**Standardizing Intellectual Property Disclosure Data**

*Short paper prepared for the NBER preconference on Standards, Patents & Innovation, May 7, 2011*

Rudi Bekkers [a,b], Christian Catalini [c], Arianna Martinelli [d], Tim Simcoe [e]

[a] *Eindhoven Centre for Innovation Studies (ECIS), School of Innovation Sciences, Eindhoven University of Technology, the Netherlands. r.n.a.bekkers@tue.nl*
[b] *Dialogic Innovatie & Interactie, Utrecht, the Netherlands.*
[c] *Rotman School of Management, University of Toronto, Canada. christian@catalini.com*
[d] *LEM-Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, Italy. a.martinelli@sssup.it*
[e] *School of Management, Boston University, USA. tsimcoe@bu.edu*

***Abstract***

*Disclosure of essential patents at standard-setting organizations provides a rich source of information that can be used for various research questions related to standards and innovation. Yet this data also has some limitations and its compilation and preparation create challenges. This paper summarizes a number of recent studies using this type of data and discusses recent efforts to create an open database.*

## 1. Introduction

In the early 1990s, many standard-setting organizations (SSOs) introduced Intellectual Property Rights (IPR) policies. While these are rather heterogeneous, they all contain three elements: disclosure requirements, rules on the possible consequences of such disclosures, and licensing requirements (Lemley, 2002). Disclosure can be seen as a compliance process: firms reserve their rights and promote their own technology, whereas SSOs set the ground rules and limit their exposure to antitrust lawsuits.

While the IPR disclosure process is primarily directed at standards developers and implementers, many SSOs make disclosure data publicly available. As a result, scholars and policy makers may use disclosure data to gain better insight into several innovation-related questions. We review several recent studies have relied upon IPR disclosure data, and find that they cover a broad range of research topics. We also point out some limitations and problems with the publicly available SSO databases. Finally, we discuss we describe an ongoing effort to create an open database with IPR disclosure data from more than ten large SSOs.

This paper is structured as follows. Section 2 gives a short introduction to patents in standards, as well as the main IPR policies at SSOs. Section 3 provides a brief overview of recent studies that relied on SSO disclosure data in one way or another, whereas Section 4 discusses some inherent limitations.

The final part of this paper, Section 5, discusses the Open Essential IPR Disclosure Database, which the authors plan to make available to any interested party.

## 2. Patented technologies and standards

In the past two decades, interoperability standards have become increasingly important. This has been most visible in areas such as telecommunications (e.g. mobile telephony standards), Information Technology (IT), and Consumer Electronics (CD, DVD, MP3). Less visible, but possibly even more important, is the use of compatibility standards in other domains, where they often serve as enabling technologies. This development can be observed in sectors such as transport/logistics, services, agriculture, and biometrics. In the future, it is expected that the range of such sectors may expand considerably into areas like health (including personal health care), domotics, and energy distribution (e.g. smart metering and electrical vehicle charging).

Interoperability standards have a specific nature. In order to enable devices to work together successfully, you need very detailed specifications (for some standards well over 10,000 pages) and the integration of cumulative technologies. These standards often need to meet a challenging set of functional, technical and legal requirements, which might very well be conflicting.[1] Clearly, certain technological solutions may meet these requirements better than others. Given the high propensity to patent in many 'high tech' sectors, attractive technological solutions will often be patented. This is especially true in the field of electronics (which provides many of these enabling technologies), and to a somewhat lesser degree in the IT sector (where software patents are a relatively recent phenomenon and limited by some jurisdictions).

Most SSOs recognize that it may be appropriate to include patented technology in an industry standard, when alternative solutions are technically unattractive. At the same time, patents in standards raise several concerns, particularly from the antitrust perspective. For instance, if an owner of an essential patent would not be willing to license to some or any implementers of the standard, market entry would be hindered and the diffusion of the standard could be under threat.

