Crossref at 20 years: what do the community need?

Rachael Lammey
Crossref, Oxford, United Kingdom

Abstract
Purpose: The 20th anniversary of Crossref is a good point to evaluate where Crossref stands with the communities it supports so that it can be in a position to serve their needs in the future.
Methods: This evaluation took the form of a survey and a series of interviews run by Shift Learning in mid-2019.
Results: Results are presented in the form of a report authored by Shift Learning which is discussed in this paper.
Conclusion: Overall, Crossref is appreciated and provides value for the scholarly community. However, it needs to make sure that it continues to serve key stakeholders, ensure that core systems work smoothly for all members and that they balance the needs of its different sizes of members and those who subscribe to or use the Crossref metadata. The report from Shift Learning makes specific recommendations regarding pricing, products and services and, communications which Crossref should consider to continue to address the needs of its diverse stakeholders.

Keywords
Crossref; Publishing; Digital object identifier; Metadata; Community

Introduction

Background/rationale: As membership organizations grow and mature, it is important that they make sure that they are aligned with the diverse needs of the communities they work with. Crossref, founded in 2000 is no exception to this. In a Scholarly Kitchen blog post from October 2019, Brand [1] documents how Crossref has grown and the fact that the number of members has grown 900-fold since its founding, and is now working with over 11,000 organization. Similarly, Fairhurst [2] details the increasing geographic spread of Crossref’s newer members, explaining that “In more recent years, there has been a rapid growth in membership from outside of Crossref’s traditional West European and North American markets, with a large percentage of new members coming from Asia, Eastern Europe, and South America.”
The type of member that Crossref works with has also expanded to include those exploring different publishing models and preprint servers and funders are also joining to register preprints and research grants.

This diversity leads to both questions and challenges. How does Crossref serve such a broad group of stakeholders, how does it help them meet their goals, what value are they deriving from Crossref and where do they see value lacking? In identifying these groups, getting a sense of their priorities and their work with Crossref, and their hopes and frustrations, there are opportunities to be gleaned that can help Crossref continue to adapt and serve the organizations it works with.

**Objectives:** A value research project, commissioned by Crossref and lead by Shift Learning (https://www.shift-learning.co.uk/) in 2019 was an important exercise in researching how the organizations that work with Crossref feel about it, to give Crossref direction on how to develop its services and work with its community going forward. Specifically, the following were research questions to be solved: What is Crossref’s value to the different functions and roles in scholarly communications? What is Crossref’s value to society and the general public? Is Crossref’s mission understood and agreed upon? What are the barriers to achieving Crossref’s mission? How does the community think Crossref has changed in recent years, and what do they think about its strategic agenda? Those questions were answered via qualitative research by telephone interview and a brief survey.

**Methods**

**Ethics statement:** All interviewees followed a script prepared by Shift Learning and agreed with Crossref. No personal or sensitive information was covered in the topics of telephone interview. The survey was answered anonymously via a web form. There were no sensitive questions in the survey items.

**Study design:** It consisted of both qualitative research through an interview of 41 stakeholders and an anonymous web-based survey. Qualitative research was based on analyzing the content of the interviews and the survey.

**Participants:** A breadth of stakeholders was surveyed and interviewed as part of this work. Publisher members formed the bedrock of the Crossref ecosystem when the organization was founded, and still form the majority of the membership. The value research project reflects this.

**Measurement tools and data collection:** Interviewees and those surveyed were asked about their perceptions of Crossref. Survey questions provided as a supplementary table (Suppl. 1). For qualitative research, 41 telephone interviews were conducted with Crossref community members around the world. Shift Learning recruited, conducted, and analyzed the interviews following an email from Crossref. Each interview was 45 minutes to 1 hour long. Interviews were conducted using a guide rather than a specific set of questions. It aimed to cover the interviewees’ relationship with and understanding and perception of Crossref, the value of Crossref to them and their organization, what they thought about Crossref’s current mission, and what direction they thought it might take in the future. A brief survey, built by Crossref aided by Shift Learning was distributed to 476 contacts via email supplemented with other data-collection channels.

**Statistical methods:** All interviews were recorded and transcribed, then analyzed through qualitative data analysis software, Atlas.ti. Due to respondent confidentiality agreements, transcripts are not available to Crossref. Survey data was cleaned and processed in Q with data made available to Crossref in Q Reader.

**Results**

**Participants:** The 41 interviewees came from 10 countries, weighted towards North America and Europe, but also included responses from South America, Australia, and Asia. The interviewees consisted of nine publishers, five societies, five non-members, three university presses, three sponsoring organizations, three libraries, three researcher tools, three non-members who pay to use Crossref services (mainly metadata users), two companies, two publisher services (tools), one research funder, one government agency, and one assessment service. There were 437 survey respondents.

