the research community to provide scalable services. The challenge of innovation is to ensure sustainability. This will require efforts and collaboration of the consortium not only to maintain the platform in the long term but also to improve its services by indexing new repositories, reinforcing the coverage of some languages, and maintaining integrated services. To reach this needed long-term sustainability the project applied Alan Cooper's interaction design principles, namely *desirability*, *feasibility*, and *viability*, in developing a suitable platform Business Model (BM) (Cooper at al., 2007). The developed business model is backed up by a suitable governance model and a well elaborated business plan.

Notes

1. The TRIPLE EU-funded project (Targeting Researchers through Innovative Practices and Linked Exploration) started in Oct. 2019 under the framework H2020-INFRAESOC-2019-1 and under the 86420 GA number. https://gotriple.eu
3. Austria, Belgium, Croatia, Czech Republic, France, Germany, Greece, Italy, Netherlands, Norway, Poland, Portugal, Slovenia, United-Kingdom
4. https://project.gotriple.eu/wp-content/uploads/2020/12/MORESS-categories-for-GOTRIPLE-2020_12.pdf
5. https://id.loc.gov/authorities/subjects.html
6. Research Infrastructure supporting open scholarly communication in the social sciences and humanities (SSH) in the European Research Area - https://www.operas-eu.org/

Bibliography

Cooper, A. / **Reimann, R.** / **Cronin, D.** (2007): About face 3: the essentials of interaction design. John Wiley & Sons. DOI 10.5555/1406522

Dumouchel, S. / **Blotière, E.** / **Breitfuss, G.** / **Chen, Y.** / **Donato, F. D.** / **Eskevich, M.** / **...** / **Toth-Czifra, E.** (2020a): "GOTRIPLE: a user-centric process to develop a discovery platform". In: Information, 11(12), 563.DOI 10.3390/info11120563

Dumouchel, S. / **Blotière, E.** / **Barbot, L.** / **Breitfuss, G.** / **Chen, Y.** / **Di Donato, F.** / **...** / **Pohle, S.** (2020b): "TRIPLE project: building a discovery platform to enhance collaboration". In: ITM Web of Conferences (Vol. 33, p. 03005). EDP Sciences. DOI 10.5281/zenodo.3925735

Olmos-Peñuela, J. / **Benneworth, P.** / **Castro-Martinez, E.** (2014): "Are 'STEM from Mars and SSH from Venus'?: Challenging disciplinary stereotypes of research's social value". In: Science and Public Policy, 41(3), 384-400.DOI 10.1016/j.respol.2018.07.

Ross-Hellauer, T. / **Reichmann, S.** / **Cole, N. L.** / **Fessl, A.** / **Klebel, T.** / **Pontika, N.** (2022): "Dynamics of cumulative advantage and threats to equity in open science: a scoping review". In: Royal Society open science, 9(1), 211032.DOI 10.1098/rsos.211032