Silicon Valley Community Foundation - Essential Open Source Software for Science Cycle 5 Progress Report

Grant Overview

Grant title: Improving OpenRefine's reproducibility

Grant number: EOSS5-0000000404

Date report submitted: 2026-02-28

Lead applicant name: Antonin Delpeuch, Martin Magdinier

List of key personnel directly funded by this EOSS award (including the lead)

Name	Email	Open source project affiliation
Antonin Delpeuch	antonin@delpeuch.eu	OpenRefine
Freexian	raphael@freexian.com	Debian
Zoe Cooper	zoe@zoe-cooper.com	OpenRefine
Rory Sawyer	rory@sawyer.dev	OpenRefine
Martin Magdinier	martin@openrefine.org	OpenRefine

Financial Overview
Provide original estimated and actual spending for each category for the reporting period. You can also submit this information as a separate spreadsheet.

	Original Yearly Budget	Original Total Budget	Year 1 Actual	Year 2 Actual	Year 3 Actual	Total Actual	Diff
(1) Development	$63,000.00	$126,000	$66,387.11	$31,141.20	$90,570.00	$188,098.31	-$62,098.31
(2) Design and design research	$8,000.00	$16,000		$38,400.00		$38,400	-$22,400.00
(3) Training and documentation	$10,000.00	$20,000					$20,000.00
(4) Product management	$28,800.00	$57,600					$57,600.00
(5) Community outreach	$10,000.00	$20,000			$28,814.52	$28,814.52	-$8,814.52
(6) Travel meeting equipment and PR	$12,000.00	$24,000	$122.94	$1,632.75	$3,420.07	$5,175.76	$18,824.24
(7) Personnel fringe							
Fiscal sponsorship fee (15%)	$23,250.00	$46,500	$23,257.50	$23,257.50		$46,515.00	-$15.00
TOTAL	$155,050.00	$310,100.00	$89,767.55	$94,431.45	$122,804.59	$307,003.59	$3,096.41

Budget narrative: Briefly explain any changes to the original budget, any challenges impacting spending; unspent budget; and plans for use of remaining funds.

Final financial figures reflect reconciliation with the fiscal sponsor's General Ledger export at project closeout. Variations between categories primarily result from staffing transitions and operational adjustments made during the extended final phase of the project, while total spending remained within the approved grant budget.

(1) **Development**
During the final phase of the project, development capacity became the primary operational need to ensure delivery and stabilization of the reproducibility improvements. Antonin Delpeuch progressively transitioned out of the developer role, ending funded work on April 1, 2025. A new Developer and Contributor Engagement contractor (Rory) was hired starting March 1, 2025 (20 hours/week, fully funded by EOSS-5). Concentrating resources on sustained development and maintenance explains the higher proportion of spending allocated to this budget line compared to the original projection.

(2) **Design and design research / Training and documentation**
As described in the interim report, design, research, and documentation activities were grouped into a single engagement to enable longer and more coherent design work rather than multiple short contracts. This regrouping supported completion of user research, interface design work, and

preparation of documentation materials aligned with the project's reproducibility goals.

(3) Product management and Community outreach
Throughout earlier reporting periods, product and project management activities led by Martin Magdinier were funded through other grants. Beginning in Year 3, part of this role was allocated to EOSS-5 funding, combining product management and community outreach responsibilities into a single coordination role. This reflects the increased emphasis on release coordination, contributor onboarding, and project sustainability during the closing phase of the grant.

(4) Travel, meetings, equipment and PR
Travel spending remained significantly below the initial budget. Several events were financed through complementary funding sources, and project activities favored local or remote community engagement. In the final phase, resources were intentionally prioritized toward developer time and community support activities rather than conference attendance, that benefitted a small number of project staff.

(5) Fiscal sponsorship fee
Fiscal sponsorship costs remained consistent with the planned 15% rate and reflect administrative, accounting, and operational support provided by Code for Science & Society throughout the grant period.

