The Methods Hub: A Community Portal for Finding and Re-Using Methods

Johannes Kiesel
Christina Viehmann
Felix Victor Münch
Arnim Bleier
Chung-hong Chan
Po-Chun Chang
Raniere Gaia Costa da Silva
Ahrabhi Kathirgamalingam
Muhammad Taimoor Khan
Stephan Linzbach
Fakhri Momeni
Ran Yu
Claudia Wagner
Stefan Dietze
methodshub@gesis.org
GESIS – Leibniz Institute for the Social Sciences
Cologne, Germany

Abstract

Computational methods have opened up new avenues for research in many disciplines, especially in those concerned with large volumes of data like web science. However, many researchers have difficulties to find or apply novel methods or reproduce results because the method documentation is often not useful—especially for less technically oriented scientists—or incorrect. This paper introduces the Methods Hub, a new service dedicated to bring advances from computer and data science to other research communities, with a focus on social scientists. The Methods Hub assists researchers in finding, sharing, and applying computational methods. The service provides computational methods and tutorials with structured and carefully edited descriptions and instructions. To assist scientists in using the methods, the Methods Hub integrates interactive environments, in which methods can be executed and tested online. To increase the reproducibility of research, the Methods Hub allows for exact references to each version of a method via DOIs. Furthermore, the Methods Hub provides method guidelines and templates to simplify the implementation of best practices. The Methods Hub thus serves as community-based toolbox to find, share, and apply methods to social science questions, like those asked in web science.

CCS Concepts

• **Information systems** → **Collaborative and social computing systems and tools**; • **Software and its engineering** → *Software post-development issues.*

Keywords

Online Execution, Reproducibility, Software Portal, Software Reuse

1 Introduction

Research software plays a central role in advancing research, especially in data-driven research as it is typical for web science. Research software is more than just a specific functionality; it is an expression of the collective experience gained during technical fieldwork on a research problem [4]—and is often lost when funding for a research project ends and the developed software is not prepared for reuse. An impressive example of the importance of research software as a research enabler is the Hugging Face Hub, which now provides access to more than 2 million AI models.[1] However, it is difficult to both find and use the best model for a specific task on Hugging Face Hub, especially for researchers who

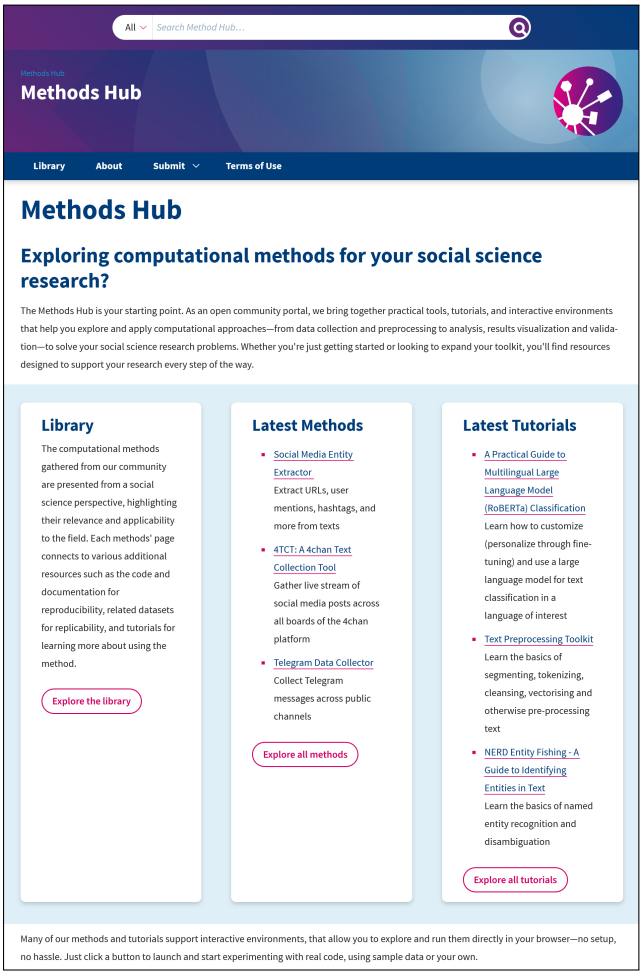

Figure 1: The Methods Hub index page, detailing basic information and latest content, with links to explore the entire library or only methods or tutorials specifically.

