Support Scholars Who Share: Combating the Mismatch between Openness Policies and Professional Rewards

Ali Krzton*

Introduction
Progress in scholarship is dependent upon access to knowledge. The internet has connected information seekers to scholarly works faster and more easily than at any other time in human history. That said, librarians and other information professionals have consistently called attention to the aspects of the new status quo in scholarly communication that fail to live up to the promise of offering broader opportunities for meaningful participation in research. In 2002, stakeholders frustrated at the persistence of cost barriers within the ecosystem drafted the Budapest Open Access Initiative, which helped the open access movement to coalesce around a common set of principles. Since its release, much progress has been made towards its goals; there are more products of research available online to anyone with a browser than ever, including scholarly articles, datasets, software code, and related web apps. For librarians and other advocates of open research, pushing for change is about more than controlling serials costs. It is fundamentally about a commitment to values: equitable access, reproducibility, and the integrity of research. This is why the slow pace of change and the persistence of restrictions and limitations have caused such frustration.

One strategy to speed up the adoption of open research practices has been mandates. Influential institutions, especially funding agencies, have started to require the products of research to be made openly and freely available, often as a condition of continued funding. The focus of these mandates is typically minimally-processed research data, articles published in peer-reviewed journals, or both. The imposition of these requirements from above has been necessary to spur change because researchers themselves, up to this point, have not chosen to prioritize openness when releasing articles and data. This is because researchers primarily perform according to expectations set within their academic discipline. Each discipline is in essence a community of practice that generates its own culture and enforces its own norms. Crucially, from the individual's perspective, their discipline's evaluation of the quality of their work is what determines their professional success. While open scholarship has been integrated into the culture of some disciplines as a matter of course, it is not considered important in others, and in a few the push for openness has engendered hostility, especially when it has been driven by top-down mandates.

This raises the question of whether institutional policies designed to advance open scholarship are capable of accomplishing that end, as well as what other consequences they might have for the practice of research. These mandates undoubtedly increase the number of publicly-available articles and datasets in the short-term, but there is also an expectation that as researchers become accustomed to the new requirements, the culture within their disciplines will shift to support open research, for instance through the founding of new open-access journals or platforms. Is this reasonable? The answer depends on the fate of individual scholars who support openness within their fields. While disciplinary norms are not biological traits, they can be said to evolve as new

* Ali Krzton, Research Data Management Librarian, Auburn University Libraries, alk0043@auburn.edu
attitudes take root and proliferate in the population of members. Cultural change is only sustainable when those who hold the new attitudes can remain in the population.

In this paper, I argue that the choices that determine professional gain and loss facing each researcher who wants to advance in their areas of study can be modeled as a prisoner’s dilemma. I first discuss why patterns in researcher behavior and beliefs suggest a prisoner’s dilemma in the area of data sharing, then apply this line of reasoning to the problem of perverse incentives in open access publishing. Finally, I explain what makes certain open access mandates counterproductive and propose alternative solutions that reward researchers who choose to make their work open. This provides a path towards reforming disciplinary norms concerning publishing.

Open Culture and the Prisoner’s Dilemma
A hypothetical cohort of scholars enter graduate school together. As the years pass, they earn their degrees, perhaps work as postdocs, become new professors, get tenure, and finally become senior leaders in their field. But at each stage of the process, the cohort is smaller. The academy is a competitive enterprise, and only those whose research achievements are judged favorably by their peers advance. What does this have to do with open scholarship? If openness is to become acceptable, and then valued, in a discipline, it is not sufficient to win over its current leaders. Achieving long-term, stable cultural change requires that junior researchers who are more committed to openness have a better than average chance of being promoted and going on to influence the field. In the current scholarly ecosystem, do they? The behaviors that contribute to the development of open research practices can overlap to a greater or a lesser degree with the behaviors that contribute to advancing along the career path. Thus, it can be useful to frame the adoption of openness as a coordination problem.