Overall, the project concluded with a small underspend relative to the total approved budget. Variations across categories reflect operational adjustments made to maintain project continuity, support staffing transitions, and prioritize development work necessary to complete the planned deliverables during the no-cost extension period ending December 2025.

Please list any other funding received during this reporting period that <u>extends the work you are doing on this project</u> (if relevant):

Award Title	Funding Source	Award Amount (US$)
Training & Sustainability Grant (2023–2024)	Wikimedia Foundation	50,000
Vanguard Charitable (2024)	Unrestricted donation	2,500
FLOSS Fund (2025)	Unrestricted donation	25,000

Progress Overview

Progress towards the deliverables:
Please provide a summary of milestones / deliverables achieved or significantly contributed to through this grant as a bulleted list (short description + optional notes)

UX/UI and community support capacity established (Deliverable 1)
Contractor support was engaged to strengthen UX/UI design and community engagement activities. Design and community input were successfully integrated into development workflows.

Community-informed user research conducted (Deliverable 2)
User research was carried out through one-to-one walkthroughs, forum discussions, and community events including OpenRefine BarCamp (2024 and 2025). Feedback-informed architectural and usability decisions related to workflow reproducibility and reuse.

Public UI/UX design process completed (Deliverable 3)
Design mock-ups, exploratory prototypes, and technical specifications were published through GitHub and community discussions, incorporating iterative feedback from users and developers. Design outcomes informed implemented improvements in workflow legibility and usability.

Stable releases delivering reproducibility improvements (Deliverable 4)
Development culminated in the release of OpenRefine 3.10.0 (following the 3.9 release series), delivering enhanced workflow validation, recipe-reuse safeguards, column-dependency tracking, improved undo and safety mechanisms, and usability improvements to support reproducible data workflows.

Expanded Continuous Integration and testing infrastructure (Deliverable 4)
Automated test coverage increased substantially, including expanded Java unit tests and end-to-end workflow tests. CI practices were extended to test compatibility with extensions prior to release, strengthening ecosystem stability.

Improved packaging and installer reliability (Deliverable 4)
Distribution processes were strengthened across platforms, including improved Windows and macOS packaging workflows, maintained container images, and continued Debian packaging support, improving reproducible installation environments.

Technical documentation and contributor onboarding enhanced (Deliverable 5)
Contributor-facing documentation and onboarding guidance were expanded to include documentation supporting extension testing and development workflows. Release notes, walkthrough materials, and progress reports documented reproducibility-related improvements.

Community dissemination and governance strengthened (Deliverable 6)
Reproducibility improvements were presented through public events and release walkthroughs, including community BarCamps and online discussions. Advisory governance was expanded to strengthen representation from research-oriented communities, and existing Code of Conduct practices were maintained and clarified.

Major changes in scope or project plan:
Please describe any problems encountered and steps taken to overcome them, as well as any general scope or plan changes.

During the course of the project, several adjustments were made to the original plan to better align with technical realities, community governance processes, and sustainability goals typical of mature open-source infrastructure projects. These changes did not alter the overall objectives of the grant but influenced how milestones were achieved.

Shift from sequential phases to iterative development
The proposal initially described clearly separated research, design, and implementation phases. In practice, these activities proceeded iteratively and in parallel. Continuous feedback from users and contributors proved most effective, allowing ideas to be validated incrementally and reducing implementation risk. This adaptation aligned with OpenRefine's established open governance model and supported sustained community participation throughout development.

User research conducted as an ongoing process
Early plans envisioned a bounded user research phase involving a fixed number of consultations. Instead, feedback was gathered continuously through forum discussions, one-to-one walkthroughs with researchers, community-wide surveys, and community events such as OpenRefine BarCamp (2024 and 2025). This approach better reflected how consensus and requirements emerge in open-source communities through sustained asynchronous collaboration.

Adjustment of technical roadmap from planned 4.x architecture to incremental 3.x releases
Initial exploration considered larger architectural changes targeting a future 4.x release. Through community governance discussions and implementation experience, the team determined that incremental improvements within the stable 3.x release series would deliver reproducibility benefits sooner while minimizing disruption to users and extensions. Development therefore focused on shipping validated improvements through the 3.9 and 3.10 release cycles.