[1] https://huggingface.co/models

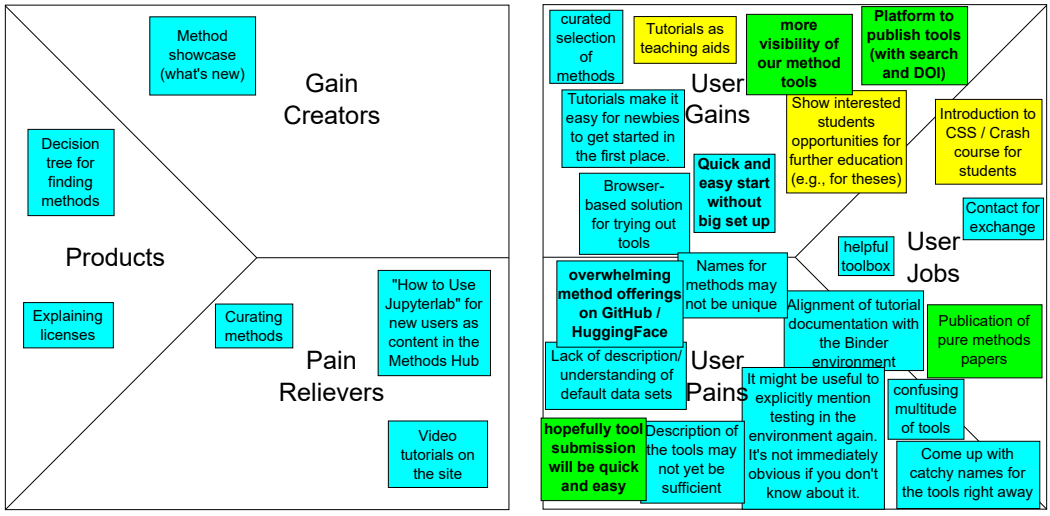

Figure 2: Example of feedback gathered during our community-driven design process. Here, the categories of the Value Proposition Canvas [3] are used to organize the feedback: service (left) products / gain creators / pain relievers; user (or "customer", right) gains (from using the service) / jobs / pains (without service). The color of the sticky notes indicates feedback related to method developers (green), educators (yellow), and method users (blue). This specific feedback was gathered as part of a workshop at the *Annual Methods Division of the German Communication Association Conference (DGPuK Methoden)* in September 2025. Feedback translated from German, except bolded which was in English originally.

are not experts in the field of AI—and there are many more methods that do not fit into the Hugging Face Hub scheme.

This paper introduces the Methods Hub (cf. Figure 1), an online portal for finding and re-using methods from the research community—especially methods in the range from basic preprocessing scripts to advanced AI models for analysis, as used in web science—, and for sharing own methods. The Methods Hub is more than just a methods registry. An editorial process ensures that accepted methods are reusable even for researchers with only basic programming skills. Submitted methods are preserved and assigned a DOI so that other researchers can trust their availability—a necessity for reproducibility. In addition, the Methods Hub offers interactive tutorials that teach basic and advanced methods.

The Methods Hub, which is developed in a community-driven design process (Section 2), differs from similar initiatives such as the Social Media Observatory wiki [5] or the Digital Methods Initiative Tool Archive[2] in that it provides a customized search interface for finding methods and method tutorials (Section 3), a content curation process to ensure the appropriateness and usability of this content even for researchers outside the field of computer science (Section 4), and integration of interactive environments for easy access and quick testing without installation (Section 5).

2 Design Process

To ensure that our service is aligned with the needs of the community, we follow a community-driven design process that involves reaching out to researchers and constantly gathering feedback. To this end, we have organized several workshops at community conferences to present the Methods Hub (cf. Figure 2 for an example of such workshop feedback). To position the Methods Hub at the intersection of computer science methods and social science perspectives, the team integrates expertise from both fields. We thus identified three target groups: method developers, method users, but also "multipliers"—for example educators—that spread content for their own purpose and benefit from a well-organized repository. In a nutshell, the community feedback clearly showed that (1) method developers are interested in enlarging their target audience through submission to online portals; (2) method users are interested in a portal that assists them in finding, learning about, and applying methods, especially when they trust that the methods in the portal are curated well; and (3) multipliers are interested in a portal that offers to their own customers (e.g., students) both entry-level tutorials and opportunities for self-directed exploration of methods and tutorials. The following sections describe the core functions of the Methods Hub that serve to address these interests.