<table>
<thead>
<tr>
<th></th>
<th>Cooperate</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>3, 3</td>
<td>−1, 5</td>
</tr>
<tr>
<td>Defect</td>
<td>5, −1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

* The first number in each square is the outcome for the player on the left. The second number is the outcome for the player on the top. The sum of the two numbers produces the global or systemic outcome. Adapted from Principles of Cooperation Across Systems by Athena Aktipis.37
The prisoner’s dilemma is a game theory concept used fruitfully in research concerning evolutionary strategies. It is designed to explore whether cooperation can be maintained within a population when it results in community benefit, but at some cost to the individual who cooperates. In this example, there are only two participants, and each one chooses between two strategies: cooperate with the partner, or refuse (known as defection). What each partner does determines an outcome for each of them, as well as for both in total (See Figure 1).

In a prisoner’s dilemma, the payoffs for cooperation and defection are structured so that the best community outcome occurs when both cooperate, the next best occurs when one cooperates, and the worst is when both defect. However, from the perspective of the individual, the payoff is highest when they defect on a partner that cooperates, then slightly lower when both cooperate, then lower still if both defect. Critically, the worst outcome for the individual is when they cooperate with a partner that defects.

This means that if reward and punishment for cooperating vs. defecting are structured as a prisoner’s dilemma, cooperation cannot persist in the system even if it generates the best overall outcome. This is because partners who always cooperate are punished by repeated “betrayals” that disproportionately reward the partner that defects. Eventually, only defectors get ahead and remain in the system, but the overall outcome is worse, since the community reward was higher when partners cooperated with each other. In the scholarly ecosystem, practicing openness can be viewed as the cooperative strategy. More openness results in better scholarship for everyone because of faster and more widespread access to information. But if incentives are structured such that those who restrict the sharing of their work to more traditional avenues reap the biggest rewards, open practice will either become increasingly rare, or it will begin to take a form that benefit individuals at the expense of the community, defeating its purpose.

Early career researchers are particularly sensitive to the incentive structure of their discipline. In a prisoner’s dilemma situation, most will adapt their strategy towards defection to avoid falling behind. The junior researchers who cooperate regardless of incentives, the most likely candidates to inspire lasting cultural change if promoted, will be selected out of the population.

**Coordination Problems in Data Sharing**

A wide range of disciplines (for example: psychology, genetics, neuroscience, and conservation biology) have endorsed data sharing as a community goal. The benefits of data sharing include accelerated discovery, reproducibility of results, and new collaborations between researchers. Unfortunately, data sharing entails costs in time, effort, money, and sometimes lost publication opportunities which are born by individual researchers. If the only benefit they receive from data sharing is participating in a stronger research community, this is a recipe for a prisoner’s dilemma, because those who do not share still have access to shared data. To see whether incentives in data sharing were structured as a prisoner’s dilemma, Pronk and colleagues simulated a research community where the payoff was measured in productivity, i.e. number of publications. No matter how they parameterized the model, “cheaters” who took advantage of data shared by others but did not share data themselves were the most productive in all cases. Even when the total amount of data sharing in the community was high (reflecting widespread cooperation) it only resulted in increased total productivity; it did not change the incentive structure, and cheaters were still more productive than sharers. In short, as long as the direct benefits of data sharing are not enough to offset the costs to each researcher who shares, the optimal strategy is to avoid those costs personally while taking advantage of available shared data.

Although the study described above was a simulation, the results are consistent with the observed attitudes and behaviors of real scholars. Early career researchers, in particular, express the most support for data sharing as a general concept but were found to be less likely to share their own data. This is consistent with the predic-
tion that they would be more sensitive to the possibility of others advancing at their expense (they cooperate, others defect) than researchers who were more secure in the system. It also suggests a reason why data sharing mandates on the part of funders, journals, and other institutions sometimes produce resistance in a discipline as opposed to shifting norms towards openness.\textsuperscript{16} The stated goals of these mandates include making data more readily available to stakeholders and the public generally,\textsuperscript{17} but from the perspective of the coordination problem, data sharing would continue to gain momentum in the system if the mandates increased the costs of not sharing data to match or exceed the costs of sharing data, erasing the competitive advantage of those who defect.