Prioritization of workflow integrity over new visualization and automation features
Early project concepts emphasized graphical workflow visualization and more extensive automation capabilities. Prototype evaluation and community feedback indicated that preventing workflow errors and improving reliability were more immediate user needs. Effort was therefore redirected toward validation mechanisms, safeguards, and usability improvements, while visualization and advanced automation work remained at prototype or foundational stages for future development.

Delayed recruitment of UX/UI contractor
Recruitment took longer than anticipated due to the need to identify candidates with both design expertise and familiarity with open-source collaboration practices. Once onboarded, design activities were integrated directly with development workflows, ensuring outputs were validated against technical constraints and community feedback.

Adjustment of planned workflow edition exploration
The original proposal referenced exploratory work toward extended workflow edition capabilities (GitHub issue #2255). During implementation, investigation showed that delivering this functionality in a robust and maintainable way would require broader architectural changes than initially anticipated. Following community discussion, development effort was prioritized toward workflow validation, safeguards, and reproducibility improvements that addressed more immediate user needs. The exploratory work nevertheless informed architectural decisions made during the project and remains a potential area for future development.

Maintainer transition during the grant period
Changes in core developer availability required the redistribution of responsibilities and the onboarding of a new developer. Responsibilities related to development, release management, and community

coordination were transferred gradually, ensuring continuity while strengthening long-term project sustainability.

Expanded focus on ecosystem stability and packaging
During implementation, the team recognized that reproducibility also depended on reliable installation environments and extension compatibility. Additional effort was therefore invested in packaging improvements, CI infrastructure, and extension testing practices, strengthening long-term maintainability beyond the initially scoped feature work.

These adaptations enabled the project to meet its core objectives while improving reliability, maintainability, and community alignment, ultimately culminating in the release of OpenRefine 3.10 shortly after the grant period.

CZI Community Building: *Please describe any ways in which you feel you and/or your team benefited from general CZI and/or open science community building activities (e.g., convenings, webinars, trainings, CZI's Movement and Capacity Building offerings, etc.). If you don't feel you benefited, please describe what CZI could change or offer that would be more valuable to you.*

Members of the OpenRefine team benefited from participation in CZI community-building activities and the broader open science ecosystem supported by the program.

Participation in CZI convenings:
Team members attended multiple in-person convenings organized for CZI grantees during the project period, including meetings in North California (2022), Boston (2024), and Carlsbad (2025). These events provided opportunities to exchange experiences with peer projects working on related open-source infrastructure challenges. Discussions with other grantees were particularly valuable in sharing approaches to sustainability, governance models, and long-term maintenance of research software. Ongoing interaction through the CZI grantee communication channels, including the shared Slack workspace, further supported peer exchange between convenings. Together, these interactions helped contextualize OpenRefine's organizational evolution and informed ongoing work on governance and sustainability planning.

Engagement through the Code for Science & Society (CS&S) network
As OpenRefine's fiscal sponsor, Code for Science & Society (CS&S) connects the project to a broader network of open-source initiatives, several of which participate in CZI-funded programs. This environment enabled informal peer learning on operational practices such as fiscal sponsorship, project coordination, and community governance. These exchanges complemented formal CZI activities and supported knowledge sharing across projects facing similar organizational challenges.

Participation in community leadership training
Team members participated in community management and leadership training delivered by the Center for Scientific Collaboration and Community Engagement (CSCCE), supported by CZI. The training focused on community facilitation, inclusive collaboration practices, and sustainable community development. Lessons from these programs informed improvements to contributor engagement practices and ongoing governance work within the OpenRefine project.

Diversity, Equity and Inclusion

Briefly describe the activities undertaken during this reporting period to **increase the diversity of contributors, leaders, and users, and/or to improve the inclusivity of the project(s)**.