3 The Methods Hub Portal

The Methods Hub is developed as part of our research institute's service offer to social scientists, and its design builds on top of decade-long experience of providing services to social scientists. For familiarity, its interface aligns with all other parts of the offer. However, also other scholars than social scientists can benefit from our resulting focus on practical use cases and streamlined execution.

The library is the heart of the Methods Hub, providing access to all methods and tutorials through a search interface (cf. Figure 3). The Methods Hub's library is based on and integrates with a search engine for social scients [1] with about 30,000 monthly users on, amongst others, publications, datasets, survey variables, and tools. The library provides several filter facets, and features result snippets that are tailored towards quickly scanning methods and

[2]https://wiki.digitalmethods.net/Dmi/ToolDatabase

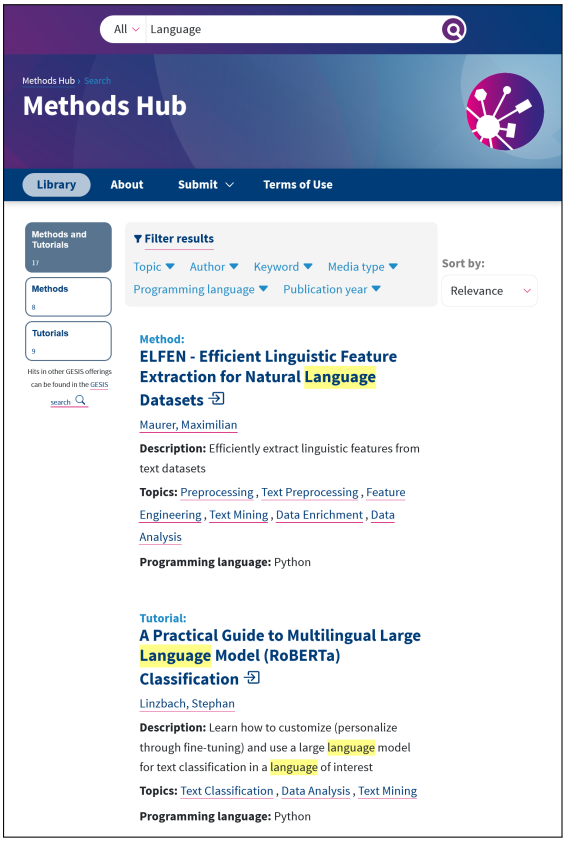

Figure 3: Excerpt from the Methods Hub search interface for the search query "Language" with the two best results and options for filtering the results and changing the ranking.

tutorials. The portal also provides action links to execute methods and tutorials in online interactive environments (cf. Section 5).

Submission to the Methods Hub comes with several benefits to the submitters, especially when it comes to visibility, citability and usability of the content. Beyond the search, a separate team at our institute ensures that the content is findable through the standard search engines and generative models. Moreover, the Methods Hub ensures that the content is linked with other research knowledge graphs [6]. For citability, the Methods Hub automatically registers a DOI with each new version of the content and preserve old versions via Git version control and associated commit hash, so that other researchers can cite the content appropriately, reliably and easily. For usability, the Methods Hub team assists during the submission process with advice and best practices (cf. Section 4), and ensures the content is compatible with interactive environments (cf. Section 5) to make it easier to use and thus more attractive for others.

4 The Methods Hub Content

The Methods Hub offers computational methods and tutorials on how to apply such methods for every step of a data-driven research project: from data collection and preprocessing to analysis, visualization and validation of results. Whether researchers are just getting started with data-driven research (e.g., in need of a tutorial on text preprocessing) or want to expand their toolkit (e.g., with fine-tuning of language models), they will find resources to support them every step of the way. The computational methods compiled from the research community are presented from a non-technical perspective, highlighting first their relevance and applicability to specific research questions. This focus enables less technically oriented scientists to make informed decisions on which method to use, before they dive deeper into the how-to-use sections.