In the late 1980s a major conflict arose over patents essential to the GSM standard for mobile telephony (Bekkers, 2001; Iversen, 1999; Garrard, 1998). In response, several formal SSOs adopted IPR policies to address concerns over patents in standards. Many smaller standards fora and consortia followed this trend. (At the same time, antitrust policies and jurisprudence also helped to address a number of these concerns.) The most common IPR policy to emerge from these debates was a requirement to grant licenses on "Fair Reasonable and Non-Discriminatory" terms. This is often called a F/RAND policy. Some standards organizations (e.g. the World Wide Web Consortium) require a Royalty Free licenses. Other organizations offer a list of options, which may also include voluntary disclosure of the most restrictive licensing terms.

Most F/RAND policies require active participants to disclose the patents that they believe to be essential for the implementation of standards. Most SSOs make a list of such declarations public

[1] For example, in the area of telecommunications, a standard for mobile internet data services might aim to (1) offer a high data rate (speed), (2) in a large, continuous coverage area, (3) allowing the user to move with speeds up to 300 km/h, (4) consuming low power in order to optimize battery life, (5) requiring a minimum number of cell sites or antenna towers, (6) while being robust to noise and other types of interference and (7) at low costs for base stations and terminals.

('IPR databases'), though there are exceptions. These IPR databases are the most tangible manifestation of IPRs in standards, though they also have a number of limitations (which we will discuss below).

**3. Studies that rely on disclosure data**

In the recent past, a number of studies relied on SDO disclosure data to address topics as: general trends in standardization, characteristics of SSOs patents and SSOs performance and determinants of essentiality, performance of SSOs, firms' IPR policies and participation to standard development, and the emergence of specific institutions such as patent pools. This section provides a brief overview of selected studies.

First of all, there are several studies that focus on IPR trends in standards. They include aspects such as disclosure trends, industry fields, and ownership distribution. In Simcoe (2005), a strong growth in the number of disclosures was reported, with a particular surge around 1994 (see Figure 1). In a recent report commissioned by the European Commission, various aspects of IPR in standards have been documented, including the distribution over SSOs, over specific standards, over technological areas, over time, and by ownership (Interplay consortium, 2010). In addition, features of the IPR owners were studied such as the company size, country of origin, the sectorial distribution, R&D intensity, and business models.

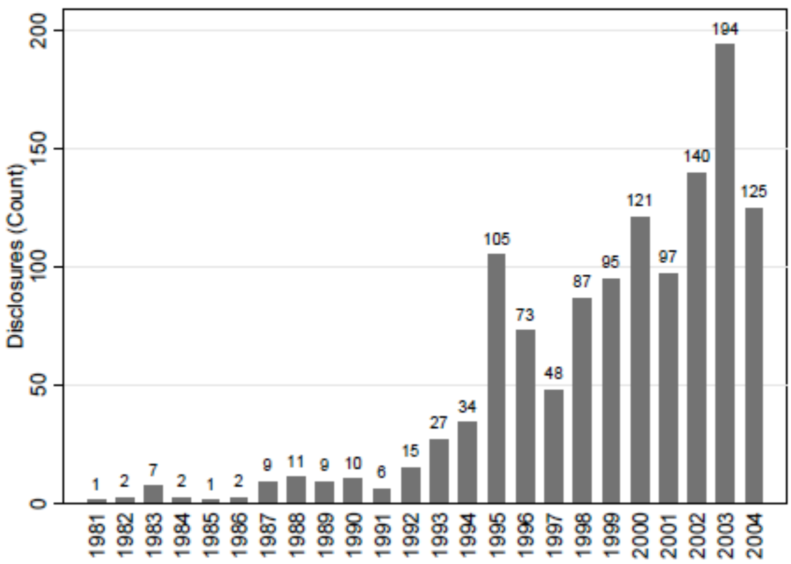


*Figure 1: Total IPR Disclosures ( Simcoe, 2005)*

Other studies assessed to what degree SSOs were able to attract valuable technologies, by analysing the citation-age profile of essential patents (Simcoe et al., 2009; Rysman & Simcoe, 2008). It was found that essential patents indeed receive more inward citations and do have a longer citation tail (see Figure 2, left), suggesting that standard-setting organizations attract/include technologies with a higher than average value. Furthermore, it was found also a significant "disclosure" effect, with an increase in citation rate after SSO disclosure (see Figure 2, right).
Essential patents show higher quality along several dimensions: they receive more citations, they contain more claims, they are more general, and they are more litigated (Table 1).