However, simply increasing the cost of noncompliance cannot put researchers on equal footing if these costs are not experienced equally to begin with. Certain barriers to data sharing can have different impacts on researchers at different ranks; one study of research data management found that tenured faculty were significantly more likely to agree that the time and effort involved stopped them from sharing than untenured faculty.\textsuperscript{18} This implies that junior researchers have different reasons for choosing not to share data. Another problem with focusing on increasing the costs of noncompliance has to do with the nature of data sharing. Researchers can comply with the letter, but not the spirit, of a mandate by submitting data in substandard condition. A survey of publicly available datasets in the field of ecology and evolution showed that over half of these datasets were incomplete, and 64\% were partially or completely unusable due to metadata or file format issues.\textsuperscript{19} Data sharing mandates that increase costs uniformly across the system could further disadvantage researchers who put in the extra effort to cooperate while contaminating the pool of open data with low-quality datasets of limited utility. This not only punishes those who share, it also serves as a drag on the system.

Solving these coordination problems and breaking out of the prisoner’s dilemma involves both reducing costs and increasing the direct benefits to researchers who share. Librarians have done much good in the area of reducing costs by introducing researchers to documentation tools, developing metadata standards, identifying trustworthy repositories for referral, and providing direct support during data deposit and archiving, whether requirements are in place or not. But increasing direct benefits to the researchers who share is the only way to align the individual incentive structure with the optimal community outcome. In fact, researchers acknowledge that they would share more if the impact of their open datasets could be captured in a metric that mattered to their evaluation.\textsuperscript{20}

Data citation offers the potential to reward researchers for sharing, but it is not yet a widely accepted measure of scholarly impact. Librarians cannot make academic departments value highly-cited datasets as they do publications, but we can improve the chances that they will in the future. Right now, the bibliometric infrastructure to track data citations is inadequate,\textsuperscript{21} and as a result, informal citation practices predominate.\textsuperscript{22} Metrics-conscious junior researchers will not voluntarily take on the costs of sharing until they can give review committees hard evidence of the impact of their data. Advocates for open data need to promote standards in this area with the same level of urgency that has been devoted to building other research data infrastructure, and in fact, existing repositories and journals are well-positioned to support mainstream data citation.\textsuperscript{23} Librarians should target education about data citation towards junior researchers to help them start building their competitive advantage early.

**Reward and Punishment in Scholarly Publishing**

A mismatch between institutional requirements and professional rewards in data sharing has produced suboptimal community outcomes and disproportionately pressured junior researchers to choose between advancing their fields and advancing their own careers. Trends in scholarly publishing suggest that similar conflicting incentives are affecting the adoption of open access. More than any other factor, the number of high-quality publi-
cations a scholar authors determines whether they will be promoted. While the quality of a publication is ideally assessed based on its content, rankings-conscious committees and administrators have increasingly turned to journal-based metrics such as impact factor to determine intellectual merit. Article-based metrics, especially number of citations, are still important and arguably more accurate, but they are trailing indicators of scholarly impact; accumulating a large citation count can take years.

A researcher’s decision concerning how to disseminate their work has a tremendous impact on their career. While authors weigh multiple considerations when targeting journals for submission, it is tacitly understood that publications in journals with (at least relatively) high impact factors will count for more in performance reviews and, increasingly, in future grant applications. Junior researchers are especially attuned to these requirements, as they are often judged on an equal footing with more experienced colleagues. This explains why, in an international survey of early career researchers across multiple disciplines, respondents overwhelmingly targeted high-impact journals in their fields first. Confronted with the choice to publish faster or in the best journal for their work, early career researchers also report holding out for the better opportunity.

The fierce competition between researchers has not stopped open access publishing from being a viable, and even desirable, choice. Many authors publish open access regardless of whether they are required to do so. As a result, the number of articles that are, at minimum, free to read has grown, and the rate at which they are being published has increased. Piwowar and colleagues found that 45% of articles published in their 2015 sample were openly available in some form. The problem is that the way open access publishing developed and grew has so far failed to rein in the total costs of publishing, which was a core objective of the open access movement. Instead, the rapid growth in the ecosystem since the year 2000 has primarily been driven by gold and hybrid open access journals. As a consequence, academic publishing remains expensive overall, as the gold and hybrid models merely shift costs from readers and subscribers to authors and funders.

