

2019 arXiv Roadmap

The arXiv roadmap is a living document for the arXiv team to communicate our annual priorities to the community. Items are listed in approximate priority order, and are subject to change based upon consideration of input from the arXiv advisory boards and other stakeholders, assessment of new opportunities and initiatives as they arise, and progress on next generation arXiv development.

Our overall goal of 2019 is to continue with the [Next Gen arXiv](#) (arXiv- NG) development plan to migrate from the existing arXiv system ("Classic Renewal"). We aim to progressively evolve arXiv into a modern and architecturally sound software system while maintaining the level of consistency and availability of the system as a whole.

Related documents: [arXiv software releases](#) || [arXiv development blog posts](#)

Technical

Launch NG Submission Interface: In 2018, we began work on a new user interface and API for submitting e-prints that will make it easier for us to make improvements and provide new features. In 2019, we will launch the new interface, first for public beta evaluation and eventually to replace the classic submission interface.

Status: In progress

New Registration and Account Management Interfaces : We are rebuilding the user registration and account management interfaces to provide a more user-friendly experience, and to integrate with the NG submission system, API gateway, and other new components.

Status: In progress

Better Support for Bulk/Programmatic Access: We are rolling out a new API gateway for accessing arXiv metadata and content programmatically. This includes a refreshed search API, and better tools for accessing e-print content (e.g., for text and data mining).

Status: In progress; Search API in alpha

Author Name Disambiguation: Users have increasingly noticed the limitations of how we represent author names in e-print metadata (see [this blog post for details](#)). We will introduce new features that make it easier for submitters to associate author and institution names across e-prints, and improve the precision of author name queries in search.

Status: Not started

Under the Hood: Modernizing the Core Announcement Record: We have begun to outgrow the way that we store and access the canonical record of e-print submissions. We are dedicating a significant effort in 2019 to migrating the canonical record to a system that can scale with arXiv over the next decade, and make it easier to handle higher volumes of traffic.

Status: In progress

Continued Improvements to Search: We will continue to improve our search interfaces, and introduce new features throughout the year. For example, we plan to include full text search as part of the native search interface.

Status: In progress

Better Support for External Links: Users expect to be able to link to relevant source code, datasets, and other resources from their arXiv e-prints. We currently support limited linking from e-print abstracts and submission comments. We plan to release an interface that allows users to curate external links to their e-prints in a more organized fashion without requiring replacement submissions.

Status: Deferred until 2020

Archiving arXiv: In order to ensure that we have a broad network of preservation strategies, in 2018 we began assessing several archival service providers with the intent to establish a third-party dark archive for our core e-print metadata and content. The dark archive will help to ensure perpetual long-term access to arXiv content in the event of a sustained outage to the [arXiv.org](#) platform. In 2019, we will begin working with one or more providers to implement this dark archive solution.

Status: In progress

Maintenance and Operational Support: A significant amount of developer effort is devoted to a range of activities that ensure that the service continues to run smoothly. Activities include (but are not limited to) server maintenance and troubleshooting, service monitoring, analytics, internal support requests, bug fixes, and ad hoc meetings and conferences

Status: Ongoing with additional allocation of effort in Q4

Moderation and User Support

Code of Conduct: We are developing a Code of Conduct, based on shared values and shared responsibility within the arXiv community. Throughout our process, we will engage with stakeholders, including staff, advisers, moderators and authors. The Code of Conduct is intended to complement arXiv's [Operating Principles](#) as a broader set of community values.

Status: Complete https://arxiv.org/help/policies/code_of_conduct

Help Page Updates: We plan to update and organize our help pages and other reference content to improve the user's access and experience. We will first migrate help pages to a new repository, making future updates easier, then we will then re-organize and update the site map and help search.

Status: Migration complete; re-organizing pages is ongoing

Planning Improvements for NG: This year we will create plans for additional updates to submission and moderation systems ahead of future NG software development in those areas. Some topics we are considering: support for various content types (proceedings, etc.), improved support for bulk and proxy submissions, explore alternatives for handling and displaying withdrawn submissions, improved workflows and communication for moderators, and update internal documentation on user account policies and identify ways to automate actions.

Status: Ongoing

Business & Governance

arXiv Strategic Vision Planning: We will develop a strategic plan to articulate arXiv's mission, vision, and values. Based on this scope, we will start creating a 5-year plan to identify priorities, new development areas, and set strategic directions to create a framework for our fundraising, collaboration, and communication processes.

Status: Initial assessment then deferred until after Executive Director is hired

arXiv vis-à-vis Scholarly Communication Ecology: arXiv is a part of the broader scientific communication ecology and needs to monitor developments in related areas (e.g., public access domains such as Plan S, implementation of standards such as DOI & Orcid, etc.) and assess implications to ensure that arXiv will continue to meet the users' needs.

Status: In progress

Society-arXiv Collaboration Pilot: arXiv and scientific societies have a shared mission to serve the scientific community, and they hold a common interest in working together to ensure that arXiv has a sustainable future that is well positioned to create an even stronger value proposition for researchers. We will implement [a pilot Society membership tier](#) and develop an assessment process and success criteria to help us assess the outcomes.

Status: Completed

Fundraising: We will strengthen and further diversify our fundraising strategies to support arXiv NG plan and align with the development of a 5-year Vision (see above), including inviting additional arXiv users (libraries, research organizations, and corporations) to become members and increase individual and corporation giving.

Status: Ongoing

arXiv Labs: In 2018 we created an R&D environment in arXiv Labs to explore new arXiv features and understand technical, policy, maintenance and user support needs prior to integration with the production system. We will continue our exploration and look into identifying alternative revenue sources in support of arXiv Labs.

Status: In progress

Priority Setting & Communication: We will review our current processes for priority setting, feedback seeking, and communication to improve the transparency of how decisions are made and implemented.

Status: Ongoing

See: [arXiv Roadmaps](#)