Overcoming inefficiencies of patent licensing

Overcoming inefficiencies of patent licensing: A method to assess a patent's essentiality for technical standards

Rudi Bekkers[*,a], Elena M. Tur[a], Joachim Henkel[b], Tommy van der Vorst[c],
Menno Driesse[c], Jorge L. Contreras[d]

[a]*Eindhoven University of Technology*
[b]*Technical University of Munich*
[c]*Dialogic innovation & interaction*
[d]*University of Utah*
** Corresponding author.*

Abstract

The market for patent licenses, despite its paramount importance for technological innovation, shows various inefficiencies. A particular problem with widely used technical standards such as LTE and Wi-Fi, is the lack of information regarding which patents are 'essential' to implement the standard. This information is crucial because it simplifies determining infringement and implies specific 'FRAND' licensing rules. While many standards-developing organisations stipulate that such patents are explicitly declared, little is known about which are actually essential. The absence of publicly available information on essentiality incurs significant social costs due to the resulting friction in the licensing market. With the growing use and importance of standards to mobility and energy markets, and to the Internet of Things, these costs are likely to increase. Responding to calls from industry, courts, and policymakers, commercial and academic studies have attempted to assess essentiality, but they all have limitations. This paper reports on the technical feasibility of a system of expert assessments for patent essentiality. In an experiment based on a factorial design, experts comprising many patent examiners conducted assessments over 100 working days. Comparing the outcomes to a high-quality reference point shows that sufficiently accurate expert assessments, at a price level that allows large scale testing, are certainly technically feasible, and we identify routes to further improvement.

Disclaimer / funding

This research was part of a larger study on the feasibility of introducing a large-scale system of essentiality assessments, funded by the European Commission's Joint Research Centre, JRC/SVQ/2018/B.6/0024/OC. The larger study is published as (Bekkers et al, 2020a).

Acknowledgements

We would like to thank the six European patent offices that made their examiners available to carry out essentiality assessments in our pilot experiment. We are also very grateful to the SEP owners and others involved in this project through discussions and providing the information required to conduct the experiment. We specifically thank Christian Loyau, Edmund Mangold, Yann Ménière, Ruud Peters, Serge Raes, and Mats Sågfors for their valuable advice, patent pool administrators Avanci, One-Blue, SISVEL, Via Licensing and 3G3P for sharing their extensive experiences with essentiality assessments, and the European Commission for funding and guidance.

1 Introduction

In the field of high-tech products, academic studies focus a great deal not only on the product market but also on the markets for knowledge and related patents (Arora et al., 2001a, 2001b). An extensive body of literature has emerged on the patents required to implement technical standards. Such 'standard-essential patents' (SEPs) are of particular interest: unlike with regular patents, it is not possible for a party making products that incorporate these standards (for instance, for a mobile phone, video player or Wi-Fi device) to design around such patents, thus creating an unusually strong bargaining position for the patent owners. Given this situation and the societal importance of technical standards, the question asked by this Special Issue – "Are intellectual property rights working for society?" – is of particular relevance in the context of SEPs.

Implementors of such standards often face thousands of patents that their owners claim to be potentially standard essential. The market for licenses for these patents shows signs of imperfections, related to (among other things) transaction costs and information asymmetry. More and more calls can be heard for increased transparency to address these market inefficiencies. The European Commission, in its 2017 communication on Standard Essential Patents, stated: '*There is therefore a need for a higher degree of scrutiny on essentiality claims. This would require scrutiny being performed by an independent party with technical capabilities and market recognition, at the right point in time.*' Equally the Commission recognizes that such a system must be balanced against the costs. While highly accurate assessments for single patents are feasible if cost and time are not an issue (as is often the case for single patents examined in the context of litigation), a large-scale system that systematically assesses the essentiality of patents for a given standard, requiring highly qualified individuals, would necessitate appropriate costs and time. Without pre-empting who would finance – or co-finance – such a large-scale system, and the willingness (and self-interest) of these parties to create budget for this, we are talking of potentially thousands of patents to be examined for standards for mobile and wireless communications and therefore assume any financier will be, to some degree, sensitive to cost. Arguably even more important is capacity: these assessment tasks require highly qualified staff, and hence, resources are limited. Should thousands of patents be assessed in a reasonable amount of time, then time efficiency is paramount.