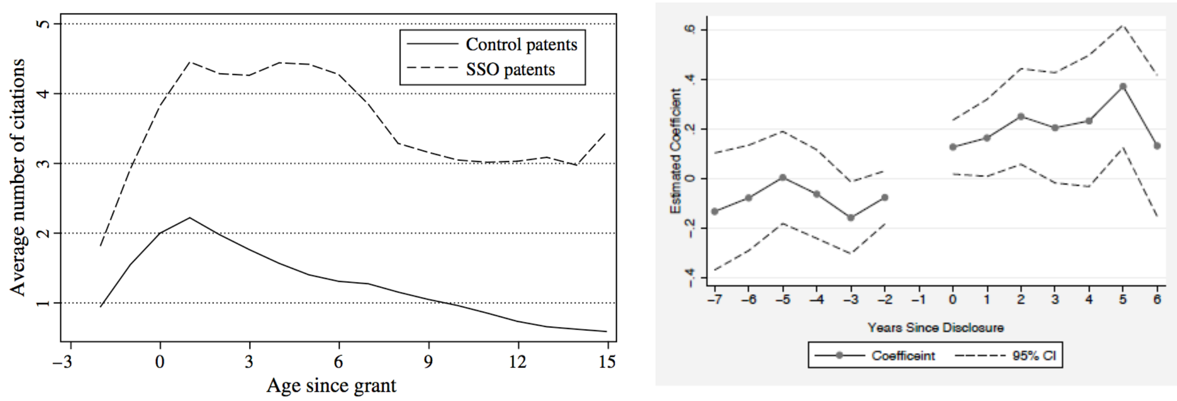


*Figure 2: Citation tail for SSO and Control Patents (Rysman & Simcoe, 2008) and Estimated Pre and Post-disclosure coefficient for SSO Patents (Rysman & Simcoe, 2008).*

| | *SSO Patents* | *Random Match* | *P-value* |
|---|---|---|---|
| Litigation Rate (percent) | 9.38 | 1.69 | 0.00 |
| Lit Rate (grant pre-94) | 14.24 | 1.99 | 0.00 |
| Lit Rate (grant 94-98) | 7.11 | 1.55 | 0.00 |
| Lit Rate (grant post-98) | 4.46 | 0.96 | 0.01 |
| Lawsuits (count)** | 1.97 | 2.06 | 0.87 |
| Defendants (count)** | 2.69 | 2.56 | 0.90 |
| Litigation Age (years)** | 6.20 | 3.75 | 0.02 |
| Pre-disclosure (percent)** | 28.09 | | |
| Forward Cites 63-06 | 33.88 | 16.84 | 0.00 |
| Backward Cites | 10.75 | 10.36 | 0.56 |
| Non-patent Cites | 9.07 | 4.69 | 0.00 |
| Claims | 22.24 | 17.89 | 0.00 |
| Continuation | 0.43 | 0.31 | 0.00 |
| Generality | 0.51 | 0.44 | 0.00 |

*Table 1: Patent litigation rates (Simcoe et al., 2009).*

Other studies focused on the impact of essential patents. Bekkers et al. (2002) discussed a case in the late 1980s and early 1990s, where in the absence of specific IPR policies at SSOs, a single right holder was able to dictate market entry and thereby market structure (this right holder was Motorola, see Figure 3). It was this event in particular that prompted SSOs to develop IPR policies, and many firms to develop explicit strategies in order to build up large essential patent portfolios. In Bekkers and West (2009), the essential patent landscape for the successor of GSM was studied and compared with that of GSM. An eightfold increase in patents was observed, and a threefold increase in patent holders. The ownership of patents, however, became more concentrated, which suggests increasingly strategic patent use by a small proportion of firms. Interestingly, this study showed distinctly different timing patterns in different types of firms: some major technology contributors owned patent portfolios that had long preceded the standardization efforts, while others basically built up their portfolio while the standard was being specified (see Figure 4). Layne-Farrar (2008) focuses on the "ex-post" patenting activity (i.e. after the first version of a standard is published) in order to distinguish truly innovative contributions from pure strategic patenting. Using essential patents declared at ETSI she finds evidence for some incremental innovation.

