

Transition FAQ: Move to Cornell Computing and Information Science

What is arXiv?

Started in August 1991 by Paul Ginsparg, [arXiv.org](http://arxiv.org) is internationally acknowledged as a pioneering digital archive and open access distribution service for research articles. The e-print repository has transformed the scholarly communication and knowledge dissemination of multiple fields of physics and plays an increasingly prominent role in mathematics, computer science, quantitative biology, quantitative finance, and statistics, recently adding electrical engineering and systems science (eess) and economics (econ) as new subject domains. arXiv is truly a global resource, with almost 90% of supporting funds coming from sources other than Cornell and 70% of institutional use coming from countries other than the USA.

What's happening?

In January 2019, the administrative and operational responsibilities for the IT infrastructure and user support collaboration for arXiv is moving from [Cornell University Library](#) (CUL) to [Cornell Computing and Information Science](#) (CIS). Cornell has hosted arXiv since 2001 when its founder, Paul Ginsparg, left the Los Alamos National Laboratory and joined the Cornell faculty in the Physics and Information Science departments. Running the service has always involved a collaboration with CIS as Professor Ginsparg and other CIS faculty lead R&D efforts through several NSF grants and other funding sources, and contribute to the development of moderation policies. This transition is a natural stage in the evolution of arXiv, required for optimum service delivery and infrastructure sustainability.

Will researchers and arXiv users be affected by the move?

As far as researchers and arXiv users are concerned, the stewardship transition will be seamless and unnoticeable, not affecting any aspects of the operation. There are no shared financial lines between arXiv and other CUL programs, so from an administrative perspective, we expect the move to be straightforward. Although the arXiv team is currently situated at the Library, the operation is quite portable and independent of the Library's organizational and technical infrastructure. Both the development and operations teams have successfully operated in a distributed mode, with both on-site and remote participants. For the next year the arXiv team will continue to reside at their current location at Olin Library to allow time for space planning at CIS.

What is Cornell's commitment?

Cornell's commitment to arXiv is unchanged. Cornell's mission is to discover, preserve, and disseminate knowledge, to educate the next generation of global citizens, and to promote a culture of broad inquiry throughout and beyond the Cornell community. arXiv is well aligned with this mission, and Cornell commits to ensure that arXiv, as a prominent scientific forum, is sustainable and continues to meet scientists' needs. Through this administrative change, Cornell reaffirms its continuing commitment to the service's stability and openness and to its core values of transparency, accountability, and community engagement. The University will continue to provide a cash subsidy of \$170,000 per year in support of arXiv's operational costs, in addition to making an in-kind contribution to cover some indirect costs, which currently represents 37% of total operating expenses. It continues to be a great privilege for Cornell to host arXiv. We believe that arXiv is a public good and provides an important infrastructure to support scientists from all over the world to rapidly communicate their research and seek timely feedback from their colleagues.

How does arXiv align with CIS's mission and values?

arXiv aligns very well with CIS's mission and values. CIS will provide a dynamic technology and sociocultural context for the operation of arXiv as a scholarly scientific enterprise:

- Cornell's CIS not only spans core contemporary IT realms, such as networks, systems, robotics, and machine learning, but is also concerned with the human aspects of computing, such as human-computer interaction and human-centered design.
- CIS blurs the boundaries of the traditional college and is a place of radical collaboration where computer science thrives side-by-side with emerging fields like computational social science and human-centered design.
- CIS has taken a leadership role in creating and collaborating with Cornell Tech, a groundbreaking initiative that infuses technology innovation with business and creative thinking.
- With a mission to lead the research and educational agenda for the information age, CIS creates an environment for arXiv that will further strengthen the arXiv team's efforts to build a sustainable infrastructure and service model.
- Paul Ginsparg, arXiv's founder who continues to contribute to the service, is a faculty member both at CIS and the Physics department.

Why change?

In this phase of arXiv's development, the arXiv Program Director, the University Librarian, and the Dean of CIS believe CIS is better positioned to move arXiv forward because of the closer ties CIS has with the computer and information science community as it relates to technical infrastructure, tools and services, and with opportunities for identifying new funding streams. The original arXiv business model approached arXiv as a production system with the goal of "keeping the lights on," not factoring in the need to maintain steady R&D efforts. An important component of arXiv's sustainability model is the periodic assessment of the service's technical, operational, and business needs to ensure an adequate level of stewardship and to keep up with new requirements that emerge as the service expands in scope and usage.

From the users' perspective, as demonstrated by the 2016 user study, arXiv continues to be a successful, prominent subject repository system serving the needs of many scientists around the world. However, under the hood, the service has been facing significant pressures as the demand for services grows, scientific communication evolves, and the legacy architecture ages. The landscape of scholarly communications is changing rapidly, and arXiv must be able to keep pace with the evolving ecosystem by refining its current sustainability model in order to leverage advances that will benefit arXiv users. Responding to these concerns, in 2016 the arXiv team embarked on an initiative to modernize arXiv's architecture (arXiv-NG) and move the service into a new business model that is appropriate for its scale and the growth trajectory. The arXiv team has benefited from the CUL's stewardship and is well positioned for a smooth transition by having built a strong, cohesive team and a blueprint for technology renewal.