During this reporting period, OpenRefine continued efforts to improve inclusivity and broaden participation across contributors, project leadership, and user communities, building on initiatives established in earlier project phases, including work supported through the CZI EOSS Diversity grant (grant number 2021-237149, *OpenRefine for Everyone*).

A significant focus was placed on strengthening contributor pathways and improving onboarding practices. Through the Developer and Contributor Engagement role introduced during the final phase of the project, the team expanded community support activities such as contributor guidance, pull request review support, and regular interaction with community members. These efforts aimed to lower barriers to participation for new contributors and support a more sustainable and diverse contributor base.

User research and design activities conducted as part of the reproducibility work also incorporated feedback from a broad range of users and disciplines, helping ensure that new features address needs beyond the project's traditional technical audience. Earlier diversity-oriented outreach and interview activities informed ongoing usability and documentation improvements, continuing work initiated in prior reporting periods.

During the grant period, the community developed and adopted updated **Mission, Vision, and Values** statements alongside a broader revision of project governance practices. The first value emphasizes respect for diverse backgrounds and expertise, formalizing inclusion as a core principle guiding collaboration and decision-making. The governance updates clarified roles, responsibilities, and participation pathways, contributing to a more transparent and accessible environment for contributors and community members.

Together, these activities contributed to improving accessibility of participation, supporting contributors from varied backgrounds, and strengthening inclusive community practices around the OpenRefine project.

Key Deliverables and Project Recognition

If new preprints, journal or conference publications, methods/protocols, datasets, documentation, new code repositories were created or deposited through this grant, please provide identifiers/URLs.

Identifier (URL or DOI)	Title	Notes (optional)
https://doi.org/10.5281/zenodo.7643281	OpenRefine 3.7 Release (Feb 15, 2023)	Early grant-period release including architectural and usability improvements supporting later reproducibility work.
https://doi.org/10.5281/zenodo.11082153	OpenRefine 3.8 Release (Apr 29, 2024)	Release incorporating infrastructure and usability improvements aligned with reproducibility objectives.
https://doi.org/10.5281/zenodo.14865763	OpenRefine 3.9.0 Release (Feb 13, 2025)	Major stable release delivered during the grant, including improvements to workflow reliability and maintenance capacity.

https://doi.org/10.5281/zenodo.17042116	OpenRefine 3.9.5 Release (Sep 2, 2025)	Stabilization and maintenance release supporting adoption and long-term sustainability of project improvements.
https://doi.org/10.5281/zenodo.18791258	OpenRefine 3.10.0 Release (Feb 26, 2026)	Culminating release integrating reproducibility, workflow, and maintenance improvements supported by the grant.

Additional recognition for the open source project(s) and/or key personnel:

During the reporting period, OpenRefine continued to receive broad recognition through sustained adoption in research and applied data workflows worldwide. Analysis of scholarly publication data indicates that over 250 publications between 2023 and 2026 mention or cite OpenRefine across a wide range of disciplines and geographic regions.

These publications span fields including life sciences, biodiversity research, digital humanities, research data management, library and information science, and data-intensive computational research, reflecting OpenRefine's role as enabling infrastructure for data cleaning, interoperability, and reproducible research workflows. The diversity of languages, domains, and institutional contexts represented in these publications illustrates the project's global reach and continued relevance to modern research and open data ecosystems.

Collaboration

Have you collaborated with, benefited from, or used resources produced by other CZI associated projects or researchers (e.g., other patient or research organizations, community partners, etc.) outside of your immediate funded project team? If so, please complete the table below:

Name	Organization	Description of resource exchange or collaboration	Was this collaboration started, renewed, or continued as a result of a CZI-sponsored event/ activity? Please respond yes/no, and if yes, which event or activity.
CZI Open Science Grantee Community	Chan Zuckerberg Initiative ecosystem	Participation in CZI convenings enabled exchange with peer grantees on sustainability, governance models, and long-term maintenance strategies for open-source infrastructure projects, informing OpenRefine's	Yes, started/expanded through CZI Open Science convenings.