Motivated by the above, this paper investigates whether essentiality assessments can be sufficiently efficient (in terms of time and costs) as well as sufficiently accurate, to enable the establishment of a large-scale system of such assessments, and thus overcome important inefficiencies in the SEP licensing market. Considering the complexity of the assessment task, but also the need for efficiency, this paper focuses on assessments that each require roughly eight hours work (i.e. a budget of one day), acknowledging that existing some assessments, while considered accurate, utilize five to ten times more resources (see Section 3.2 below).

We formulated the following research questions:

> Q1. Given an average time budget of eight hours per patent, how accurately can qualified assessors determine the essentiality of a given patent for a specific standard?
>
> Q2. How does the availability of claim charts[1] and the chosen definition of essentiality affect the above relationship?

[1] In the context of this paper, the term 'claim chart' refers to a document produced by the patent owner that maps, in detail, the claim(s) and claim feature(s) in the patent to those sections of the standard document to which they are believed to be essential.

Note that we focus on the 'technical feasibility' of a large-scale system of essentiality testing. For questions concerning 'institutional feasibility', such as who should set up such a system, who should carry out the assessments, who would finance it, and whether there is sufficient support from stakeholders, we refer to the complementary work presented in (Bekkers et al, 2020a). While we focus on patent essentiality, we acknowledge that also patent validity, enforceability, infringement, and patent value are important aspects in terms of transparency and licensing discussion, and sometimes related to essentiality[2], these are out of the scope of this study.

In an experiment based on a factorial design, twenty-eight experts, including many patent examiners, performed 109 assessments over a total of 100 working days. Using patent pool data as a high-quality reference point, we obtained not only data on patents accepted by pools, but also on those rejected by pools and the full information set provided at that time to pools – including claim charts prepared by the patent owner. This setting enabled us to replicate the earlier pool assessment as accurately as possible. Based on our analysis, we conclude that sufficiently accurate assessments, at a price level that allows large scale testing, are certainly technically feasible, and we identify routes to further improvement.

In Section 2 we discuss market inefficiencies in the licensing of patents, focussing on SEPs. Section 3 looks at existing attempts to determine essentiality, including patent pools, commercial and academic studies, and court case analyses, and we review recent endeavours using Artificial Intelligence (AI) approaches. Section 4 presents the experiment design and data for our research, followed by the results in Section 5, then conclusions and a discussion in Section 6.

2 Market inefficiencies in licensing standard-essential patents

2.1 The imperfect market for patent licenses

As economists have agreed since Adam Smith (1774), well-functioning markets increase efficiency through specialization and division of labour. However, efficient markets require parties to have full information. This assumption is frequently violated by information asymmetry, a lack of transparency, and uncertainty about future events. The classic example is Akerlof's (1970) 'Market for Lemons.'

In the market for technologies (MFT) (Arora et al., 2001a, 2001b), potential efficiency gains, but also market imperfections, are particularly pronounced. Efficiency gains arise, for instance, in the division of labour between start-ups generating new technology and the large incumbents commercializing them (Baumol, 2010), or through general-purpose technologies that can be fruitfully employed in many markets (Bresnahan & Trajtenberg, 1995). Accordingly, the MFT has acquired high relevance for innovative industries.

A challenge with trading knowledge is that in many cases it can easily be expropriated. Intellectual property rights, and patents in particular, can alleviate this problem by increasing the appropriability

[2] There can be a complex relation between essentiality and validity. A broad interpretation of the claims in a patent may lead a patent to be essential; however, that same broadness might actually render the patent invalid. Patent essentiality and infringement are related but not identical. Many standards define the behaviour of different device categories (for instance mobile phones, and network switching equipment, or CD discs and CD players). A patent that is essential to the standard may in fact only be required for network switching equipment and will not be infringed by a mobile phone that fully complies with the standard. Moreover, standards often define optional features. While many SDOs have definitions of essentiality that include patents required for optional features, an implementer that chooses not to include this optional feature will not infringe the related essential patents for this feature.

of knowledge (Lamoreaux & Sokoloff, 1999; Arora et al., 2001a, 2001b; Arora & Ceccagnoli, 2006; Gans et al., 2008). Accordingly, the market for patent licenses (including, for the sake of simplicity, both patent licenses and assignments) has acquired a high importance in its own right (Madiès et al., 2014), and scholars often measure transactions on the MFT through patent transactions (Lamoreaux & Sokoloff, 1999; Gambardella et al., 2007; Serrano, 2010). In fact, for publicly available but patent-protected knowledge, the relevant market is not the MFT, but the market for patents (Fischer and Henkel, 2012: 1531).