- Title
- Authors - with ORCID for unambiguous attribution
- Meta data plus actions for referencing and execution
- Description - brief overview of what it is and how to use it
- Use Cases - who used or could have used it for what research
- Input Data - definition and example data
- Output Data - definition and output for input example data
- Hardware Requirements - both specs and in layman terms
- Environment Setup - environment and package installation
- How to Use - steps to produce the output data and others
- Technical Details - how it works and how to adapt it
- References
- Acknowledgements
- Disclaimer
- Contact Details

Figure 4: Overview of the developed section schema for method READMEs. Optional (suggested) sections are shown in gray and more sections can be added if needed.

To ensure consistency and the use of best practices and thereby help the target audience in comprehending the specifics of each method and tutorial more quickly, we developed a checklist and templates for researchers who want to submit their methods and tutorials.[3] For methods, we developed a README specification that is compatible with the large source code platforms (especially GitHub and GitLab) and guides readers, while allowing the respective developers to describe intricacies and advanced use in a "Technical Details" section (cf. Figure 4). This specification of the "Methods Hub Friendly README" is formatted itself as README with examples, thus serving as a template for new methods at the same time. The checklist and the specification of the README, are based on prior work on ensuring method reproducibility and best practices [2]. Similarly, we provide templates for writing tutorials.

5 The Methods Hub Interactive Environments

The Methods Hub provides access to interactive environments where users can directly apply, test, and adapt methods. These environments allow for hands-on learning and immediate application of methods to own data without installation on ones machine, avoiding technical issues and restrictions from institution machines—typical errors when using methods from method registries or blog posts. However, the interactive execution is not required for taking a first look at a method or tutorial, as the Methods Hub renders a complete execution including outputs on its web page.

[3] https://github.com/GESIS-Methods-Hub/guidelines

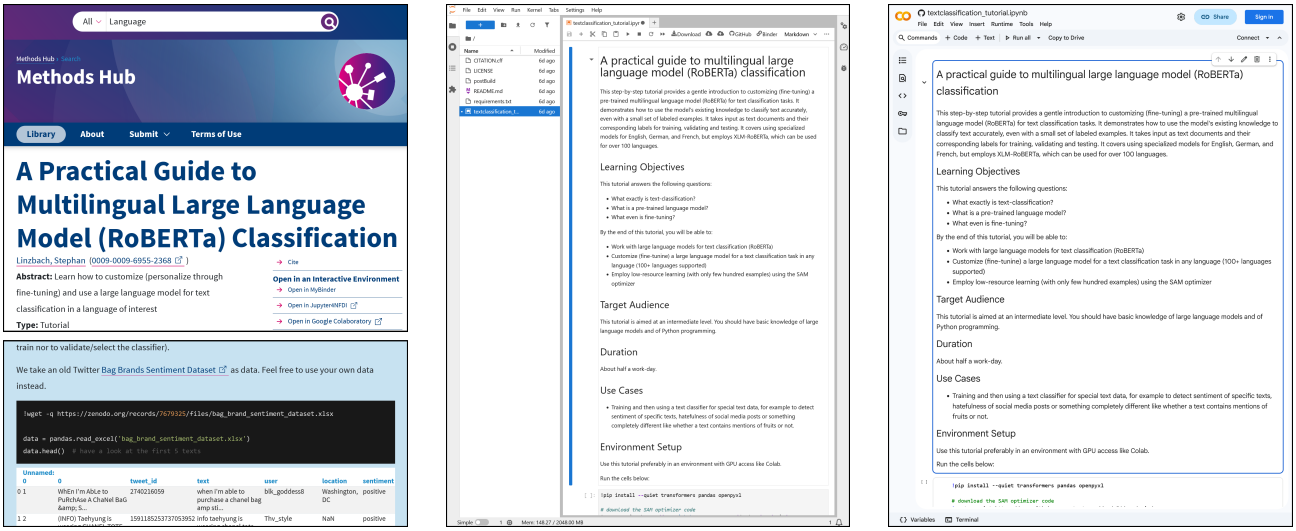

Figure 5: Several excerpts of the Methods Hub and linked interactive environments for one tutorial. The content page in Methods Hub (shown top left) displays four action links on the right-hand side, including three links for opening the tutorial in interactive environments for execution ("Open in *X*"—only available if the method or tutorial supports the respective environment). The content page displays the results of a full tutorial execution (shown bottom left; the table is the result of the executing the code above), allowing users to quickly see whether the method or tutorial meets their needs. The MyBinder and Jupyter4NFDI environments (latter shown center) use the same interface. But, while the MyBinder environment does not require an account or registration, the Jupyter4NFDI environment requires an ORCID or university account, but offers better hardware. Using Google Colaboratory (shown right) requires a Google account, but also provides access to GPU resources.

