Basic information

First name

Last name

Email

Phone

Link to your profile In case you want to provide your profile in the code forge of your choice, like GitHub or Codeberg.

Country of residence of the person who will sign the contract.

How did you hear about the Sovereign Tech Fellowship and/or the Sovereign Tech Agency?

Legal Requirements

I am legally authorized and competent to sign contractual agreements.

- [ ] Yes
- [ ] No

- [ ] I acknowledge: All code and documentation to be supported within the Sovereign Tech fellowship must be licensed such that it may be freely reusable, changeable and redistributable.

- [ ] I acknowledge: I will not invoice other private or public organizations for activities carried out within the Sovereign Tech Fellowship.

- [ ] I acknowledge: I am an open source maintainer, community manager, or tech writer myself and I am not applying on behalf of a business entity that is planning to employ or contract fellows using this program.

Attachments (PDF only)

1. Cover letter We want to hear more about your motivation to apply for the fellowship. Please create a cover letter and tell us why you are applying, including your background in Open Source and your past journey with the projects you are maintaining or contributing to. The document does not need to follow a formal format and **should not be more text than one page** (PDF only).

Upload PDF

Martin Magdinier
Montreal, QC
Canada

April 3rd, 2026

To the Sovereign Tech Fund Selection Committee,

I am applying to the Sovereign Tech Fellowship to support my work on OpenRefine, where I contribute as a community manager.

I began contributing to OpenRefine in 2012 by writing tutorials and documentation. Since its transition to a volunteer-led project in 2013, my contribution has focused on exploring sustainability models for open-source projects. In 2019, I helped establish a grant-based model and onboarded Code for Science and Society as a fiscal sponsor. I served five years on the Advisory Committee and, in 2023, took a paid role to dedicate more time to community activities. Over time, my focus shifted from using the tool to supporting the conditions that allow the project to function, with a particular interest in coordination, governance, and contributor support for small maintainer teams.

OpenRefine is maintained by a dedicated group of core contributors and a wider volunteer community. As the community manager, I facilitate communication, maintain governance processes, support contributor onboarding, and handle maintenance tasks, such as reviewing contributions to the documentation and website, preparing release notes, and triaging issues. This coordination is essential for effective collaboration among stakeholders and for the optimal use of maintainers' capacity. While institutional reliance on OpenRefine is increasing, it remains fragile, underscoring the need for continued coordination and support for contributors.

The fellowship will support 7 hours for sustaining coordination and contributor infrastructure for OpenRefine. Code for Science and Society will provide 3 hours for additional administrative and fundraising activities. My goal is to maintain an environment that enables effective participation from both volunteers and institution-supported contributors, while keeping the project accessible for users without expanding the centralized team. I am eager to engage with the STA’s maintainer community by participating, as I previously saw the benefit of such communities of practice via the CSCCE programs.

I am based in Canada and operate as an independent contractor, able to engage internationally.

Martin Magdinier

2. Technology Details

2.1 Please provide us with some high level information about the projects you are maintaining or contributing to. If available please include up to three repository URLs below the description. Describe each open source project in a sentence. Please highlight which role you have in each of them (maintainer, contributor, community manager or tech writer). You have 10000 characters remaining

I primarily contribute to **OpenRefine** as a **community manager.**

OpenRefine is an open-source tool for working with structured data, enabling users to clean, transform, and reconcile datasets using shared identifiers and linked data standards. It is used to prepare data for integration into knowledge systems, including authority files and knowledge graphs, and supports workflows that align datasets and make them interoperable across repositories.

The project sits at the intersection of data preparation and knowledge infrastructure, providing an accessible interface for domain experts to work with structured data without requiring custom development. It supports reproducible data transformations and reconciliation workflows, making it suitable for use in research, cultural heritage, and open knowledge contexts.

My role focuses on supporting the project's coordination and sustainability. This includes facilitating communication between maintainers, contributors, and users; maintaining governance processes; and supporting contributor onboarding and documentation. I help surface project priorities and support a shared understanding of them across the community.

I also contribute to maintenance workflows by reviewing documentation and website pull requests, supporting issue triage and clarification of user requirements, and helping communicate releases through walkthroughs of new features. This work supports maintainers by reducing coordination overhead and improving the project's accessibility for contributors and users.

In addition to the core repository, OpenRefine includes an ecosystem of extensions and reconciliation services that enable integration with external systems such as Wikibase and authority files (e.g. VIAF, GND, the Getty Vocabularies, ORCID, and ROR). These components are developed and maintained by contributors and partner institutions. My work supports this broader ecosystem by helping coordinate contributions, maintain shared practices, and facilitate collaboration.

If you want to reference more than 3 repositories, please use the description text box above.

2.2 Repository URL

#1 https://github.com/OpenRefine/OpenRefine

#2 https://github.com/OpenRefine/openrefine.org

#3

3. Project Descriptions

3.1 We would like to learn more about the projects where you hold the role of a maintainer, community manager or tech writer. Please be clear about which project you are referring to within the text fields.