What are the advantages of operating arXiv from CIS?

Transitioning the stewardship of arXiv to CIS will offer a number of advantages:

- arXiv is a service, a mode of scholarly communication, a corpus of information, and a community. Situating arXiv at CIS will bring a range of R&D opportunities to support the enhancement of the service, use it as a data source for exploring a rich range of scholarly communication trends, and expand arXiv's potential as a nexus for building digital communities.
- From work in human-computer interaction to computer-aided sustainable design, CIS will provide the arXiv team with a rich context to further strengthen the team's efforts to build a sustainable infrastructure and service model for arXiv. For the team, the focus of "innovation" is not merely on adding new features but creating an open, scalable, and resilient service framework to ensure arXiv will continue to meet scientific communities' evolving needs and continue to be a reliable and trusted information provider.
- Positioning arXiv within CIS will open up collaborative opportunities between subject domain and computer and information scientists to seek large grants from U.S. and international agencies (e.g., NSF) in support of sustainable infrastructure development and innovation (including data science).
- Information science as a discipline has grown out of the library science domain, and we will continue to consult with both CUL and other member libraries on many issues.
- arXiv is moving toward a more open, evolvable, and collaborative codebase. A close relationship with CIS may increase the potential to realize the benefits of an evolutionary architecture through collaborations with researchers in computer and data sciences, which may catalyze development of a long-term sustainability model involving community development.
- There will be opportunities for extending arXiv's technical expertise. For instance, we'll be able to hire postdocs through academic appointments to blend technical and academic skills. This is one of the strategies used in similar services that are situated in academic units (e.g., Astrophysics Data System (ADS) at Harvard and Inspire at CERN).
- arXiv offers a potential research corpus (papers, usage logs, and applications) for data science and is a high-value asset for studying the social aspects and trends in science. CIS would provide a sandbox and laboratory setting for the operation to foster data science and scientific communication research based on the corpus, while adhering to arXiv's values and policies related to privacy and confidentiality of use.
- Through CIS, we anticipate engaging in new partnerships to add new revenue sources and further strengthen arXiv's current business model, invigorating community commitment to maintaining and developing arXiv as a public good.

Will there be any changes to the current Organizational and Governance Model?

Reporting lines will be affected by this change but staffing levels will not. Starting in January 2019, CIS will hold the overall administrative and financial responsibility for arXiv's operation and development, with continuing guidance from its Member Advisory Board (MAB) and its Scientific Advisory Board (SAB). The arXiv team will maintain the same configuration and will continue to run arXiv's technical infrastructure, oversee the moderation and user support functions, lead development and implementation of policies and procedures, and establish and maintain collaborations with related initiatives to improve services for the scientific community through interoperability and tool-sharing. Instead of reporting to the University Librarian, the arXiv Program Director will report to the Dean of CIS, as she continues to collaborate with the Scientific Director to oversee the operation. See [2019 org chart](#)

What will be the role of CUL in the new stewardship model?

Similar to other academic units, arXiv will benefit from the expertise of the Library staff, especially in the service areas of metadata, copyright, and scholarly communication policies and best practices. For instance, the arXiv team will continue to collaborate with CUL's Physics, Astronomy, and Mathematics Librarian, in implementing and interpreting usability studies, especially as new features are developed. As the arXiv team explores third-party archival agencies, such as Portico, we will continue to rely on the CUL's archival system to support arXiv's preservation. Also, for the next year, the arXiv team will continue to reside in their current location at Olin Library to allow time for space planning at CIS.

Are there any examples of running scientific repositories such as arXiv from academic units?

There are several successful examples of how this can be achieved; for instance, the Astrophysics Data System (ADS) is operated within the Harvard-Smithsonian Center for Astrophysics. As noted previously, the arXiv development team is portable, and the technology infrastructure is moving to the cloud, incrementally. Operational dependencies for technical infrastructure would not be tied tightly to an institution.

How will the member libraries continue to participate in arXiv's governance?

The engagement of member libraries in the arXiv's governance is essential and the move will not have any impact on the strong alliance we've build with our library partners. The financial support from libraries and research laboratories will continue to be an important component of arXiv's business model. In addition to generating revenues to support the operation, it helps us be eligible for the \$300,000 per year matching funds from the Simons Foundation. The arXiv membership system is instrumental in leveraging the expertise and commitment of librarians in arXiv's oversight as it provides an essential public infrastructure in support of open science. Member libraries will continue to participate in the arXiv's governance through the Member Advisory Board, which provides input for project prioritization, new service offerings, financial planning, use of discretionary funds, future technical developments, and policy decisions. See: [2018-2011 Sustainability Plan](#).