		organizational planning.	
CSCCE Training Program	Center for Scientific Collaboration and Community Engagement (CSCCE)	Team members participated in community leadership and collaboration training focused on inclusive community management and sustainable governance practices applied within the OpenRefine project.	Yes, CZI-supported training program.
Wikimedia Sweden / Wikimedia ecosystem	Wikimedia Foundation ecosystem	Collaboration through complementary grant activities supporting training, sustainability work, and engagement with Wikimedia contributors using OpenRefine for structured data workflows.	No. Collaboration continued independently of a specific CZI event.
Outreachy Program	Outreachy	Continued participation in open-source internship programs supporting contributors from underrepresented groups through mentorship and onboarding within the OpenRefine community.	No. Ongoing collaboration predates this grant.
Google Summer of Code (GSoC)	Google Open Source Programs Office	Participation in the Google Summer of Code program through mentorship of contributors working on OpenRefine-related projects, supporting onboarding of new contributors and strengthening the project's contributor pipeline.	No. Ongoing collaboration predates this grant but continued during the reporting period.

Deliverables Details

Deliverable 1

Milestone: One month into the project, the OpenRefine team is supplemented with one or more contractors combining UX/UI design and community management skills who will support this project.

Status: Completed

The project successfully engaged contractor support to perform UX/UI design and community engagement activities. The designer conducted user research, supported interviews with diverse user groups, and contributed to interface and workflow design for the reproducibility features.

Community engagement objectives associated with this milestone were achieved through a combination of design-led user interviews and later contributor-support activities integrated into project coordination and development roles. These efforts strengthened communication with users and informed subsequent design and development decisions.

The adjusted hiring timeline did not prevent achievement of the intended outcomes of this milestone, which focused on incorporating user-centered design and community perspectives into the project's development process.

Deliverable 2

Milestone: At the end of phase 1 (by month 4), initial user research concludes; OpenRefine consults at least 8 scholars and potential end users of the reproducibility features.

Status: Completed

User research was conducted throughout the project as an iterative and community-driven process. In the OpenRefine open-source development model, requirements and consensus emerge through sustained asynchronous discussion and continuous validation with users and contributors.

User input was gathered through multiple complementary mechanisms:

- **Direct researcher engagement:** one-to-one walkthroughs and prototype demonstrations conducted by the technical lead collected feedback from researchers and advanced users, informing architectural decisions around reproducibility features.

- **OpenRefine Forum discussions:** public asynchronous discussions played a central role in shaping features such as workflow validation, recipe reuse, and visualization

approaches.

- **Community BarCamps (2024 and 2025):** structured community events provided opportunities to discuss workflow challenges and validate priorities with active users and contributors.

- **Public prototype discussions:** including community review of early visualization prototypes, enabling iterative refinement through open feedback.

Because project governance relies on transparent, long-form collaboration, research findings are accumulated across public discussions and documentation rather than being captured in a single standalone report. The resulting community-informed research directly shaped the implementation of workflow safeguards and reproducibility improvements delivered in OpenRefine 3.9 and 3.10.

Deliverable 3

Milestone: At the end of phase 2 (by month 6), initial UI/UX design work concludes.

Status: Completed

Initial UI/UX design work was conducted and published through OpenRefine's public development channels, including GitHub and community forum discussions. Rather than being finalized in a single design phase, mock-ups, exploratory prototypes, and technical specifications were developed transparently and refined through multiple rounds of feedback from both users and core developers.

Design work addressed key areas identified during the user research phase, including improvements to operation visualization and iconography, workflow legibility, and exploration of graphical representations of workflow history. Public prototype discussions and pull requests enabled iterative feedback cycles consistent with the project's open-source governance model.

The open publication of design artifacts allowed alignment between usability goals and technical feasibility, ensuring that implemented changes were grounded in community validation. Several design outcomes were subsequently implemented and included in later stable releases, particularly improvements to operation icons and workflow readability delivered in OpenRefine 3.10.

