Open IP Workgroup Report

Joann Delenick, Donald Guy, Laurel Haak, Patrick Herron, Joyce Ogburn, Crispin Taylor

Abstract / Workgroup Question

As a new issue for OSI2017, this workgroup (originally designated as the Patent Literature workgroup) will look at patent literature, research reports, databases and other published information. OSI by design has a university-centric and journal-centric bias to the perspectives being considered. Patent literature, research reports, and databases are also important sources of research information—more so than journals in some disciplines (although these still reference journal articles). As with journal articles, this information isn't always free or easy to find and is suffering from some of the same usability issues as journal articles.

I. Framing and Scoping

From the beginning of our workgroup discussions, we realized that the scope of our assigned topic, *patent literature*, was too narrow in comparison to the range of intellectual property specified in the topic assignment. With patent literature, research reports, and databases in mind, we looked at the topic as a broader continuum of intellectual property (IP). Our group began by defining *intellectual property* as the set of objects comprised of artifacts created or otherwise contributed by researchers that either potentially or presently are part of the scholarly record. This broad definition led us to rethink our topic and, in keeping with the conference theme of open scholarship, rename it as *Open IP*. Further workgroup discussion revealed that, beyond patents, many types of IP lacked information on existing models, structures, workflows, or standards, highlighting the need for more time-consuming exploration and concept development. Despite the expanded scope of interest, due to time constraints our workgroup concentrated our efforts on developing recommendations relevant to improving the discovery, access and use of patent data and closely-related IP.

II. What is Open IP?

In the context of our discussions, the concept of Open IP could certainly relate to ideas of open innovation; we recognize, however, that intellectual property, especially when operating within scholarly domains, can far exceed its role as a foundation for commercial innovation. The patenting process is itself a tradeoff between publishing and protection. To be granted a patent requires the invention to be published. Patent files are maintained as a public resource, by national and transnational organizations. The group agreed that one of the major challenges is that while the patent files are openly accessible, they are not easily used due to structural issues in the way the data is collected and the file is constructed and presented. This makes it difficult to understand who are the inventors and assignees, which in turn has an impact on our ability

to track innovation and develop policies to better support innovation.

Given that patent files are openly accessible, another discussion thread was how Open IP might be practiced. This relates to but is different from licensing of IP. It is possible for organizations, as Tesla and NASA have demonstrated, to implement open patenting practices. As a federal government agency of the United States, NASA has a public domain collection of IP available to users, and has created a website to facilitate the transfer and translation of their open technology portfolio.[1] In private Industry, Tesla has opened the licensing of their patents, citing the greater advantage of finding engineers already familiar with Tesla's technology over the restricted use of their patents. Tesla claims they innovate much too quickly for infringement of their patents or patentable technologies to harm their firm in a substantive way [2] Both organizations argue that opening their patentable research accelerates knowledge production that simultaneously serves the interests of the individual organizations, educational and training institutions, workers such as scientists and engineers, and the broader public interest. While opening patent licensing in the ways such examples illustrate may have a number of advantages concurrent with the goals of open scholarship, the scope of our interest is currently limited to the opening of IP-related data.

III. Discussions

Several ideas and frameworks emerged soon after we started that helped us shape and hone our discussions and potential recommendations. These centered on the accessibility of documents rather than on licensing, and included content types, principles, standards, stakeholders and incentives. We talked briefly about recent literature regarding the content of scholarship and the scholarly record, referencing the OCLC report on The Evolving Scholarly Record[3] and NISO's work on standards and best practices.[4]. We also determined that we would consider all digital formats of scholarship.

Our workgroup first identified a set of principles and values appropriate for making and evaluating recommendations, leading to an effort to create an inventory of contemporary or potential types of intellectual property. These two initial discussions led to a greater discussion of scholarly communications, stakeholders, incentives, standards and practices, setting the path for making recommendations.

A. Guiding Principles and Values for Open IP

The group quickly agreed that the elucidation of fundamental Open IP principles would be an important first step toward exploring this topic. The group thought that a combination of open and public domain principles should apply both to content as well as the computational analytics developed to understand such content. One reason for this is because content is increasingly valuable for use in discovery and content creation applications in medicine, the pharmaceutical industry, engineering, and even in the humanities and arts. Moreover, we observed that analytics are becoming essential for critical computational analysis tasks such as disambiguation, text structuring, and basic bibliometrics, scientometrics, infometrics, and altmetrics. In response to these ideas, we developed the principles noted below in Table 1.