The market for patent licenses shows various imperfections due to transaction costs, both in terms of motivation and coordination costs (Milgrom and Roberts, 1992: 29). Motivation costs cause inefficiencies through information asymmetry (Caves et al., 1983), which in the market for patent licenses is typically bilateral. The patent owner might have private information about the value of the invention, its limitations, or required complementary knowledge not evident from the patent text. In turn, the prospective licensee will generally know more about the invention's potential applications and resulting economic value. To some extent, the consequences of such information asymmetries can be mitigated by a suitable choice of licensing terms (Gallini & Wright, 1990; Beggs, 1992; Macho-Stadler et al., 1996), but inefficiencies remain.

In addition, coordination costs arise due to uncertainty and lack of transparency. To start with, patent applications are typically not disclosed to the public until 18 months after filing. Often a considerable amount of time passes before a grant decision: at the EPO, a patent grant comes on average 6.0 years after the filing date,[3] and only then are the final patent text and scope known. Even then, a granted patent may be 'latently invalid' due to prior art not found by the examiner or disclosed by the applicant (Farrell & Shapiro, 2008; Miller, 2013; Henkel & Zischka, 2019). Furthermore, being written in natural language, a patent allows room for interpretation (Bessen & Meurer, 2005). As information and communication technologies (ICT) are complex, it may be difficult to ascertain whether a given patent claims features of a particular product (Lemley & Shapiro, 2007; Magliocca, 2007). Finally, the sheer number of patents in a given field may hinder identifying those patents that are relevant to a given technology. This problem is aggravated by the issuance of patents on trivial (e.g., Reitzig et al., 2007: 147) or non-novel inventions (e.g., Graham & Mowery, 2003: 226). These difficulties complicate licensing negotiations and involve the related issue of identifying potential licensors.

The inefficiencies in the market for patent licenses can have serious welfare consequences, including the under-utilization of existing technologies, inadvertent or intentional infringement, and unfounded or, in the case of inadvertent infringement and subsequent lock-in, excessive royalty demands (e.g., Jaffe & Lerner, 2006; Lemley & Shapiro, 2007). These consequences are particularly severe when it comes to widely used, standardized technologies, as we discuss next.

2.2 Specific problems posed by licensing Standard Essential Patents

A significant number of product categories in the ICT sector typically use of interoperability standards. Every smartphone, tablet and laptop sold today incorporates hundreds of different standards (Biddle et al., 2007). Likewise, hundreds of different firms and research organizations collaborate to develop widely deployed telecommunications standards such as UMTS (3G), LTE (4G), 5G, and Wi-Fi. At the same time, these standards are implemented by a large number of equipment makers. The rise of the Internet of Things (IoT) and smart technologies in the context of Grand

[3] For all patent applications at the EPO between 2001 and 2008, the grant delay was on average 6.02 years (with annual variations averaging 0.2 years). Subsequently, as a result of truncation, the average gradually dropped to 4.0 years in 2014. Source: authors' calculations based on PATSTAT 2020a, comparing filing date with day of B1 publication.

Societal Challenges is expected to broaden the application domain for telecommunications standards-based technology, and thereby the number of implementers, significantly.

Interoperability standards are typically developed in voluntary associations known as standards-development organizations (SDOs), where firms collaborate in developing standards of interest to the industry. As we explain in Section 3, most SDOs in the ICT sector stipulate in their policies that for any (known) patents required to implement their standards, the patent owner must has committed to offer licenses on terms that are 'fair, reasonable and non-discriminatory' (FRAND). These FRAND licenses may be royalty-free or royalty-bearing. Royalty-free patent licenses such as with standards for Bluetooth, HTTP, TCP/IP, and USB, are comparatively non-controversial. However, if FRAND licenses are royalty-bearing, SEP licensing takes on the characteristics of a market transaction.