The milestone objective, to produce and validate initial UI/UX designs through community feedback, was achieved, though through a sustained and iterative process rather than a fixed design conclusion at month six.

Deliverable 4

Milestone: At the end of phase 3 (by month 22), development concludes.

Deliverable: Release of a new stable version of OpenRefine including improved visualization, customization, reproducibility and automation features.

Development activities culminated in the release of OpenRefine 3.10.0 (February 26, 2026), following iterative delivery through the 3.9 release series. During the project, early exploration of larger architectural changes intended for a future 4.x version led to a community decision to prioritize incremental improvements within the stable 3.x release line. This approach supported ecosystem stability, backward compatibility, and earlier delivery of reproducibility improvements to users.

While the original proposal also envisioned deeper workflow automation capabilities, including more formal support for non-interactive execution, community discussions led to prioritizing improvements that strengthened workflow reliability, reuse, and transparency. As a result, automation-related work focused on foundational infrastructure that enables safer, more adaptable workflows was deferred to future development outside the scope of this grant.

Key outcomes delivered through these releases include:

- Workflow pre-validation safeguards preventing partially applied recipes and improving reproducibility integrity (PR #7095).
- Column dependency metadata and recipe abstraction enabling prediction of workflow effects prior to execution (PR #7056, PR #7116).
- Column mapping interface supporting reuse of workflows across datasets (Issue #7500).
- Workflow safety and undo improvements, including notifications when applying recipes and safeguards preventing loss of project history (#7172 and related fixes).
- Improved handling of long-running operations, reducing the risk of data loss during interrupted processes.
- Workflow legibility improvements, including updated operation iconography and groundwork for visualizations (#6416, #6538).

Graph-based workflow visualization remained at a prototype stage following community consensus to defer full integration, while enabling infrastructure and design groundwork were completed. Together, these changes substantially improved the reliability, interpretability, and reproducibility of OpenRefine workflows, establishing technical foundations for future automation work.

Deliverable: Continuous Integration test suite includes tests of these features.

Status: Completed

Testing infrastructure expanded significantly during the grant period to support reproducibility and release reliability:

- Java test methods increased from **1,259 to 1,790** (+42%).
- Cypress end-to-end workflow tests increased from **242 to 292** (+21%).
- Continuous Integration workflows received sustained improvements across multiple development cycles.

In addition, testing practices were extended to the extension ecosystem. Development builds were tested against community extensions prior to the 3.10 release, and extension testing guidance was published to reduce ecosystem breakage and improve release stability.

Deliverable: Installer packages include the new software version.

Status: Completed

Distribution and packaging processes were strengthened across supported platforms to ensure reliable and reproducible installations:

- Continued collaboration supporting Debian packaging.
- Maintenance of official container images through the OpenRefine containers project.
- Improvements to Windows and macOS packaging and distribution workflows, including enhanced signing and release processes.

These improvements support reproducible installation environments, an important requirement for reliable reuse of data workflows.

Deliverable 5

Milestone: Documentation updated.

Status: Partially completed

Documentation activities evolved during the project to reflect the iterative delivery model adopted for development work. Rather than producing a single documentation expansion at the end of the project, documentation updates were integrated continuously alongside releases and community development processes.

Technical and contributor-facing documentation saw substantial improvement during the grant period. Updates included expanded onboarding guidance for developers, documentation supporting extension testing practices, and improved transparency around development workflows and release processes. These materials supported contributor engagement and helped sustain participation as the project introduced reproducibility-related changes.

User-facing documentation for reproducibility features was delivered primarily through detailed release notes, public progress reports, walkthrough presentations, and recorded demonstrations accompanying the OpenRefine 3.9 and 3.10 releases. These resources provided practical explanations of new capabilities as they were introduced, consistent with documentation practices commonly used in active open-source projects.

While the originally planned expansion of dedicated end-user documentation and reusable tutorial videos was only partially completed, the project significantly improved technical documentation and onboarding resources, strengthening long-term maintainability and enabling future documentation work to build on clearer contributor processes.

