



ITHAKA S+R

Innovation and Shared Infrastructure in the Adoption of AI

Roger C. Schonfeld

November 2024

Key Takeaways

1. Shared infrastructure is a vital category for publishers and libraries.
2. Generative AI is going to be provided mostly via shared infrastructure rather than through services that individual publishers or libraries build themselves.
3. Generative AI may scramble the sector but ultimately ensuring a competitive marketplace will be essential.

Background and Context

What Is Digital Infrastructure?

For libraries

- ILS/LSP
- Library discovery
- Interlibrary sharing
- Institutional repository
- And more

For publishers

- Manuscript submission, research integrity, and peer review
- Hosting
- And more

Important interconnections with one another and with research infrastructures, teaching infrastructures, and business infrastructures.

Digital Infrastructure and Competitiveness

- Digital infrastructure is the essential foundation on which universities, libraries, and publishing organizations conduct operations and provide services.
- For libraries and publishers, digital infrastructure is foundational, enabling innovation, collaboration, and competitiveness
- When we underinvest in infrastructure, it impedes innovation and causes us to fall behind in their service offerings.

Scale Effects

- Digital infrastructure is best provided at a scale where the level of investment can provide the quality required.
- The largest organizations can develop their own digital infrastructure. In publishing, consolidation strategies benefit from the opportunities for big publishers to control their own infrastructure.
- Smaller organizations, including the vast majority of publishers and libraries, are existentially reliant on shared infrastructure. Without it, they cannot provide modern services or meet modern user expectations.

Shared Investment

- Any model to launch and maintain shared infrastructure must:
 - generate access to capital for launch and ongoing reinvestment,
 - provide trustworthy services,
 - innovate strategically,
 - ensure agility and flexibility.
- Several models, each with strengths and challenges, include:
 - commercial,
 - not-for-profit,
 - open-source.

The First Digital Transformation

- The first digital transformation saw a massive shift from paper to digital, but otherwise publishing and libraries retained many of the structures, workflows, incentives, and outputs that characterized the print era.
- One notable exception was that libraries steadily lost control of discovery as a result both of Google Scholar and websearch more broadly.

The Second Digital Transformation

- In the second digital transformation, many of the structures, workflows, incentives, and outputs that characterized the print era are being revamped in favor of new approaches that bring tremendous opportunities, and also non-trivial risks, to scholarly communication.
- Now, AI is scrambling everything. It may generate a third digital transformation.

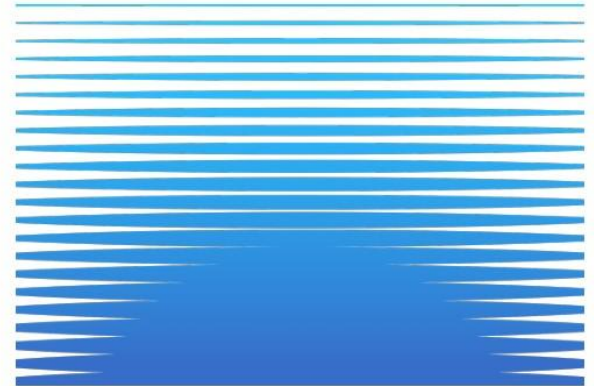
The Second Digital Transformation of Scholarly Publishing



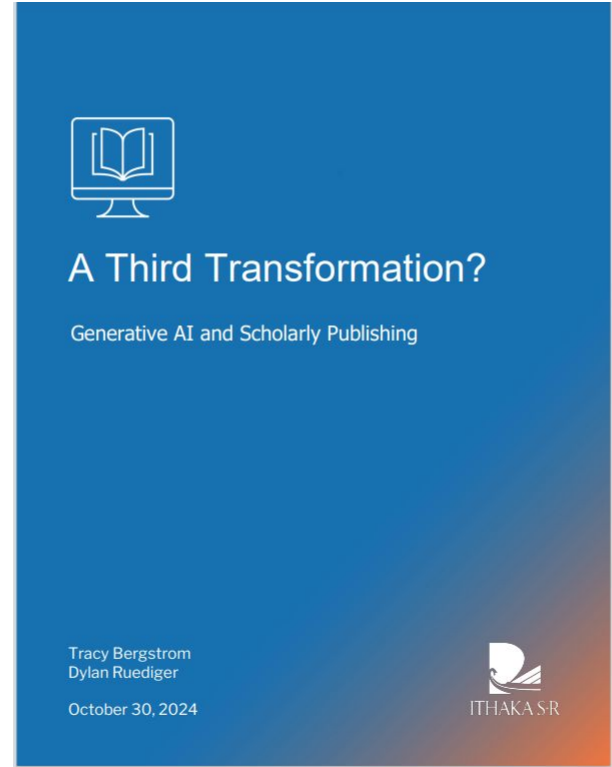
The Second Digital Transformation of Scholarly Publishing

Strategic Context and Shared Infrastructure

Tracy Bergstrom
Oya Y. Rieger
Roger C. Schonfeld



A Third Transformation? Generative AI and Scholarly Publishing



A Framework for Situating AI Services

Generative AI Is Not a Service

- Generative AI is a set of underlying technologies, not a product.
- The key question for libraries and publishers is what services will generate value and therefore be sustainable.
- The types of services we are seeing so far fall into several categories.