The market for SEPs has a number of peculiarities. First, although standardized products presumably implement a large number of SEPs, licenses are often not sought or finalized until months or years after products are on the market (Contreras, 2013: 59-62). Second, any given mobile telecommunication product is likely covered by thousands of patents declared (potential) SEPs (Baron & Pohlmann, 2018; Bekkers et al., 2020). Yet due to broad or ambiguous claims and the complexity of the technology, it is often unclear which patents precisely cover a given product, component or standard. Finally, it is often uncertain whether patents declared as 'potentially essential' (see Section 3) for a given standard are, in fact, essential, even assuming they are otherwise valid and enforceable. This latter issue is crucial, given that non-essential patents can often be worked around or omitted from a product, unlike essential patents whereby a product by definition has to comply with the respective standard and thus will need to use these essential patents (unless these patents are only essential to a not implemented optional feature, or only essential to other device categories). Thus, reliable information about a patent's essentiality goes a long way toward creating transparency regarding its use in a given product.

Yet, as discussed in Section 3.1, studies have reported significant 'over-declaration' of SEPs relating to mobile telecommunications (see studies referenced in Contreras, 2019: 211), whether or not intentional. It is believed that many, if not the majority, of patents declared essential to standards in the mobile telecommunications industry are actually *not* essential. The widespread occurrence of SEP over-declaration creates significant inefficiencies in the market for SEPs. This seems at least partially driven by information asymmetry: owners of (potential) SEPs usually have intimate knowledge about their own patented inventions and whether they are likely to be essential or not. Implementors, on the other hand, are confronted with dozens of SEP holders with thousands of patents, and typically have limited or no knowledge about the details of individual patents claiming to be SEPs. This asymmetry is complicated by long supply chains where products implementing standards range from generic chips, application-specific chips and modules, to intermediate and end products. Downstream firms in these supply chains are often less knowledgeable about the technologies covered by the SEPs in question, which are often implemented in modules or chipsets (see Henkel, 2021).

Asymmetric information and associated uncertainties hamper licensing negotiations for SEPs and invite opportunistic behaviour, creating friction on the MFT and reducing societal welfare. Recognizing these issues, the European Commission (2017: 5) states in its Communication 'Setting out the EU approach to Standard Essential Patents': *'Evidence points to the risk of broad over-declarations and makes a strong case for more reliability with respect to SEP essentiality. Stakeholders report that recorded declarations create a de facto presumption of essentiality in negotiations with licensees. This scenario places a high burden on any willing licensee, especially SMEs and start-ups, to check the essentiality of a large number of SEPs in licensing negotiations. There is therefore a need for a higher degree of scrutiny on essentiality claims.'*

A final inefficiency arising from declarations of patent essentiality is the lack of a consistent standard for defining essentiality among SDOs. Bekkers & Updegrove (2012: 35) identify thirteen different features of essentiality definitions in twelve major SDOs (see also Contreras, 2007: 12, on additional variants and exclusions). There are two major definitional axes where SDO essentiality definitions differ: (1) the degree to which they cover *optional* portions of a standard, and (2) whether they refer to '*technical'* or '*commercial'* essentiality (commercial meaning that alternatives to the patented solution technically exist but are unattractive or even prohibitive for cost reasons). Yet these terms' definitions are vague (Contreras, 2017: 218-219), leading to uncertainty among patent holders regarding which patents to declare as essential, and causing implementors to question whether patents declared essential in one SDO would be also under another SDO's policies.

These problems are particularly serious given the uptake of IoT applications, where compared to smartphone market implementors, IoT implementors are much more numerous and heterogeneous, yet much less knowledgeable about the SEPs that might cover standards used in their IoT devices (Henkel, 2021). Also the growing use and importance of standards to mobility and energy markets, where prospective licensees have less knowledge about patents that may or may not be essential and where different business cultures exist, increases tension and transaction costs. Thus, processes to reliably assess the actual essentiality of declared SEPs are urgently needed to ensure an efficient SEP licensing market. We review existing attempts at large-scale essentiality assessment, before introducing our own study.

It is important to acknowledge that full information on essentiality does not remove all possible sources of information asymmetry surrounding licensing essential patents. Other important topics are the validity of patents, the technical merit of the patented technology (impacting the appropriate monetary compensation), and the patents' enforceability (which depends on whether patents are granted, not expired, and their renewal fee being paid, also affected by claims of inequitable conduct, patent misuse, competition violations, etc.). These are all important dimensions in licensing discussions, and for each of them, more transparency may also lower transaction costs. This said, we believe that information on essentiality is the first step in increasing transparency, and other dimensions would come later (when it is determined whether or not a patent is essential in the first place). In that sense, we see essentiality testing as a necessary, but not sufficient, step toward reducing information asymmetry and lowering social costs.