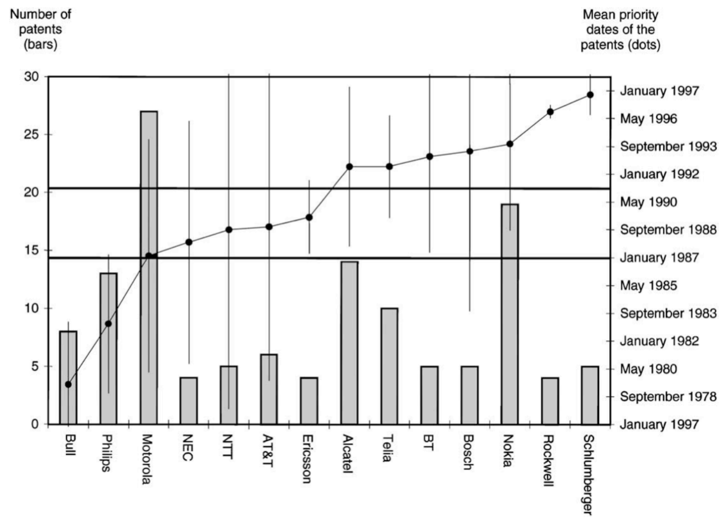


*Figure 3: Timing of firms with regard to essential IPRs in GSM (Bekkers et al., 2002)*

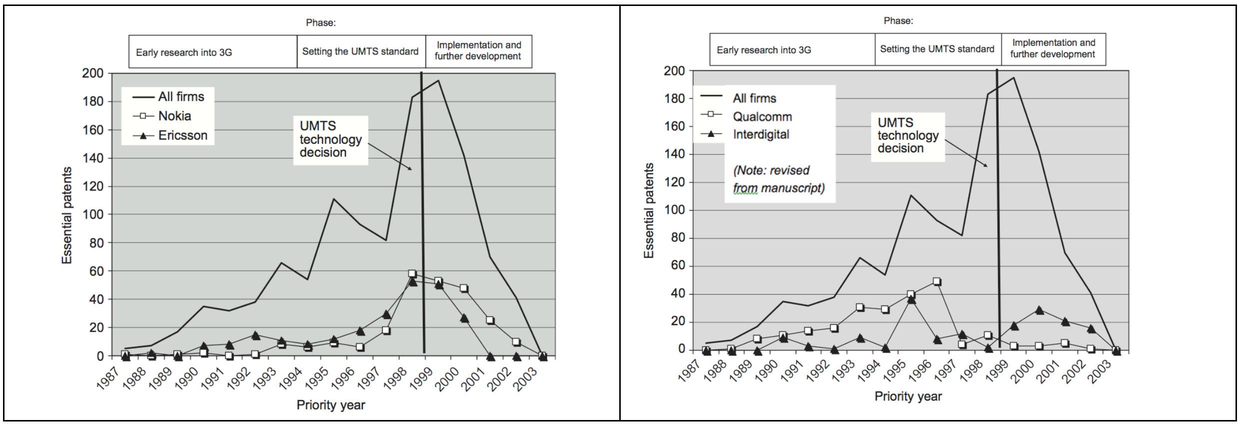


*Figure 4: Timing of firms with regard to essential IPRs in UMTS (Bekkers & West, 2009)*

Such findings prompted questions on the actual determinants of essential patent claims. Were the included technologies chosen on the basis of their technical merit (as suggested by Rysman & Simcoe, 2008) or were they perhaps the result of strategic involvement in the standard setting process by their owners, being able to 'direct the standards towards their own technologies'? These two hypotheses were tested in Bekkers et al. (2011). In this paper, patents claimed essential for the WCDMA standard are compared with a control group. It was found that both hypotheses could be accepted, but that involvement in the standardization process is a stronger determinant than the technical value of the patent in question. Along the same line, Leiponen (2008) finds that political capital acquired through participation in external consortia has a significant impact on the ability to influence technical standard setting.