Describe your projects more in-depth. Why are they critical? You have 1500 characters remaining

OpenRefine is critical as it provides a key layer in data infrastructure workflows, enabling the reconciliation, transformation, and alignment of data using shared identifiers and linked data standards.

It is widely used in the digitization and curation of collections, where institutions need to prepare, clean, and enrich data and metadata before publication. OpenRefine enables domain experts to work directly with structured data, reconcile it against authority files, and transform it into interoperable formats.

This includes workflows involving library and archival formats such as MARC, as well as integration with knowledge graph systems, including RDF-based data and SPARQL endpoints, such as Wikibase. By supporting these processes, OpenRefine helps ensure that data can be consistently structured, linked, and reused across systems.

As a result, it plays a role in enabling datasets to move from local, heterogeneous formats into shared, interoperable knowledge infrastructure.

3.2 Please provide a brief overview of projects that depend on your technology. You have 1500 characters remaining

OpenRefine is used across research and open knowledge ecosystems to support data reconciliation and integration workflows.

Within the German National Research Data Infrastructure (NFDI4Culture), OpenRefine is formally endorsed as a tool for data curation and enrichment across a network of universities, archives, and cultural institutions.

The Coordination Centre for Scientific University Collections (SODa) at Humboldt-Universität zu Berlin and its network use the software to support linked data workflows and persistent identifiers across academic collections, reaching hundreds of institutions.

OpenRefine is also a primary interface for reconciliation services provided by hbz des Landes NRW and partners within the Gemeinsame Normdatei (GND) network, supporting memory institutions in the DACH region.

Finally, the Wikimedia Foundation recognizes OpenRefine as critical public interest infrastructure. It has enabled cultural heritage institutions to upload nearly 3 million files to Wikimedia Commons, providing multimedia content for Wikipedia in multiple languages, and is essential for increasing machine-readable information on Wikidata (supported by Wikimedia Deutschland).

These examples reflect a broader pattern of reliance on OpenRefine for data preparation and interoperability across global research infrastructures. Formal letters of support from NFDI4Culture, hbz des Landes NRW, and the Wikimedia Foundation are available upon request.

3.3 Which target groups do your projects address (who are its users?) and how would they benefit from a paid maintainer fellowship (directly and indirectly)? You have 1500 characters remaining

OpenRefine is used by a range of communities working with structured data, in particular GLAM institutions (galleries, libraries, archives, museums), research infrastructures, and open knowledge communities such as Wikimedia. Its users are typically domain experts, data stewards, and contributors who need to prepare, reconcile, and publish datasets without relying on custom development.

These users depend on OpenRefine to work with authority files, linked data systems, and heterogeneous data sources. In many cases, it is used as a bridge between local collections and shared knowledge infrastructures.

A paid community manager fellowship would enhance project coordination and sustainability for these groups. It directly supports maintainers by reducing coordination overhead, improving documentation, and clarifying contribution pathways, thereby making it easier for new contributors to get involved.

It would also support the development of a community of practice through regular community calls, enabling users from different institutions and domains to exchange workflows, share knowledge, and learn from each other.

Indirectly, it improves the reliability and continuity of the tool by ensuring that user needs are surfaced, prioritized, and communicated to maintainers. This helps align contributions with real use cases and supports the long-term stability of workflows that institutions depend on.

3.4 How was the work on the project made possible so far (structurally, financially, including volunteer work)? If applicable, list other sources of funding that you applied for and/or received. You have 1500 characters remaining

OpenRefine is maintained through volunteer contributions, institutional support, and targeted grants. Initially supported by dedicated funding, the project was fully volunteer-led from 2013 to 2019. Since 2019, it has operated under a hybrid model of grant-funded work and community contributions.

Today, OpenRefine relies on a small coordinating core and a broad community. Rather than building a large paid team, we enable contributions from individuals and institutions as part of their professional or voluntary work. This includes developers, documentation writers, and domain experts.

Financially, the project has received grants from the Chan Zuckerberg Initiative (EOSS) and the Wikimedia Foundation for Wikidata and Wikimedia Commons integrations. These funds supported specific development and ecosystem improvements.

Currently, OpenRefine operates on a limited annual budget (around USD 50,000) for part-time coordination and community infrastructure. This model prioritizes maintaining the conditions for participation, including coordination, guidance, and shared priorities, over expanding a centralized team. This resourceful approach remains active and responsive but depends on minimal, consistent coordination to align diverse contributions over time.

This approach has enabled the project to remain active and responsive, but it depends on maintaining a minimal level of coordination to support and align contributions over time.

3.5 What are possible alternatives to your project and how does your project compare to them? You have 1500 characters remaining

Possible alternatives to OpenRefine include spreadsheet tools (such as Excel), scripting environments (such as Python with pandas), and data transformation or integration platforms.

Spreadsheet tools are widely accessible but do not support reproducible workflows or systematic reconciliation against external data sources. Scripting environments provide flexibility and scalability but require programming skills, limiting accessibility for many domain experts. ETL tools can handle complex pipelines, but are often not designed for interactive data exploration or reconciliation with authoritative data sources.