3 Existing attempts to determine essentiality

Recognizing the nature and consequences of inefficiencies in the markets for SEPs, various studies have attempted to determine the essentiality of patents for given technical standards. While we can learn from these attempts, none to date has provided the market with an assessment of a known accuracy level *and* a cost per patent that would make it feasible to assess all patents potentially essential for a standard. We first discuss how declarations of potential essentiality are made pursuant to SDO IPR policies (Section 3.1), then recent attempts to assess essentiality using expert-based approaches (Section 3.2), Artificial Intelligence (AI) and other automated approaches (Section 3.3).

3.1 Procedures for declaring SEPs and resulting over-declaration

In most SDOs, the disclosure of a patent as 'essential' to a standard is based entirely on the self-declaration by the patent holder. But a self-declaration does not necessarily mean that a patent is *actually* essential, and the studies discussed below have found significant 'over-declaration' of SEPs, particularly at SDOs focusing on mobile telecommunications (e.g., Contreras, 2019: 211, collected studies). For example, studies of the GSM (2G) standard and the 3GPP standards for 3G and 4G have found over-declaration rates between 8% and 58%, with individual patent holders over-declaring at rates as high as 82% (see Unwired Planet [2017 EWHC 711 at 324-329], citing numerous studies).

Such over-declarations may be unintentional. A patent holder may, for instance, have submitted a technical proposal to an SDO and accurately declared patents covering the proposal as essential, after which the standard may have evolved to exclude the technology from the proposal. Or a party may have made a declaration on the basis of a patent *application*, and the patent that is ultimately granted no longer includes those patent claims that covered the technology included in the standard (see Lerner, 1994; Marco et al., 2019, observing that the scope of an issued patent is often significantly narrower than the original patent application). Or a patent holder, unsure whether a patent is essential, may prefer to err on the side of over-declaration, given the significant legal consequences of not declaring an essential patent (e.g. punished by court for nor respecting SDO IPR policies) and the few legal consequences for over-declaration (Contreras, 2017: 223).

However, opportunistic, deliberate over-declarations also occur. A firm's share in overall royalties that can be charged for a standard is often approximated, for lack of a better criterion, based on its numerical share of declared SEP families (Unwired Planet [2017] EWHC 711, at 182 (portfolio strength is often approximated by "some kind of patent counting")). Accordingly, patent holders have an economic incentive to over-declare – a form of strategic behaviour (Dewatripont & Legros, 2013; Bekkers & West, 2009, Aoki & Arai, 2018). As noted by Justice Birss in *Unwired Planet v Huawei, 'Very many more patents are declared to be essential than in fact are essential'*, and '*it must also be recognised that the fact that rates are negotiated by counting patents creates a perverse incentive to declare as many patents as possible, making over-declaration worse.*'[4] Finally, even if a patent owner is convinced of the essentiality of a given patent, this might not actually be the case. Lemley & Simcoe (2019) show that a substantial proportion of declared essential patents are found non-essential even after they have been carefully chosen for litigation (which arguably indicates their owner believes they are essential).

For the sake of completeness, we must also acknowledge that there are reasons why some essential patents are not declared as potentially essential. First, parties that are not participants in the relevant SDO are not under any obligation to declare potentially essential patents at that SDO (Contreras, 2016, discusses the concept of standards 'outsiders'). Second, SDO disclosure policies are not 'absolute', and what needs to be declared may depend on actual participation in relevant working groups and on the knowledge of individual participants – even if room to 'manoeuvre' might be limited by the 'good faith' or other requirements that SDOs have in their policies or that are imposed by law (see Bekkers & Updegrove, 2013: 78-80, 82-85).

3.2 Large scale essentiality tests using expert-based approaches

Given the uncertainties and inefficiencies relating to the declaration of SEPs noted above, various mechanisms have been developed to assess the essentiality of sets of patents for a given technical standard. Various parties have conducted such analyses for different purposes. We review these existing approaches and focus on larger scale assessments where hundreds or at least dozens of patents are assessed.[5]

The first patent pools for technology standards, such as those for the MPEG-2 video compression standard and the DVD (digital video disc), appeared in the 1990s (Den Uijl et al., 2013). To ensure

---

[4] Unwired Planet v Huawei [2017] EWHC 2988 (Pat).