Guiding Principles and Values for Open IP
Accessible
Discoverable as academic content
Interoperable
Machine-readable optimization for non-consumptive use
Text and data mining
Persistent, i.e., long-term availability in the scholarly record
Pragmatic
Global in scope and perspective
Generalizable
Multidisciplinary
Stakeholder commitment to Open IP
Anticipatory, forward looking toward new developments and concepts

Table 1. Guiding Principles

B. Open IP Content Types, Intellectual Outputs, and Creative Activities

We enumerated types of intellectual property and added them to either of two groups: *proximal*, i.e., those objects not only more likely to be thought of as IP but also more likely to be within the scope of this workgroup; and *distal*, i.e., objects less likely to be equated with patent-centric notions of IP, and hence more challenging to make recommendations for openness. Table 2 below contains a list of intellectual property types we identified.

Proximal Types of IP	Distal Types of IP
Patents (including design)	Archives
Patent-related specimen and materials repositories	Artifact and biological collections
Databases and datasets	Art installations
Software	Blogs
Clinical trials	Court testimony/opinions
Research reports	Expert panel
Regulatory government filings	Peer review
Grant abstracts/proposals	Analytics
Technical specifications	Music
	Simulations
	Tweets and other social media posts
	Games
	Interactives
	Algorithms
	Trade secrets
	Copyright

Proximal Types of IP	Distal Types of IP
	Trademarks
	Virtual and augmented reality

Table 2. Types of IP

C. Differences between Patent Literature and Scholarly Research Literature

Scholarly research literature is generally well-structured digitally, and such structures are becoming widely standardized, which greatly aids discoverability, utility, and interoperability. Both content and standards are often distributed by digital means, thereby making them at least potentially more readily discoverable than other types of scholarly outputs. Although technological improvements could be made to digital formats, the content therein would also become increasingly discoverable with further standardization of digital publication document structures.

Patent literature, however, can be structurally complex. The diversity of patent granting offices results in diverse patent literature structures. Although the adoption of WIPO XML standards (e.g., ST.36) by patenting agencies reduces structural complexity, the fact that metadata fields are largely manually-entered text strings means that patent-based discovery remains difficult even for the offices of patent prosecutors and patent examiners.

Variations in naming conventions and classification standards within and between patenting agencies only add to these already significant challenges. Combining patent and research literature poses even greater complexities. For example, although both scholarly publications and patent records are easily citable, in filing patent applications and writing research papers alike it is often difficult to integrate and incorporate resources from both types of literature.

The increasing adoption of metadata standards and discovery technologies, such as open crosswalks, open APIs, persistent identifiers, and controlled vocabularies, helps stakeholders connect to siloed information and data types. Improvements to open crosswalks and APIs between literature and patents would lead to better discoverability and more complete utilization of journal and patent literature.

D. Stakeholders

Any conversation about IP requires the identification of and engagement with stakeholders, of which there are many. IP stakeholders are those groups or classes of people or organizations who have an expressed interest and and/or concern with regards to the accessibility of patent data files and other IP.

To begin with, there are individual stakeholders who create or manage research and the resulting IP— inventors, research scholars, scientists, artists, performers, librarians, and the like. Other stakeholders are organizations or institutions, such as publishers, database owners, governments, patent and technology offices (both governmental and academic), research parks, corporations, and business startups. The members of the general public also have a stake in the application of IP, and may have specific cultural ideas regarding research property. One example

is the different views and values that indigenous peoples may hold in regard to IP.[5] Professional groups also have a stake: small business advisors, patent attorneys, investors, and incubators. A trickle-down effect also occurs for groups and individuals who, while not directly involved, are impacted by IP policy and assignments. This stakeholder group might include downstream licensees and subcontractors, real estate developers, city and regional planners, architects, designers, environmental impact managers and potentially more. We need to remember that stakeholder groups can transcend states and nations and that IP can be both legislated and negotiated in treaties, contracts, and licenses.