More recently, scholars started to study technological change and standards from a patent network perspective (Fontana et al., 2009), adopting the concept of technological paradigms and trajectories (Dosi, 1982). Here, the methodology proposed by Hummon & Doreian (1989) was employed in order to identify breakthrough technological contributions from a large set of patents. In Bekkers & Martinelli (2011), the same methodology was applied and the outcomes were compared to essential patent claims (Figure 5). An historical narrative and an analysis of licensing payments confirmed that the so-called top path – the set of breakthrough patents – indeed closely reflected the actual technology development and knowledge position of the companies. Out of the 2987 claimed essential patents at the USPTO and EPO (1729 distinct IMPADOC families), only seven were actually part of the top path of important technological contributions, whereas the top path also included seven patents that were not claimed essential. On the basis of this study, we can conclude that counts of essential patents are not necessarily a good indicator of a firm's knowledge position.

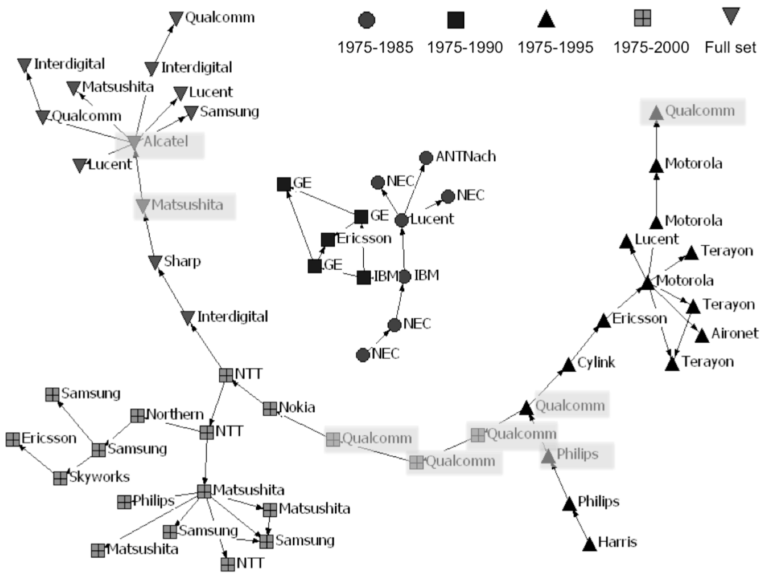


*Figure 5. Main Top Path of patented technologies in 2G and 3G mobile telecommunications. Shaded patents are claimed to be essential by their holders (Bekkers & Martinelli, 2011)*

Finally, some authors have recently studied standard-based patent pools. Although the procedures to include patents in pools are much more demanding than a self-declaration, Baron and Delcamp (2010) showed that firms that are already members of the pool are more able to include lower value patents than 'outsiders'. Layne-Farrar and Lerner (2011) empirically investigate patent pools participation and they find that both firm's characteristics (i.e. presence in the downstream market and symmetry of the patent portfolio) and pool's characteristic (i.e. rent-sharing rule) affect the likelihood of a company to join. The interesting feature of this study is the use of essential patents. They constitute the control sample for the patents included in the pool. In fact, disclosed patents can be considered as patents that potentially could have been included in the pool but they were not.