[5] We did not review company in-house assessment mechanisms because this information is not shared publicly.

their operations were compatible with competition (antitrust) law, some of these pools sought Business Review Letters from the U.S. Department of Justice Antitrust Division. The analysis conducted under these letters concluded that in order to prevent anti-competitive effects, it was important to ensure that patents included in a pool were essential to the standards in question, and that the patented technologies in the pool complemented rather than substituted one another (Gilbert, 2017).[6] As a result, almost all pools set up formal mechanisms for assessing the essentiality of patents proposed for inclusion in the pool. Typically, these assessment procedures (1) require patent owners to propose patents and submit claim charts that demonstrate why the proposed patents are indeed essential, (2) outsource the assessments to independent, external experts (usually at specialized law firms), and (3) have formal appeals procedures for patent owners and – sometimes – for other pool members and/or licensees. While details of the procedures used by specific pools are usually not made public, an interesting exception is the 3GPP, also known as the 'WCDMA pool' or '3G Patent Platform', whose initiators published an extensive book describing their approach (Goldstein & Kearsey, 2004). Among other things, this pool (part of the reference set we discuss below) involves independent, parallel assessments and we compare their results. A recent study commissioned by the EC includes a review of essentiality assessment mechanisms in patent pools (Bekkers et al., 2020a; see also Merges & Mattioli, 2017).

Pool organizers have extensive experience with such essentiality assessments. Given the strong legal incentives to include only essential patents, pools may be expected to implement diligently performed, high-quality essentiality assessment mechanisms. Also, a pool licensing a portfolio of SEPs for a certain standard should be motivated to build a reputation of licensing actual SEPs; and existing pool members should not be willing to accept a dilution of their portfolio share through newly added patents unless these are actually essential (Merges & Mattioli 2017; for royalty allocating mechanisms in pools, see Layne-Farrar & Lerner, 2011). The resources they use per patent are in line with such a high-quality assessment.[7] This does not mean that pool assessments are perfect. There is inherent uncertainty associated with essentiality assessment. In cases of uncertainty about the actual essentiality of a patent submitted by a member, patent pools may be subject to incentives to include that patent, which would lead to over-inclusion. Also, communication between a submitting member and the evaluator as well as appeal opportunities may lead to over-inclusion. And on an aggregate level, a pool might benefit from more lenient inclusion criteria as it increases its portfolio size compared to other patent holders, thus justifying higher royalty demands.

In addition to essentiality assessments for the formation of patent pools, third parties have assessed the essentiality of patents to different standards. Sometimes, this work has been carried out by academics, who often publish it openly, but more often it is by private consulting firms that only make the results available to parties purchasing their reports. One of the first (published) attempts was by Goodman & Myers (2005), conducted in the context of a conflict between several companies over patent portfolio value (and also sponsored by one of these companies). Many later studies

[6] A specific case is where a standard can only be implemented by using one of a set of two (or more) technologies, each patented, making the patents in this set substitutes. In the ETSI definition of essentiality (See Section 4.1), such patents are finally deemed essential, but these are exceptional cases; neither during the 50+ hours of our face-to-face talks with experts regarding their work, nor from the feedback we collected from the assessors in our experiment, was there any indication that this was happening in practice (even though we asked about it explicitly).

[7] These resources are estimated at EUR 5,000 to 10,000 for a single European patent, and up to twice as much for a single US patent. The price per evaluation is typically pre-agreed with the experts that perform it and is based on the actual costs based on the average time spent. (See Merges & Mattioli, 2017 and Bekkers et al., 2020a).

refer to this work, but it was also criticised (see Martin & De Meyer, 2006). From around 2007, a stream of commercial studies followed this approach, including Fairfield Resources International (Fairfield, 2007, 2009, 2010), which are continuations of the Goodman & Myers (2005) study, as well as studies from Article One Partners (2011), Cyber Creative Institute (2011), Jefferies & Company (2011), iRunway (2012), PA Consulting Group (2015) and Charles River Associates (2016). These studies mostly begin with lists of patents declared as potentially essential – for instance patents declared to ETSI – and perform manual assessments of essentiality. It is difficult to assess the quality of these efforts: the underlying methodology, working assumptions and data processing steps are not generally made public, and there is no evaluation, such as a comparison of the results with a benchmark of known accuracy, or inter-rater consistency. Validity of outcomes is not extensively discussed in these works. It is also difficult to compare the outcomes of these studies with each other since they differ in terms of the standard investigated, data selection and cut-off dates.