E. Incentives

Along with other topics, we explored the incentives that might apply to Open IP. In some cases, they may be associated with a particular stakeholder group, and in other cases they may cross multiple groups; all stakeholder participation and behaviors, however, are influenced by incentives, rewards and motivations. Generally speaking, Open IP can greatly enhance new discoveries, further research, improve upon existing intellectual property, advance and stimulate education, and otherwise enrich the world and our daily lives. Specific deliverables from and incentives in support of Open IP may be conceived to include:

- Increasing numbers of citations or references for a disseminated work
- Increasing economic gain or revenue—for academic organizations, grant funders, researchers, publishers, software and service providers, university spin offs, and the like
- Achieving higher impact it in terms of citations or the public interest
- Increasing reputation and visibility

- Acquiring more funding for research and development, provided by both individual donors and corporate entities
- Supporting state and regional economies via innovation, diffusion or translation of research into practice.

In addition, we discussed core values and drivers that could make IP more open:
- Sharing so others can build on results
- Mitigating or reducing risk
- Making systems and content interoperable to improve access to data across institutional silos ("Circle of gifts"/Open data)
- Supporting monetization, cost recovery, and return on investment (ROI)
- Serving the long-tail public good, wherein a small population/underserved/narrow band of people benefit.[6]

F. Standards and Identifiers

We identified standards as one means for easing or enhancing use of the patent literature, particularly by assigning unique identifiers—much like DOIs and other identifiers are deployed. Possibilities for new standards include the "who, what, and where" of the patent or other IP resource. We agreed that such standards and identifiers should be persistent, global, and applied to academic, governmental, and private entities alike. There is a rich set of examples and cases from which to draw inspiration for developing effective standards and with which new standards could be integrated or related.

G. Acceptance as Scholarship

Patents and other forms of documentation of commercializable or otherwise

translatable technologies are being recognized as forms of scholarship in so far as they are increasingly being incorporated into promotion and tenure reviews, practices, and guidelines. Texas A&M, University of Arizona, Virginia Tech, and a few other universities have incorporated these products in their efforts to improve and modernize promotion and tenure guidelines. An excerpt from the Iowa State University's faculty promotion and tenure guidelines serves as an example:

> Scholarship results in a product that is shared with others and is subject to the criticism of individuals qualified to judge the product. This product may take the form of a book, journal article, critical review, annotated bibliography, lecture, review of existing research on a topic, or speech synthesizing the thinking on a topic. Also falling under the umbrella of scholarship are original materials designed for use with the computer; inventions on which patents are obtained; codes and standards; art exhibits by teacher-artists; musical concerts with original scores; novels, essays, short stories, poems; and scholarly articles published in non-research based periodicals, newspapers, and other publications; etc. In short, scholarship includes materials that are generally called "intellectual property.[7]

The continuum of scholarship our workroup has proposed in Section B, above, is reflected in the Iowa State guidelines. As other universities expand the scope of scholarly artifacts to include forms beyond research publications, they are likely to become ever more inclusive of the full set of scholarly intellectual property objects enumerated by our workgroup.

G. Public Access and Taxpayer-Funded Research

We noted that patents and inventions are protected as IP but are also intended to support the public interest. Their protection is more limited in time than that of copyrighted materials. Although somewhat different, it is worth commenting that directives and policies for public access to articles and data promote the dissemination of ideas and stimulate innovation. Sources of funding (such as the U.S. taxpayer) and considerations of the public interest have become factors in deciding whether IP should be open or restricted for a specified period.

IV. Recommendations

We perceived overlaps with the discussions and recommendations of other OSI workgroups and stakeholder groups, such as Standards, Norms, and Best Practices; Scholarly Research Infrastructures; and Promotion & Tenure. We should ensure that overlaps are addressed, and look for commonalities and differences. In direct response to our work, we recommend that OSI delegates do the following:

- Affirm and add to the guiding principles as they relate to all forms of Open IP.
- Establish internationally-recognized standards, including persistent IDs, for name, organization, and citation (the who, where, and what referred to above). Promote the widening adoption of both metadata standards and discovery technologies, such as open crosswalks, open APIs, and controlled vocabularies, that help stakeholders connect to siloed in-

formation and data types. Improvements to open crosswalks and APIs between literature and patents would lead to a more complete discoverability and utilization of journals and patents alike. Engage with WIPO in these efforts.

- Define the IP continuum as a first step in providing a larger context for Open IP issues and seeding future conversations. Then, shift focus from patents to other IP and scholarly products, including software, databases and reports that require consideration of standards, identifiers, and other structures that would facilitate their discovery and use (see Table 2).[8]

- Explore patent licensing as an alternative means of use, and discuss applications of "Open patenting" or "patent-left." To enforce the principle of transparency, ensure that license details are published. Examine whether open licensing creates an incentive to researchers to take otherwise "dead" outputs and revive them.[9]

- Emphasize the public good to encourage a different understanding of IP and how it can be deployed. Assess the impact of investments in patents that are tied to grants.

