

The new dimension in scholarly communications: How a global scholarly community collaboration created the world's largest linked research knowledge system

Ashlea Higgs*

ÜberResearch/Digital Science, 625 Massachusetts Ave, 2nd floor, Cambridge, MA 02139, USA

Abstract. Digital Science and more than one hundred global research institutions have spent the better part of the last two years collaborating to solve three distinct challenges in the existing research landscape: (1) Research evaluation focuses almost exclusively on publications and citations data; (2) Research evaluation tools are siloed in proprietary applications that rarely speak to each other; and (3) The gaps amongst proprietary data sources made generating a complete picture of impact extremely difficult (and expensive). The goal of this collaboration amongst publishers, funders, research administrators, libraries, and Digital Science was to transform the research landscape by breaking down the barriers to discovery and innovation.

1. Introduction

The National Federation of Advanced Information Services 2018 conference sought to examine the changes to the scholarly landscape brought about by increased adoption of open access, taking an in-depth view of this information transformation, and the drive towards a more open, global, and collaborative future.

The lightning talks session offered sensible solutions to problems; overviews of new projects, the latest news about “open,” and other critical issues to spark debate. Managing Director of ÜberResearch, Ashlea Higgs, updated the conference on the development of new research information platform; Dimensions.

2. Lightning talk

Dimensions is the new research information platform from Digital Science. In both its features and approach it is very different from existing databases, which provide access to one or two aspects of research information, but may struggle to offer a complete picture.

Dimensions was created to address two specific barriers in the provision of research information. First, research evaluation has traditionally focused almost exclusively on publications and citations data, but

*E-mail: ashlea@uberresearch.com.

relevant information comes in many more forms, such as clinical trials data, policy documents, and mentions in mainstream press. Second, research information has been held in proprietary systems that do not easily communicate with one another – creating silos of data when the research community needs a complete picture.

Together these barriers created a context for research information in which gaps between proprietary data sources made establishing a comprehensive view of impact for scholarly work – both at the level of individual articles and across an organization – both difficult and expensive.

Along with over one hundred partners from organizations across the world Digital Science set out to transform the research information landscape by addressing the challenges resulting from expensive, siloed research evaluation data.

The result is a new research information platform, Dimensions. Dimensions is ...

2.1. *Open*

Explore the openly-available publication records and related metrics in the new platform at app.dimensions.ai.

2.2. *Collaborative*

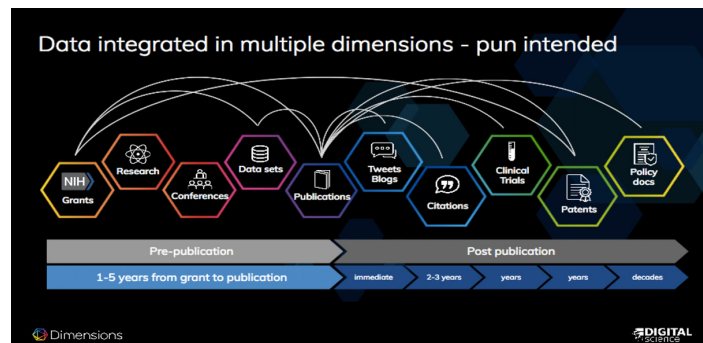
For over a year before the launch of the new Dimensions database, Digital Science worked with more than one hundred research organization and funder development partners to develop the new platform. Development partners came from leading institutions across the world and represented the entire research landscape, so the platform has been built on real world use cases.

2.3. *Transformative*

Which brings us to the details of the new platform. One of the most important elements of Dimensions is that it brings different types of research information together, and links them consistently to provide a rich resource full of connected information – from the beginning of the research process to the end, from funding through to output. Dimensions enables users to explore:

- Over \$1.3trillion of funded grants
- 96 million+ publications
- Over 1 billion citations
- 400,000 clinical trial records
- 36 million patents
- 350,000 policy documents

Anyone interested in learning in detail about the inclusive approach to data within Dimensions can read *A Guide to the Dimensions Data Approach* [1]. By collecting together all this data we can provide a view of the research landscape that begins to look like this:



The platform links grants awarded at the beginning of the research process to different kinds of research output that occur at unique points in the research lifecycle.