Finally, large-scale essentiality tests have been performed by economics experts in the context of patent litigation. Such cases include *Unwired Planet v Huawei, TCL v Ericsson*, and *In re Innovatio IP Ventures*.[8] The purpose of such analyses was usually to provide input to a court's top-down calculation of FRAND royalties, where the aggregate royalty for all SEPs covering a specific standard is first determined, and then allocated among individual SEPs and SEP holders (Siebrasse & Cotter, 2017). Compared to the analysis by consulting firms discussed above, litigation analyses are somewhat more transparent, yet their procedures vary considerably in design and parameters. While the courts in these three cases recognised that the large-scale essentiality assessments performed by experts were not perfect, they were useful to the courts for their FRAND royalty determinations.

3.3 Large scale essentiality tests using AI and other automated approaches

The use of automated approach, including Artificial Intelligence (AI), has attracted attention from scholars, and we also see emerging commercial interest such developments (for instance, see IPLytics, 2021). Inspired by work on the computation of semantic similarity between patents (Younge & Kuhn, 2016; Arts et. al., 2017), Brachtendorf et al. (2020) investigated the semantic similarity between patents and standards documents to assess actual essentiality of declared SEPs.[9] The algorithm is validated by comparing the findings with the results of the manual essentiality assessments for the TCL v Ericsson court case mentioned above. At the individual patent level, the consistency the Brachtendorf study found between its outcomes and the court case data is limited. From the set of 166 patents assessed essential by manual evaluators, the automated system predicted only 40 (24%) were essential. From the set of 236 patents assessed not essential by manual evaluators, the automated system predicted 216 (92%) were not essential. If we assume the reference point is perfect (which it may not be), then the automated system has many false negatives, and fewer false positives. Yet the authors found strong and highly significant correlations between the experts' decisions on standard essentiality and their measurement of semantic similarity, and good accuracy in predicting the share of actual SEPs in a larger portfolio. All in all, these initial results are promising but not yet satisfactory in terms of predicting essentiality on an individual patent basis.

[8] Unwired Planet v Huawei, [2017] EWHC 711 (Pat); TCL Commc'n Tech. Holdings, Ltd. v. Telefonaktiebolaget LM Ericsson, No. 8:14-cv-00341 (C.D. Cal. Dec. 21, 2017), rev'd on other grounds, 943 F.3d 1360 (Fed. Cir. 2019); Innovatio IP Ventures, LLC Patent Litigation, 2013 U.S. Dist. LEXIS 144061 (N.D. Ill. Sept 27, 2013).

[9] The authors identified standards documents based on patent declarations at ETSI, resulting in 4,796 standards documents, and compared them with 37 million patent documents, considering patent claims as well as technological descriptions. The study used an algorithm developed by Natterer (2016).

An undeniable strength of automated approaches is their scalability. Potentially, they allow the analysis of very large sets in a relatively short time span, and at low costs. But they also come with inherent limitations. Firstly, the meaning, interpretation, and precise scope of words and terminology (both in patents and standards) depend on the context, making it hard to automate. Second, semantic approaches can face difficulties dealing with changes in terminology over time. Third, the patent to be evaluated, or parts of it, may be written in a different (natural) language than the respective part of the standard. Furthermore, even with the same natural language, the vocabulary in patents (drafted by patent attorneys) often differs from standards (drafted by engineers). Fourth, a technology or solution required to implement the standard may not be explicitly mentioned in the standard's text but may still be required in order to satisfy the standard (i.e., be implied by the standard). Fifth, an essentiality analysis should consider all existing (patented[10] and unpatented) technical alternatives that may also satisfy the standard and may render the patent under investigation non-essential. For instance, when a standard requires a quasi-random code for some function, a certain patented technology may indeed be able to generate such a code, but the existence of other solutions that also generate a code that would satisfy the requirements of the standard must also be considered. An experienced human assessor, well-trained in the technical field, would be expected to have such knowledge. Current AI systems in this field, however, merely compare a focal patent and a focal standard. Adding the whole universe of external solutions would be challenging given the current state of relevant AI technology. Sixth, any automated system is prone to gaming, whereby patent owners, anticipating the workings of such a system, adapt the wording of their patent applications (which might end up in the granted patent claims) to the wording of standard documents.[11]

In sum, these approaches to essentiality assessment provide useful input in terms of designing an essentiality testing mechanism, but do not yet answer the question of whether essentiality assessments can be made sufficiently efficient (in terms of time and costs) as well as sufficiently accurate, to set up a large-scale system and thus overcome important inefficiencies in the market for SEP licensing.