- Identify stakeholders and gain a better understanding of meaningful incentives for these groups.

- Create a sandbox—a tightly-controlled, separate set of resources—and source control infrastructure for testing and sharing computational analytics implementations (e.g., code examples for scientometric-based applications).

- Provide education and IP literacy for undergraduate students through senior scholars, e.g., by developing an IP boot camp for anyone outside the legal profession who needs to keep up to date with the always-changing policies, rules, conventions, and software.

- Promote clear language in patents. Some aspects of the language may be construed as the reserve of the legal establishment and IP managers. However, as part of our mission we are equally responsible for providing clear language on how to exercise standard formats that we will recommend. As an example, we might recommend, "there is a two letter and only two letter code for countries and these letters are used in upper case font," rather than referring to any particular country code table.

- Determine what "published" means in the context of "prior art" for patents, and how this applies to various places around the globe. Is something published if it has a DOI, an abstract and title, or other markers that have been or could be determined?

- Identify the risks of Open IP in a world where resources (computational and labor power) are unevenly distributed and the best methods are frequently not open.

V. Summary

In considering our discussion topic, patent literature, the workgroup took the opportunity to expand further upon an understanding of intellectual property. In defining a keystone concept, Open IP, we developed initial recommendations concerned with guiding principles and values with respect to Open IP content and types, intersections between patent literature and scholarly research literature, and

other considerations regarding stakeholders, incentives, and motivations. As a major point of consensus, we affirmed the role of standards development, especially the implementation of unique identifiers, which would provide greater access to patents as well as other intellectual property products moving forward. Our discussions have illustrated, for the first time, a collective, open public face to all dimensions of IP and we expect our current and future recommendations to address this continuum of Open IP as OSI further examines this topic.

Open IP Workgroup

Joann Delenick, Staff bioscientist, Intralogics/PatentVision
Donald Guy, Manager, Research Collaboration & Library Services, Sandia National Labs
Laurel Haak, Executive Director, ORCID
Patrick Herron, Senior Research Scientist, Information Science + Studies, Duke University
Joyce Ogburn, Digital Strategies and Partnership Librarian, Appalachian State University
Crispin Taylor, Chief Executive Officer, American Society of Plant Biologists

End Notes

[1] https://technology.nasa.gov/publicdomain

[2] https://www.tesla.com/blog/all-our-patent-are-belong-you

[3] Lavoie, Brian, Eric Childress, Ricky Erway, Ixchel Faniel, Constance Malpas, Jennifer Schaffner, and Titia van der Werf. 2014. The Evolving Scholarly Record. Dublin, Ohio: OCLC Research. http://www.oclc.org/research/publications/library/2014/oclcresearch-evolving-scholarly-record-2014-overview.html

[4] See for example *Outputs of the NISO Alternative Assessment Metrics Project. NISO RP-25-2016*. Baltimore, MD: National Information Standards Organization, 2016. http://www.niso.org/publications/rp/rp-25-2016

[5] See for example Anderson, Jane. (2010) *Indigenous/Traditional Knowledge & Intellectual Property*. Durham NC: Center for the Study of the Public Domain Duke University School of Law. http://web.law.duke.edu/cspd/pdf/ip_indigenous-traditionalknowledge.pdf, and the Nagoya Protocol. https://www.cbd.int/abs/about/.

[6] See, for example, Open Prosthetics. https://openprosthetics.org/

[7] Iowa State University. (2017) Faculty Handbook, 58-59. http://www.provost.iastate.edu/faculty-and-staff-resources/faculty-handbook

[8] Effective January 2013, the NSF Proposal & Awards Policies & Procedures Guide (PAPPG) specified revised content for biographical sketch information for senior personnel participating in NSF research grant proposals by deleting the publication references requirement and substituting a requirement for research product information. See http://seqanswers.com/forums/showthread.php?t=22559. This research product requirement for research personnel is currently in place at NSF as stated in the PAPPG bulletin of 2017. https://www.nsf.gov/pubs/policydocs/pappg17_1/nsf17_1.pdf

[9] There are semi-open venues where this information can be deposited and discovered, e.g., flintbox, http://www.flintbox.com/.