3. Data enrichment

When users search a platform for information they expect consistent, reliable results. Dimensions is based on disparate information sources that take very different approaches to metadata, so a data enrichment process is vital to even out inconsistencies and ensure that the platform returns high-quality results.

The data enrichment process for Dimensions includes:

- Organization identification
- Concept extraction
- Researcher disambiguation
- Reference extraction
- Categorization

All this means that users can trust and rely on the information they are seeing when they search for researchers, organizations, concepts, classifications, references, and metrics.

4. Data and links driving dimensions

With functionality driven by the collaborative development partner program, inclusive data, and enriched metadata, Dimensions currently offers:

- Over 1 billion citations between publications
- 38 million links between publications and funding
- Almost 400,000 connections between clinical trials, publications and grants
- Altmetric attention data for over 9 million publications
- 20 million links between patents, publications and grants
- Over 350,000 policy document records

Dimensions allows people with different roles in the scholarly community to interrogate the data in different ways. It has a clear interface which allows people to drill down into the data, and a number of

ways to communicate it in a visually-engaging way, which makes it easier to understand and act upon. For funders and research organizations, it can be particularly useful for:

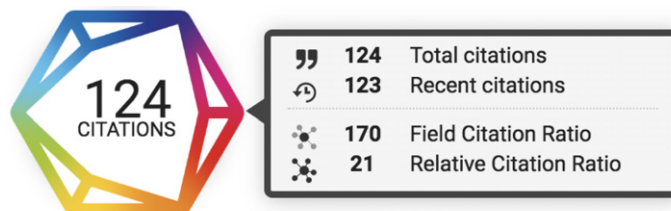
- Evaluating the outcomes of research projects
- Tracing funding, faculty, and research trends globally
- Informing future research strategy
- Identifying collaboration or growth opportunities

For inspiration on the many uses of Dimensions Data, take a look at the report: *Dimensions – A Collaborative Approach to Enhancing Research Discovery* [2].

5. Dimensions badges

Bringing different types of content from different stages in the research cycle together, and creating new connections between them, essentially means creating brand new information. It's important for users to be able to see metrics in a fast, visual way to aid their understanding of what the data is telling them.

The Dimensions badges do just this. They offer a simple citation score, which opens into a details page that offers a more sophisticated look at citations for any particular article, providing the Field Citation Ratio, Relative Citation Ratio, and a breakdown of citing research publications.



6. API

As part of our commitment to putting control back into the hands of the research community, the full Dimensions API allows for flexible use of this enriched data.

It can be used without constraints for internal purposes – particularly useful for enriching administrative systems or analytical software. Our custom [Dimensions Search Language](#) makes it easy to use even with little technical knowledge. Using these APIs, you can extract data for reports not covered by the application not only as raw data, but also as the foundation for your report – meaning constructing complex analysis and visualizations becomes an efficient and effortless task. This means that users can innovate on both the data and the metrics, creating in-house applications based on the Dimensions API that fulfill their unique needs. We're most intrigued to find out what users do with the API that we're not expecting. Innovation will come from asking new questions of the data.

7. Conclusion

Dimensions was created to solve a number of the following challenges facing the scholarly community: an over reliance on publications and citations data; research evaluation tools that were siloed in proprietary applications; and gaps in data sources that made generating a complete picture of impact both difficult and expensive. We have created a new platform that is *open*, and *collaborative*. We hope it will most transformative when new metrics and uses are led by the scholarly community.

About the Author

Ashlea Higgs is one of the founders of ÜberResearch, part of Digital Science, the team behind Dimensions. They look after Government and Funders for Digital Science. Email: ashlea@uberresearch.com.

References

- [1] G. Bode, C. Herzog, D. Hook and R. McGrath, *A Guide to the Dimensions Data Approach*, Digital Science, January 2018, <https://www.digital-science.com/resources/portfolio-reports/a-guide-to-the-dimensions-data-approach/> (last accessed 27 June 2018).
- [2] J. Adams, P. Jones, S. Porter, M. Szomszor, H. Draux and I. Osipov, *Dimensions – A Collaborative Approach to Enhancing Research Discovery*, Digital Science, January 2018, https://figshare.com/articles/Dimensions_A_Collaborative_Approach_to_Enhancing_Research_Discovery/5783160 (last accessed 28 June 2018).