**4. Implications of Strategic Behaviour in the Disclosure Process**

The studies reviewed above suggest that disclosure is a strategic decision. This creates a number of challenges for any researcher who might use IPR disclosure data to study innovation or standard setting. While firms' incentives to distort the information they disclose will have different implications for different studies, this section lists some of the issues that are likely to arise in any study that hopes to use these data:

1. Some companies submit 'blanket claims', in which they indicate that they do own essential patents (and will comply with F/RAND licenses), but do not provide any identities of the patents they own. No research systematically investigates the incentives for or incidence of these 'blanket claims'. However, discussions with SSO participants suggest two broad reasons why firms are not specific about IPR: (a) Search Costs. Large firms may own tens of thousands of patents, and participate in hundreds of standard setting efforts. For these companies, the costs of conducting a comprehensive patent search for each standards development effort may be prohibitive. Thus, large firms often prefer to make a blanket promise to license any essential patents on RAND terms. (b) Strategic Behavior: Firms often resist disclosing information about unpublished patent applications, arguing that it may give away valuable information about firm's current research projects (Chiao et al., 2007). A blanket RAND claim may also favor firms that hope to have their proprietary technology included in a standard, since the lack of specificity makes it harder for other participants to locate and assess the relevant IPR.

2. Even when firms provide a specific patent or application number as part of their IPR disclosure, the timing of that announcement may be carefully chosen. Most SSOs encourage early disclosure, so standards committees have time to weigh the potential trade-off between technical quality and cost of implementation. But SSOs cannot easily enforce such a rule, and typically have no sanctions to prevent a firm from "disclosing" essential patents long after the costs of switching technologies has increased substantially. The economic literature on patent hold-up in standards recognizes that extremely late IPR disclosures may be worthless, and today's case law is oriented in ruling against companies strategically seeking to hold-up other companies (Farrell et al., 2004). However, standard typically go through many revisions, and it may be hard to tell from publicly available data whether a given announcement represents a "late" disclosure on the original standard, or an "early" announcement on the net iteration.

3. Firms may have an incentive to "over-claim" by disclosing patents that are very unlikely to be essential in order to comply with the standard. One reason firms may over-claim is to increase the value of their patents. Owning several essential patents increases a firm's bargaining power in license sharing and enhances the possibility for cross-licensing. Another possibility is that firms with large portfolios may simply "dump" long lists of patents into the disclosure process rather than conduct a thorough search, since there is no cost of over-disclosing, and it presumably protects them form the larger risk of forfeiting their property rights in litigation. SSOs typically make no effort to check the quality of IPR disclosures, but simply post the information they have received from their members as it arrives. Finally, this issue is further complicated by debates over the precise meaning of "essential." Most SSOs adopt a narrow definition that says an essential patent would necessarily be infringed by any fully compliant implementation of the standard. However, some IPR policies that refer to IPR 'believed to be essential' by its owner, and this has opened a debate between the former "technical" definition, and the latter, which may include "commercially essential" patents that are needed to make a competitive implementation, even if low-quality technical work-arounds exist.

4. Declarations are sought before the standard is finalised, and the final standard might be different from earlier draft versions. Disclosures that were appropriate for a certain drafts might not be essential for the final version of the standard. Similarly, firms may disclose pending patents, and the granted patent may not be as broad as the original application. Since SSOs usually do not require parties to update or withdraw earlier disclosures, such declarations remain in the IPR database.

5. The previous points regard the strategic behaviour related to the duties of SSO's members. However, essential patents can also belong to companies not involved in the standardization procedure. Since IPR policies are only binding for members of a standard-setting organization, firms that own potentially essential patents may forgo membership in the hope of avoiding any F/RAND licensing requirement. An external company owning essential patents will be free to grant exclusive licenses, seek injunctions and set any sort of royalty rate (Bekkers et al. 2002). The cost of this strategy is that non-participants will have a poor sense of whether their technology is actually essential until the standard is published. Nevertheless, as pointed out by Rysman & Simcoe (2009), any dataset of essential declarations will contain both "false positives" (included non-essential patents) and "false negatives" (missing essential patents owned by non-participants).

**5. The Open Essential IPR Disclosure Database**

Disclosure data requires considerable investment in processing and 'cleaning'. To date, almost all studies that have used such data prepared a new, proprietary data set in order to perform the analysis. The authors of this paper have taken the initiative to develop a comprehensive disclosure data set and make it freely available to any interested party. This new data set builds upon several earlier efforts to work with IPR disclosure data.