Deliverable 6

Community-focused deliverables

Status: Completed

Project outcomes were disseminated through multiple public activities conducted throughout the grant period. These included OpenRefine BarCamp events (2024 and 2025), public presentations and walkthroughs accompanying major releases, and ongoing open development discussions conducted through community forums and online events. These activities provided structured opportunities to present reproducibility improvements, gather feedback, and engage users working in research and data stewardship contexts, including biomedical and life-science domains.

In addition to direct presentations, public project reports, and release walkthrough materials, documented reproducibility improvements were introduced in OpenRefine 3.9 and 3.10. Continued academic adoption across research disciplines, including data stewardship, biodiversity research, epidemiology, and health data analysis, demonstrates sustained engagement with communities concerned with reproducible scientific workflows.

Biomedical community representation

Status: Completed

During the grant period, OpenRefine updated and expanded its advisory governance structure to better reflect the diversity of its user communities. This included strengthening representation from research-oriented communities, including biomedical and scientific data practitioners. These governance updates improved alignment between project decision-making and the needs of research users and contributed to broader community representation within project leadership.

Code of Conduct

Status: Completed

OpenRefine operated under an established Code of Conduct prior to the start of the grant. During the project period, governance documentation and community practices were clarified and strengthened, improving transparency around roles, expectations, and participation processes. These updates reinforced inclusive collaboration practices and supported a welcoming environment for contributors and users.

Closing questions

Last screen after upload of the report

Please share any additional comments on the progress reported in your deliverables section, or any additional relevant progress beyond the specified deliverables.

Beyond the specific deliverables, the grant strengthened OpenRefine's long-term sustainability as open research infrastructure. Development practices matured through expanded testing, improved release engineering, and strengthened packaging and extension compatibility workflows, improving reliability for users working with reproducible data workflows.

The project also advanced governance and coordination practices during a period of maintainer transition, supporting continuity and broader community participation. Experience gained during the grant clarified a sustainable delivery model combining clearly scoped technical milestones with ongoing community stewardship, positioning the project to continue advancing reproducibility improvements while maintaining independent governance and community alignment.

Major changes in scope or project plan: Please describe any problems encountered and steps taken to overcome them.

Several adjustments were made during the project to better align implementation with community needs and the realities of maintaining a mature open-source infrastructure project.

The proposal initially described sequential research, design, and development phases; in practice, these activities proceeded iteratively, allowing continuous validation through user discussions, walkthroughs, and community events. Early technical exploration toward a larger 4.x architectural evolution was also reassessed through community governance discussions,

leading the team to prioritize incremental improvements within the stable 3.x release series. This approach enabled reproducibility improvements to reach users sooner while maintaining compatibility with existing workflows and extensions.

Community feedback further shifted priorities toward workflow reliability and safeguards rather than more experimental visualization and automation features. Additional challenges included delayed recruitment of UX/UI support and a transition in core development responsibilities. These were addressed through gradual onboarding, redistribution of responsibilities, and increased investment in testing and release processes, ensuring continuity and successful delivery of project outcomes.

CZI Community Building:

Please describe any ways in which you feel you and/or your team benefited from general CZI and/or open science community building activities (e.g., convenings, webinars, trainings, CZI's Movement and Capacity Building offerings, etc.). If you don't feel you benefited, please describe what CZI could change or offer that would be more valuable to you.

The OpenRefine team benefited from participation in CZI convenings and the broader open science ecosystem supported by the program. In-person grantee meetings provided opportunities to exchange experiences with peer infrastructure projects addressing similar sustainability, governance, and funding challenges. These discussions informed aspects of OpenRefine's organizational and delivery strategy.

Team members also participated in community leadership training delivered by the Center for Scientific Collaboration and Community Engagement (CSCCE), supported through CZI initiatives. This training informed improvements to contributor engagement, governance clarity, and facilitation practices within the project.

More broadly, participation in the CZI open science community strengthened peer connections and knowledge sharing, supporting OpenRefine's evolution as sustainable open research infrastructure.