Methods and tutorials often need to be adapted for compatibility with such environments. For example, the MyBinder environment requires standard files for dependency management, while Google Colaboratory requires a standalone Notebook file.[4] The Methods Hub team assists submitters in making their content compatible with as many of the environments as possible. Figure 5 illustrates the integration of interactive environments of one tutorial.

6 Conclusion

This paper presents the Methods Hub, an online portal for finding and re-using computational methods for and from the research community, that is publicly available at https://methodshub.gesis.org. The Methods Hub is developed as part of GESIS's service offer to social scientists, but its key features—quality assurance to ensures ease-of-use also for less technically oriented researchers, and an integration into search infrastructure, the assurance of the contents' citability and streamlined execution—make it attractive also for other researchers. Computational methods researchers can contribute to the Methods Hub to advertise their work to a larger audience and to ensure the above mentioned qualities. Focussing on methods from data acquisition to result presentation, the Methods Hub is also especially of interest for the web science community.

[4]MyBinder: https://mybinder.org/; Google Colaboratory: https://colab.google/

Acknowledgments

The work presented in this paper is financially supported by the projects NFDI4DS (DFG project no. 460234259) and Jupyter4NFDI as part of Base4NFDI (DFG project no. 521453681).

References

[1] Daniel Hienert, Dagmar Kern, Katarina Boland, Benjamin Zapilko, and Peter Mutschke. 2019. A Digital Library for Research Data and Related Information in the Social Sciences. In *19th ACM/IEEE Joint Conference on Digital Libraries (JCDL 2019)*, Maria Bonn, Dan Wu, J. Stephen Downie, and Alaine Martaus (Eds.). IEEE, 148–157. doi:10.1109/JCDL.2019.00030

[2] Fakhri Momeni, Muhammad Taimoor Khan, Johannes Kiesel, and Tony Ross-Hellauer. 2025. Checklists for Computational Reproducibility in Social Sciences: Insights from Literature and Survey Evaluation. In *3rd ACM Conference for Reproducibility and Replicability (REP 2025)*, Thomas Pasquier, Ashish Gehani, and Khalid Belhajjame (Eds.). ACM, 179–191. doi:10.1145/3736731.3746161

[3] Alexander Osterwalder, Yves Pigneur, Gregory Bernarda, and Alan Smith. 2015. *Value proposition design: How to create products and services customers want.* John Wiley & Sons.

[4] Bernhard Rieder, Stijn Peeters, and Erik Borra. 2022. From Tool to Tool-Making: Reflections on Authorship in Social Media Research Software. *Convergence: The International Journal of Research into New Media Technologies* 30, 1 (Dec. 2022), 216–235. doi:10.1177/13548565221127094

[5] Gregor Wiedemann, Felix Victor Münch, Jan Philipp Rau, Phillip Kessling, and Jan-Hinrik Schmidt. 2023. Concept and Challenges of a Social Media Observatory as a DIY Research Infrastructure. *Publizistik* 68, 2 (Aug. 2023), 201–223. doi:10.1007/s11616-023-00807-6

[6] Matthäus Zloch, Danilo Dessì, Jennifer D'Souza, Leyla Jael Castro, Benjamin Zapilko, Saurav Karmakar, Brigitte Mathiak, Markus Stocker, Wolfgang Otto, Sören Auer, and Stefan Dietze. 2025. Research Knowledge Graphs: The Shifting Paradigm of Scholarly Information Representation. In *The Semantic Web - 22nd European Semantic Web Conference (ESWC 2025) (LNCS, Vol. 15719)*, Edward Curry, Maribel Acosta, María Poveda-Villalón, Marieke van Erp, Adegboyega K. Ojo, Katja Hose, Cogan Shimizu, and Pasquale Lisena (Eds.). Springer, 140–154. doi:10.1007/978-3-031-94578-6_8