No one has yet modeled scholarly publishing using a game theoretical framework, but three trends in open access that have been harmful from the perspective of its advocates are suggestive of a prisoner’s dilemma dynamic. In contrast to shared datasets, open access publications can offer a direct professional reward to researchers in the form of a citation advantage. While the magnitude of the advantage may vary, it is reasonable to suppose that in general, papers available for free to anyone with an internet connection will attract more readers. Piwowar and colleagues found a citation advantage for hybrid, green, and so-called “bronze” (publisher-hosted and free to read but without an explicit reuse license) open access articles and a disadvantage for gold open access articles compared to paywalled articles. Although the non-standard “bronze” classification confounds the issue, as most of these articles would likely be considered gold in the rest of the literature, the anomalous result for gold open access is suggestive of one of its problems if taken at face value.

The original Budapest Open Access Initiative encouraged the formation of new gold open access journals, supported by institutions and other stakeholders in scholarly publishing, to supplant traditional subscription journals. While gold open access journals have proliferated and new ones continue to appear, only a handful have high impact factors (even relative to their disciplines). These high-impact journals are typically owned by large commercial publishers and are funded by per-article charges paid by authors or, for the well-resourced, their funders. Unsurprisingly, these fees are high (See Table 1).

Gold journals that charge authors little or nothing to publish articles but have lower impact factors are not comparable substitutions. Of the journals listed in Table 1, even those with comparatively low impact factors still command high prices because of the small advantage they offer over the thousands of totally unknown journals. If the obscure gold journals were removed from Piwowar and colleagues’ analysis, the citation disadvantage for gold open access would probably disappear. The only journal in the table that does not charge authors to pub-
lish is funded by two major universities, an exception that proves the general rule: only researchers who already have adequate institutional support and funding can afford to publish repeatedly in the most desirable open access journals. The high price of publication fees in these journals can be understood as the result of a series of defections in a prisoner’s dilemma. Authors repeatedly chose to publish in high-cost, high-impact journals over affordable but low-impact journals to avoid losing the impact factor arms race, authors that were not able to do so were selected out of the system, and prices remained high for the community at large.

The popularity of publishing open access in hybrid journals, which still charge for subscriptions, is also a logical outcome of repeated defections. Authors who pay to open individual articles in high-impact subscription journals maintain their impact factor advantage and gain an additional citation advantage at the expense of authors whose manuscripts are accepted by the same journal but who cannot afford to buy out the paywall. The community as a whole suffers because more funds are diverted away from research due to publishers’ double-dipping. The hybrid model reinforces and exacerbating existing inequalities in the system. Still, as bad as hybrid journals have been for open access, the trend that has undermined its goals the most is the erosion of self-archiving.

The Budapest Open Access Initiative framed the self-archiving (green open access) strategy as equally important to the proliferation of gold open-access journals. In contrast to publishing in high-impact but expensive journals, self-archiving allows authors to make their work open for free with only a small investment of time. Unfortunately, the institutionally-supported repositories that were designed with green open access in mind are underutilized. Instead, authors overwhelmingly choose to self-archive on academic social networks like ResearchGate, even though they often violate their publishers’ copyright in the process.33 In addition to potentially being illegal, posting articles to commercial social networks hurts their discoverability in comparison to technically compliant repositories. It is a defection in the sense that even though the costs of depositing an author-accepted manuscript into an institutional repository are low, these commercial sites are even easier to use and give authors a boost in promoting themselves in addition to their work. While the community-optimal solution would have been a strengthening and normalizing of the institutional repository ecosystem, researchers repeatedly following the path of least resistance allowed a parallel, less open, and apparently legally vulnerable system to grow in its stead.

### TABLE 1

<table>
<thead>
<tr>
<th>Journal Title</th>
<th>IF (2017)</th>
<th>APC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genome Biology</td>
<td>13.214</td>
<td>$3490</td>
</tr>
<tr>
<td>Nature Communications</td>
<td>12.353</td>
<td>$5200</td>
</tr>
<tr>
<td>PLoS Biology</td>
<td>9.163</td>
<td>$3000</td>
</tr>
<tr>
<td>Scientific Reports</td>
<td>4.122</td>
<td>$1790</td>
</tr>
<tr>
<td>Geoscience Frontiers</td>
<td>4.051</td>
<td>0$</td>
</tr>
<tr>
<td>Database: Journal of Biological Databases and Curation</td>
<td>3.978</td>
<td>$1680</td>
</tr>
<tr>
<td>Frontiers in Plant Science</td>
<td>3.677</td>
<td>$2950</td>
</tr>
<tr>
<td>Remote Sensing</td>
<td>3.406</td>
<td>$1800</td>
</tr>
<tr>
<td>Ecology and Evolution</td>
<td>2.340</td>
<td>$1950</td>
</tr>
<tr>
<td>PeerJ</td>
<td>2.118</td>
<td>$1095</td>
</tr>
<tr>
<td>Geoscience Data Journal</td>
<td>1.867</td>
<td>$1500</td>
</tr>
</tbody>
</table>