**Outcome:** A summary of the results has been shared by Shift Learning [3]. These provide a selection of quotes and details of the responses from interviewees and those who completed the survey. These have been summarized in the sections below.

**Positive feedback:** Many of the organizations surveyed shared positive perceptions of Crossref. There was a strong feeling that Crossref is central to the publishing ecosystem and all publishers to be involved with it benefit from the value conferred by working together on a shared service. Researchers and research systems can consistently use the digital object identifier (DOI) as it is used by a critical mass of scholarly content and Crossref provides a central place and a set structure for the information which promotes discoverability and reuse and the persistent linking ecosystem between publishers making important links between research outputs persistent.

The feeling that Crossref is collaborative in this way was appreciated by those who participated in the research, and they felt that Crossref’s mission was being met in this way.

Among the valuable actions attributed to Crossref related to scholarly communication were connecting research, stan-
dardizing infrastructure, managing metadata, maintaining a database, improving scholarly communications, enhancing findability, preserving content, automating discovery, and facilitating access. There were also findings that publishers, direct members and respondents based in Asia and smaller publishers were more likely to agree that working with Crossref gave them credibility and worked in a transparent way. This perception wasn’t shared as strongly by all respondents, which points to regional differences in how Crossref is perceived, and also speaks to the ongoing debate about predatory publishing. Crossref doesn’t vet publishers on the quality of their content, and having a DOI in itself does not infer quality as Bilder explains [4]. However, membership of Crossref does infer responsibilities for its members in terms of the maintenance of the DOIs they register and the related metadata, and asks them to make arrangements for its long-term preservation by working with archiving providers like Portico or a national library.

Survey respondents also placed value on the additional services that Crossref offers. Similarity Check, a service provided in partnership with Turnitin which provides a user-friendly tool to help scholarly publishers detect plagiarism was specifically mentioned here, as was Cited-by, which allows a member to find out which other members are citing their content. The Funder Registry, an open registry of persistent identifiers for grant-giving organizations around the world, allows everyone to have transparency into research funding and its outcomes help organizations avoid manually tracking and reporting on this information. Organizations also rated the good quality service and support they received when they needed help working with Crossref.

Organizations (apart from members) who use Crossref metadata were appreciative of the standardized metadata that Crossref members provide so that they had the capacity to use it in different tools and services, rather than having to get the information individually from publishers which do not scale well. The Metadata Plus service which provides enhanced access to all our supported Application Programming Interfaces (APIs) guarantees service levels and support and additional features such as snapshots and priority service/rate limits. Some enterprises felt that the Crossref metadata was integral to their businesses and provided a reputable alternative to the Google search and Google Scholar.

**Criticism and suggested improvements:** Overall, the survey provided positive feedback for Crossref. However, there were areas for criticism and concern that Crossref can aim to address. Some of these concerns came from longer-term members of the organization, who were worried that Crossref had become distracted from serving their needs, and was focused on new publishing models and the long-tail of members. Crossref’s mission is broad in scope, but with such a breadth of stakeholders, this can lead to confusion in who Crossref is serving. It can also lead to a desire for more clarity in terms of how Crossref makes strategic decisions and for more communication on product developments so that they can plan their own work in line with these.

Technical developments were noted: the creation of the Metadata Manager tool for registering content without needing to work with XML directly, the new schema for registering peer reviews, and the world-class REST (Representational State Transfer) API to access the metadata. However, further scaling and simplification of services and tools to register content with Crossref were high on respondents’ wish lists. Crossref has spent time over the last eighteen months working to eliminate some of its technical debt which will leave it well-placed to address these needs.

Metadata users desired cleaner metadata with fewer errors and omissions, and it was mentioned that Crossref lacks profile in some territories like China and could improve its profile in Latin America.

The survey also took the opportunity to ask respondents for suggestions on how Crossref could improve. These spanned six main areas: communication, operation, usability, outreach and promotion, onboarding, and training, and content registration.

In terms of communication, organizations would like more information on changes and developments to services, and to highlight the benefits of services rather than just the features of them, to help non-technical users. Some metadata users explained that they would like to have more of a voice with Crossref and potentially sit on the Crossref board.

On the operations and usability side, people wanted simplicity and the modernization of Crossref’s core technology. This would look at things like improving search algorithms for finding content in Crossref databases, simplifying the language used and improving the quality of instructions in other languages, and building a better process for giving feedback and disputing metadata.

Respondents felt that Crossref could make sure it update regional sponsoring organizations lists, encourage the use of DOIs in book publishing and work to promote themselves to researchers and journalists. More outreach and training would also be useful - this would include more educational and training resources for new and prospective members and service users.