The overall goals of this project are:

- open: we intend to open-source the data
- harmonized: dealing with the specifics of each SSO and making a common framework
- cleaned: providing all patent identities in a single, standardized way, allowing for matching with patent databases such as EPO/OECD's Patstat or Thomson/Reuters Derwent Innovation Index (DIII)
- wide coverage : currently we plan to include ANSI, ATIS, BBForum, CEN, CENELEC, ETSI, IEC, IEEE, IETF, ISO, ISO/IEC JTC 1, ITU, OMA, and TIA.

The new dataset will retain the structure and the several features of the open dataset already made available by C. Catalini and T. Simcoe at http://www.ssopatents.org/. The central element in the database will be 'disclosures', which are declarations of one or more patents to a given SSO on a given data by one single IPR claimant. The main elements of a disclosure are: (a) a specific combination of SSO, firm, and date, (b) committee, specification or standard, (c) a list of one or more IPRs in the disclosure, and (d) the licensing commitments. The structure of the database will be multiple tables (relational database) in an XML-type format.

In compiling this database, we face several challenges. These are discussed below.

**Challenge 1.** The patent identities provided are often inconsistent, incomplete, and sometimes erroneous. We have to cross-check a number of patent databases.[2] We aim to translate all disclosures into a uniform patent identity at USPTO or EPO with: (1) patent number, (2) published application number or (3) serial application number. Then, each disclosed patent will be matched with the OECD/EPO PATSTAT database.[3] This matching allows us to verify patent identities, identify geographical overlap (via patent families), and add patent metadata like filing data, priority date, family size, granting status, etc. Our target is to have over 90% of all patent identities matched (a full match will not be possible because of erroneous data as well as disclosure of applications that have not yet been published by the patent offices). A particular challenge is the translation of application serial numbers to publications, which is troublesome. The available USPTO correspondence tables are incomplete, whereas the USPTO's PAIR database is not designed to look up larger number of records. We still need to make decisions on whether to include specific types of disclosures (such as provisional applications in the 61/, and 62/ series).

**Challenge 2.** Code all the firm names and licensing commitments. While many disclosures come from large, well-known firms, there are also numerous smaller entities, and some of them are hard to track. Also mergers, acquisitions, joint ventures and patent transfer can complicate decisions on the name of the claiming party. Here, we need to decide how to deal with overlapping claims. These could be the result of 'true' co-assignees, but it is also possible that the patents were traded and the former owner did not withdraw their claims from the database. Even more challenging is the coding of the actual licensing commitment that firms make in their disclosure. While some SSOs allow only one, defined type of commitment, others let the firm decide what commitment they are willing to make. Here, we find complex terms such as FRAND, FRAND-Royalty Free (FRAND/FR), non-assert, reciprocal conditions, field-of-use conditions, voluntary declarations of ex-ante licensing terms, and other specific conditions. In some SSOs, IPR databases are in fact a collection of hundreds of letters from the claimants with no prescribed structure. Finally, some SSOs also allow third parties to submit disclosures on patents (which they do not own themselves).

**Challenge 3:** Link disclosures to a specific committee, specification or standard. Also here, we see great differences between the ways SSOs have included such references in their IPR databases (if at all).

**Challenge 4:** Compiling the database, you need to deal with various types of patent overlap. The four most important ones are (1) Patent overlap between SSO, (2) Patent overlap between different standards within a single SSO, (3) Patent overlap between patent legislations and (4) Patent overlap within patent legislations. With the relational database approach, we aim to include all raw data yet allowing recognition / correction of overlap for a given research question.

We are working towards finalizing the first public version of this database in the second half of 2011. Interested parties are welcome to contact the authors for more details.

---

[2] These include the Thomson/Reuters Web of Science Derwent DII database, EPO's Espacenet database, the USPTO Patent Full-Text and Image Database (PatFT) and the USPTO 'PAIR' database, used to link application serial numbers to publication identities.