*Article publication fees funded by China University of Geosciences and Peking University*
Choosing a Sustainable Path to Openness

Rather than conceiving of scholarship in a binary state of open or not, it is more accurate to speak of openness in degrees. The ideal of a system that makes articles, data, and other products of research freely accessible, without delay, and with almost unlimited potential for reuse, is a worthy one, but the path to get there may involve compromising on some of these, at least in the near term. I have argued that protecting the interests of junior researchers is an important consideration for advocates of open scholarship because if participating in such harms their career prospects, only those who are indifferent or hostile to openness will be promoted, precluding cultural change within their disciplines. I have further argued that the costs of certain activities affect junior researchers more, disadvantaging them to the point that they find themselves in a prisoner’s dilemma where they must act in their own narrow interest to the detriment of the community as a whole.

Based on this analysis, I suggest two criteria librarians and other information professionals should consider before offering their support for any particular openness mandate, whether it concerns articles, data, code, or any other product of research:

1. Does it affect the ability of researchers to publish in the highest impact journal that will accept their work?
2. Does it push researchers into a monetary arms race to secure a citation advantage?

Mandates that fail one of the above tests are likely to be unpopular with researchers in general, if only on the grounds of academic freedom, but they will be especially harmful to junior researchers. As a general rule, mandates that privilege gold over green open access, or specify that research must be published on a specific platform, deserve extra scrutiny. This is not to say that all mandates are bad. For example, a mandate that refused to recognize articles published in hybrid journals as compliant might have a protective effect on researchers with fewer resources, regardless of their rank. But a misplaced emphasis on mandates to the exclusion of other strategies will allow the coordination problems in the scholarly ecosystem to persist indefinitely.

Early career researchers will adopt a positive attitude toward openness if it helps them succeed. To this end, librarians and other research support professionals should educate them about the citation advantage of open access articles and, more importantly, help them to take advantage of it. When it comes to articles published in journals with high impact factors, green open access is by far the most cost-effective strategy and offers at least as much of a citation advantage as gold. Self-archiving in institutional repositories should become a nearly frictionless process, with librarians providing guidance concerning copyright and even assistance with deposit when possible. Journals may adopt more liberal self-archiving policies if authors, eager to reap the benefits of increased discoverability and citation without paying thousands of dollars for the privilege, begin to demand them. The work of promoting new gold open access journals that charge little or nothing to publish should be left to leaders in the field who are no longer as reliant upon impact factors to signal the quality of their research. That said, helping junior researchers to demonstrate the impact of their work to reviewers through article-level metrics, as opposed to journal-level metrics, will make research assessment more accurate and potentially lower the premium that high-impact journals command through both subscriptions and article processing charges. In theory, all other things being equal, an open access article should have better article-level metrics than one behind a paywall.

Rewarding the scholars who share, rather than putting them at risk of falling behind for doing the right thing, will allow them to stay in the pipeline, continue the research they’ve worked to make accessible, and change the culture of their disciplines organically.
Endnotes


4. Chan et al., "Budapest Open Access Initiative."


13. Pronk et al.

14. Ibid.

15. Tenopir et al., "Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide."


17. Briney, Goben, and Zilinski, "Institutional, Funder, and Journal Data Policies."


20. Ali-Khan, Harris, and Gold, "Point of View."


28. Piwowar et al.


30. Piwowar et al., "The State of OA."

31. Chan et al., "Budapest Open Access Initiative."

32. Piwowar et al., "The State of OA."


35. Ali-Khan, Harris, and Gold, “Point of View.”