OpenRefine occupies a distinct position by combining an accessible user interface with support for reproducible data transformations and reconciliation workflows. It allows domain experts to clean, transform, and align data using shared identifiers without requiring custom development.

In addition, OpenRefine integrates with linked data systems such as Wikibase and supports workflows around authority files like the Gemeinsame Normdatei (GND), enabling data to be prepared and published as interoperable datasets.

This combination of usability, reproducibility, and integration with knowledge infrastructure makes it particularly suited for data curation and digitization workflows in research and cultural heritage contexts.

3.6 How is the maintenance and governance of the projects structured? How many people are maintaining the projects?

You have 1500 characters remaining

OpenRefine is maintained by a distributed group of contributors, with a governance structure designed to support coordination and transparency across the project.

Governance (https://github.com/OpenRefine/OpenRefine/blob/master/GOVERNANCE.md) is structured through a Core Dev Group, responsible for technical decisions, and an Advisory Committee (three volunteers), which provides governance support, including oversight of the project budget. Decisions are documented and shared publicly to ensure transparency.

Maintenance is led by the Core Developers Group, currently two volunteer maintainers, supported by a broader set of committers. Over the last 12 months, 16 contributors have been active on GitHub, with 135 open issues and 103 merged pull requests (excluding those created by Dependabot). Project metrics are available at https://openrefine.org/usage

The Release Manager (volunteer) oversees the release process, and a Project Manager (paid), who also fulfills the community manager role, supports coordination, communication, and contributor engagement.

This structure enables distributed maintenance but requires ongoing coordination to remain effective, particularly given the limited number of active maintainers and the diversity of contributors involved.

4. Scope of Work

4.1 What would be your focus within this fellowship? Please highlight some recurring and non-recurring activities that you will spend your time with. Please also describe how your work will enable new contributors to join the projects. You have 4500 characters remaining

This work focuses on maintaining and strengthening the organizational infrastructure that enables contributors to participate effectively in OpenRefine. By facilitating a community of practice and lowering barriers to entry, it supports both individual and institution-based contributors in identifying priorities, understanding how to participate, and moving from users to contributors.

1) Maintainer coordination and knowledge flow

I will facilitate structured monthly community and contributor calls. These calls provide a shared environment for contributors, maintainers, and users to exchange knowledge, surface blockers, and share ongoing work. By documenting and sharing outcomes, I turn these discussions into actionable entry points. This work reduces knowledge silos and creates visible entry points that help contributors identify how to get involved.

2) Community surveys and priority setting

In 2026, I will conduct bi-yearly community surveys to capture use cases and identify priorities. Previous cycles received 207 responses in 2022 and 226 in 2024 (see the 2024 results: https://openrefine.org/blog/2024/12/20/2024-survey-results). Results will be documented and used to update the project's goalposts and inform roadmap discussions. This process translates raw user needs into clearly defined "Goalposts," providing a roadmap that allows contributors to engage with high-impact tasks that reflect real-world usage.

3) Contributor pathways and AI-assisted contributions

I will maintain and expand contributor guidelines to support all contribution types, including documentation, support, training, and emerging practices for AI-assisted contributions. This work help reduce maintainer burnout by establishing expectations that improve the quality of incoming contributions, allowing maintainers to focus on high-level review rather than repetitive process clarification.

4) Governance transparency and documentation

I will maintain governance records (https://github.com/OpenRefine/OpenRefine/blob/master/GOVERNANCE.md) and ensure documentation regarding roles, processes, and decisions remains up to date. Transparent governance increases institutional trust, providing the stable framework necessary for organizations to confidently commit staff time or funding to the project.

5) Maintenance workflow support

I will support maintenance workflows by reviewing documentation and website pull requests, assisting with issue triage, and clarifying user requirements. This improves the quality and readiness of contributions for these essential non-code pathways, reducing iteration cycles and ensuring that incoming work can be integrated efficiently by the core team.

6) Communication and release support

I will produce walkthroughs of new features for each major release and update project documentation. By making technical changes immediately accessible, I support the broader ecosystem of trainers and consultants who rely on OpenRefine, ensuring the tool’s advancements are quickly adopted by the community.

7) Institutional engagement and ecosystem coordination

I will engage with institutional partners and extension developers to support collaboration across the ecosystem. This includes sharing institutional needs with the broader community and identifying opportunities for organizations to contribute staff time or resources to the project. This coordination aligns the efforts of major stakeholders, such as the Wikimedia Foundation and NFDI4Culture, with the core project, securing the long-term sustainability of the contributor base.

Together, these efforts create a structured yet open environment where contributors can contribute effectively to OpenRefine's long-term sustainability.

4.2 What is your preferred amount of weekly working hours? (6 - 32h) - integer

4.3 What is the hourly rate you are asking for? Please make sure to include all expenses you will have in order to perform this work (fully loaded rate, excluding taxes). This rate has to be in EUR. integer

55

4.4 What is your preferred length of the contract in months? (3-12 months) integer

12

- [ ] I have read and understood the privacy policy