Finally, while survey respondents indicated they are quite highly satisfied with registering content, they would welcome improvements to make the process more automatic and efficient, which would also help with metadata improvements, along with making it easier to correct mistakes made while registering metadata which can be difficult to do.
What if Crossref did not exist?: An interesting piece of the value research talks about what risks there are for the research ecosystem if Crossref did not exist. Some of the risks identified include 1) research outputs would be worse, because of the additional costs and time required to access the same materials; 2) the landscape would become balkanized and complex to maneuver within; 3) for those that used multiple services, this would mean negotiating a swathe of new contracts from different suppliers; 4) the end to progressive developments with the likes of preprints; and 5) an existential threat to scholarship in general, with much less likely to support this agenda if there was the scope for such catastrophe.

It is clear that some of the value that members get from Crossref is that it helps them support scholarship and the aims of their organizations, and the manual work involved in recreating the collaborative nature of Crossref services would create large overheads for individual organizations.

The vast majority of survey respondents said that they planned to continue working with Crossref in the future, due to the value of the services it provided for their organization. This is a positive message for Crossref to take, but it should also be mindful that for some services there may not be any alternatives to Crossref, so it's still important that Crossref continues to improve and add to its services over time.

Discussion

Key results and interpretation: Crossref is strongly appreciated by many stakeholders and has strong positive associations, delivering strong value for many through increasing discoverability, permanence, and linking. Secondly however, for some large stakeholders, Crossref has strayed significantly from its mission. In some of their opinions, Crossref has moved from an organization which serves members to one which serves the scholarly community more directly, using their funding to push a semi-political agenda and facilitating activities which may be detrimental to these members’ interests. Note that consistent with Crossref’s role as a nonprofit organization with 501(c)6 tax exempt status (providing benefit to the broader, non-member industry and community is consistent with, and a requirement of, 501(c)6 status—benefits to members should be incidental to the broader industry/community benefit. Third, smaller members feel strongly that Crossref is extremely good value for money, delivering the visibility that enables them to compete with larger content owners. Those organizations benefiting from metadata services also see the value, developing innovative services that are reliant on Crossref’s APIs. Some large, long-standing members again are less satisfied. They feel that they have not reaped the benefits of developments which should have reduced Crossref’s cost to them and that Crossref delivers decreasing value to them given developments in the sector and their already significant scale and visibility. Forth, while this is only a small number of publishers currently, Crossref needs to consider the decline in value of their metadata services and perhaps also in the value of membership to smaller content owners looking for credibility and contact with larger publishers that might occur were these larger publishers to follow through on their threats to leave the organization and go out on their own. Finally, proving to members that investment is going into ensuring that core systems work smoothly will be crucial to ensuring that relationships with all members are smooth moving forwards.

Limitations: All interviews were conducted in English. Respondents were self-selecting and interview incentives were not substantial and were donated to charity on behalf of the respondent. This may have resulted in a bias towards those with strong feelings towards Crossref, either positive or negative, and is likely to have resulted in a higher proportion of participation from those who already have an understanding of its activities and mission.

Recommendations on Crossref services: Crossref should be cautious in increasing prices for larger publishers (note that it was not suggested that Crossref would do this, and in 2020 Crossref has removed fees for the Crossmark service, which many publishers participate in. It has not increased membership fees for over 10 years). Crossref should be extremely cautious in increasing prices for the biggest publishers; while thinking about increasing prices for metadata users who are becoming established as businesses, and show the community how this is feeding back into further service improvements. New product development should be more transparent about its rationale, costs, and benefits to members as well as the fit to Crossref’s mission. Search algorithms should be updated to finding content in Crossref databases more accurately. Any subsidies for less developed countries should be publicized more strongly.

Conclusion: The value research survey was a useful opportunity for Crossref to make sure it is aligned with the needs of its community. It helped clearly identify key areas where members and non-members get value from Crossref, where they feel Crossref is diverging from those needs and take steps to rectify areas where it could improve. It is important that Crossref listens to this feedback as it works to continue to play a key role in supporting the research ecosystem. As Brand [1] says, ‘the Crossref of 2040 could be an even more robust, inclusive, and innovative consortium to create and sustain core infrastructures for sharing, preserving, and evaluating research information,’ and it should work towards that goal.
Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Data Availability

Due to respondent confidentiality agreements, transcripts from the interviews are not available to Crossref. Responses to the survey are confidential and anonymous.

Supplementary Material

Supplementary file is available from the Harvard Dataverse at: https://doi.org/10.7910/DVN/ESAJVG

Suppl. 1. Survey questions: the value of Crossref

References


3. Powell J; Shift Learning. Crossref value and benefits [Internet]. London: Shift Learning; 2019 [cited 2020 Jul 11]. Available from: https://docs.google.com/presentation/d/1RsqtnHssBkaFNphdWoq20_ewruYP04n8j_dYB-9wvphM/edit#slide=id.g65af51c04a_1_1002