[3] EPO Worldwide Patent Statistical Database (also known as EPO PATSTAT) has been specifically developed for use by government/intergovernmental organizations and academic institutions. It has been developed by the European Patent Office, in close cooperation with the OECD. With over 70 million records and with a file size of over 130 GBytes, it is one of the most extensive patent databases currently available.

## References


Baron, J. & Delcamp, H. 2010. Strategic inputs to patent pools, International Schumpeter Society Conference 2010, Aalborg, June 21-24.

Bekkers, R. (2001). Mobile Telecommunications Standards: GSM, UMTS, TETRA and ERMES. Boston, MA: Artech House.

Bekkers, R., Duysters, G., & Verspagen, B. (2002). Intellectual property rights, strategic technology agreements and market structure - The case of GSM. Research Policy, 31(7), 1141-1161.

Bekkers, R., Bongard, R. & Nuvolari, A. (2011, accepted for publication). An empirical study on the determinants of essential patent claims in compatibility standards. Research Policy.

Bekkers, R. & Martinelli, A. (2011). Knowledge positions in high tech markets: trajectories, strategies, standards and true innovators. Paper presented at the 7th European Meeting on Applied Evolutionary Economics (EMAEE 7), Pisa, February 14-16, 2011.

Chiao, B., Lerner, J. Tirole, J. (2007). The rules of standard setting organizations: An Empirical Analysis. RAND Journal of Economics, 38(4), 905-930.

Farrell, J., Choi, J., Edlin, A., Greenstein, S., Hall, B. & Saloner, G. (2004). Brief Amicus Curiae of Economics Professors and Scholars in the Matter of Rambus, Inc. U.S. Federal Trade Commission Docket No. 9302

Hummon, N., & Doreian, P. (1989). Connectivity in a citation network: The development of DNA theory. Social Networks, 11(1989), 39-63.

Dosi, G. (1982). The nature of the innovative process. In G. Dosi, C. Freeman, R. Nelson, G. Silverberg, & L. Soete (Eds.), Technical Change and Economic Theory (pp. 221-238). London and New York: Pinter.

Fontana, R., Nuvolari, A., & Verspagen, B. (2009). Mapping technological trajectories as patent citation networks. An application to data communication standards. Economics of Innovation and New Technology, 18(4), 311-336.

Garrard, G. A. (1998). Cellular communications: World-wide market development. Norwood, MA: Artech House.

Interplay consortium (2010). Study on the interplay between standards and intellectual property rights (IPRs); Preliminary Results. Interim results of a study commissioned by the European Commission. Presented at the Open Workshop organized by DG Enterprise of the European Commission, November 23, 2010.

Iversen, E. (1999). Standardization and Intellectual Property Rights: ETSI's controversial search for new IPR-procedures. Proceedings from SIIT'99 - IEEE Conference on Standardization and Innovation, Aachen, Germany.

Layne-Farrar, A. (2008). Innovative or Indefensible? An Empirical Assessment of Patenting within Standard Setting. Available at SSRN: http://ssrn.com/abstract=1275968

Layne-Farrar, A., & Lerner, J. (2011). To join or not to join: Examining patent pool participation and rent sharing rules. International Journal of Industrial Organization, 29(2), 294-303.

Lemley, M. A. (2002). Intellectual property rights and standard-setting organizations. California Law Review, 90(6), 1889-1980.

Leiponen, A. E. (2008). Competing Through Cooperation: The organisation of Standard Setting in Wireless Telecommunications. Management Science, 54(11), 1904-1919

Rysman, M., & Simcoe, T. (2008). Patents and the Performance of Voluntary Standard-Setting Organizations. Management Science, 54(11), 1920-1934.

Simcoe, T.S. (2005). Explaining the Increase in Intellectual Property Disclosure, in: Bolin, S. (Ed), The Standards Edge, Vol. 3. Sheridan Books.

Simcoe, T., Graham, S., & Feldman, M. (2009). Competing on Standards? Entrepreneurship, Intellectual Property, and Platform Technologies. Journal of Economics & Management Strategy, 18(3), 775-816.