Opportunities

- There are two fundamental types of opportunities
- First, publishers and libraries can build tools and services that reach scholarly and scientific users
- Second, content owners can share their content with third parties to build their tools and services

Opportunities

- There are two fundamental types of opportunities
- First, publishers and libraries can build tools and services that reach scholarly and scientific users
- Second, content owners can share their content with third parties to build their tools and services

Building Our Services....

- Scholarly written outputs --
 - Better writing, translating, and editing
- The integrity of scholarly research --
 - Streamlined peer review and editorial processes
- The discovery of scholarly and cultural materials --
 - Faster metadata creation and a better interactive discovery experience
- The understanding and acquiring of knowledge --
 - Summarization and synthesis

Scale

Developing services that utilize generative AI requires strong engineering and product leadership

- This expertise can be difficult to source at individual libraries and publishers
- Scale is a solution to spreading the cost across multiple institutions

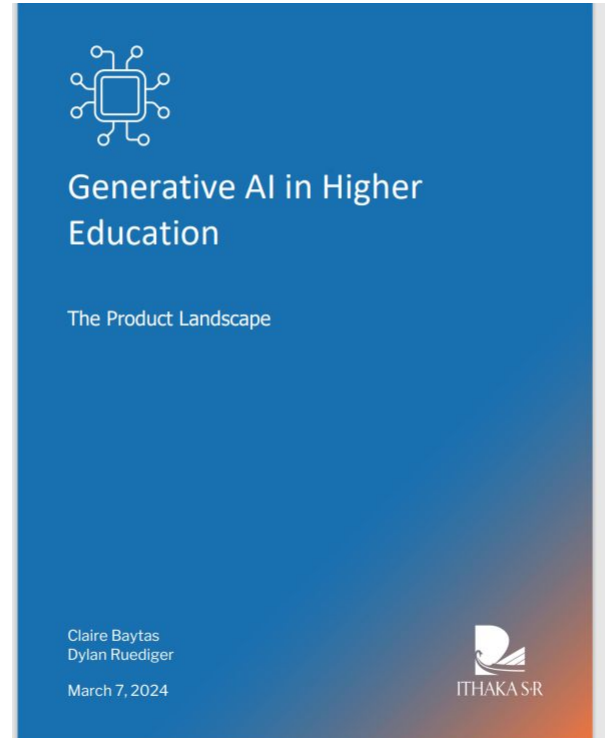
Additionally, LLM fees can be substantial and unpredictable

- Larger providers can build business models that spread this expense more evenly and predictably

Platforms, Infrastructure, and Scale

- Many types of organizations will contribute
 - Publishers, libraries, platform providers, and other intermediaries
- Provided mostly via shared infrastructure rather than through services that individual publishers or libraries build themselves.
- Most organizations will need to rely on shared infrastructure in order to implement AI tools in ways that preserve their value

Generative AI in Higher Education: The Product Landscape



Opportunities

- There are two fundamental types of opportunities
- First, publishers and libraries can build tools and services that reach scholarly and scientific users
- Second, content owners can share their content with third parties to build their tools and services

Sharing Content with Others

- AI providers primarily focusing on consumer services like search
 - Google, OpenAI, Apple, etc.
- Field-specific providers
 - Pharmaceutical and other scientific research companies
- Publishers face real strategic dilemmas in considering whether -- and under what terms -- to provide their content to third parties.

Tracking the Licensing of Scholarly Content to LLMs



Once Again: Scale Matter

- Easy for the largest publishers to engage.
- Not all of them want to -- yet.
- More challenging for smaller publishers.
- Another problem of scale.

The Marketplace

Competitiveness

The marketplace for infrastructure must not only keep up with and enable -- but ideally help to trailblaze -- library and publisher strategic directions

Infrastructure only keeps up when:

1. It is viewed as a public good and the recipient of adequate public investment; or
2. There is a healthy marketplace for its provision.

Marketplace Conditions

Infrastructure has substantial switching costs, and providers can become just as “locked in” as customers

Under such conditions, once successful, providers can reduce their investments to cash in on profits.

In the long run, no single provider matters, if we have a healthy marketplace, but in the near term many customers feel trapped

Generative AI

The marketplace for the types of infrastructure providers discussed today is developing.

It includes both well established providers and startups, as well as a few not-for-profit organizations.

Will library shared models develop?

Governance and Business Models for Collaborative Collection Development



Governance and Business Models for Collaborative Collection Development

Tracy Bergstrom
Oya Y. Rieger
Roger C. Schonfeld

August 1, 2024

Concluding

Caveats and Considerations

- Generative AI can introduce risks to
 - Research integrity
 - Learning outcomes
- And of course there are larger social and policy questions

Key Takeaways

1. Shared infrastructure is a vital category for publishers and libraries.
2. Generative AI is going to be provided mostly via shared infrastructure rather than through services that individual publishers or libraries build themselves.
3. Generative AI may scramble the sector but ultimately ensuring a competitive marketplace will be essential.



ITHAKA S+R

Thank you

Roger C. Schonfeld

Vice President

roger.schonfeld@ithaka.org