4 Experiment design and data

In this section, we discuss the experimental design and associated key choices, elaborating on the definition of essentiality used, the reference points, the selection of assessment cases (and associated data collection), the assessors, case allocation and instructions.

---

[10] Whether the existence of patented alternatives remove essentiality depends on the exact definition of essentiality adopted by the SDO in question. The definition of essentiality at ETSI is explicit on this aspect: if alternatives exist that are not patented, the patent in question is not essential; if only alternatives exist that are also patented, then the focal patent is essential (as well as the patented alternatives). (ETSI, 2020: Annex 6, §15, Item 6; see also Contreras, 2017: 218). Rules at other SDOs may differ or are not explicit (Bekkers & Updegrove, 2013: 66).

[11] For example, the U.S. Federal Trade Commission found that semiconductor designer Rambus, Inc. deliberately modified the claims of its patent applications during prosecution better to cover technology features that were concurrently being designed into JEDEC standards. In the Matter of Rambus, Inc., FTC Docket No. 9302, 2006 FTC LEXIS 101, *6 (FTC, Aug. 20, 2006) ("through its participation in JEDEC, Rambus gained information about the pending standard, and then amended its patent applications to ensure that subsequently-issued patents would cover the ultimate standard"); see also id. at *88-89 (describing meetings between Rambus CEO and its JEDEC representative concerning "how Rambus might add claims to cover JEDEC standards").

The experiment's quantitative part follows a factorial design, a type of experimental design that allows to separate the effect of several factors (and their interactions) on a response variable, even when the number of observations is limited. Moreover, it allows to distinguish between factors, which are controllable and of interest, and blocks, which are the other variables which are controllable and likely to affect the outcome but are not of interest (in technical terms, 'nuisances'). All uncontrollable nuisances can be accounted for by randomizing. In our experiment, the treatment is whether a patent has claim charts or not, and the block is whether the patent was considered essential by a pool. We also gathered qualitative outcomes by asking the assessors both closed questions and open feedback. We did so after each assessment, as well as at the end of all assessments. To ensure our assessors had in-depth expertise in the field of the standards and patents they were reviewing, we focused the experiment on a single technological area, namely ETSI/3GPP 3G and 4G standards, and selected assessors accordingly. This technical area is one of those primarily making the calls for essentiality testing (European Commission, 2017). Furthermore, there are several patent pools active in this area, allowing us to use their essentiality decisions as reference points. There are unobserved variations between cases that we cannot control, such as the level of difficulty, so we randomized the data in each cell to limit the impact on our results.

4.1 Definition of essentiality

SDOs and other organisations have adopted different definitions of essentiality (see Section 2.2). Since our experiment considered assessments of patent essentiality for ETSI standards, we in principle followed that organisation's definition of essentiality throughout the experiment: *'ESSENTIAL' as applied to IPR means that it is not possible on technical (but not commercial) grounds, taking into account normal technical practice and the state of the art generally available at the time of standardisation, to make, sell, lease, otherwise dispose of, repair, use or operate EQUIPMENT or METHODS which comply with a STANDARD without infringing that IPR. For the avoidance of doubt in exceptional cases where a STANDARD can only be implemented by technical solutions, all of which are infringements of IPRs, all such IPRs shall be considered ESSENTIAL.'* (ETSI, 2020).

While preparing the experiment and discussing this definition with the various patent offices involved, some offices raised concerns that patent examiners were not trained in determining infringement and asked if the assignment could be re-phrased. Together with these offices, we developed an alternative that we call a 'novelty-based test', based on the following: *in the hypothetical case that the standard document had already been published before the patent priority date, would that document have been novelty-destroying?* We stressed that we were *not* asking patent examiners to determine whether the patent is *valid* or not; after all, the text of the standard document is in reality published *after* the patent priority date. The various definitions were included as an additional block in our factorial experimental design. Several stakeholders indicated that they did not expect the specific definition to make a difference in essentiality assessments; nonetheless, in our analysis we compared the essentiality assessments based on the ETSI definition to those based on the novelty-based test.

4.2 Source and characteristics of the patents to be assessed

In order to interpret the outcome of the experiment, it is important to understand the features of the set of patents that are to be assessed. These features will depend on the source of these patents. This is a design choice for a large-scale assessment system (Bekkers et al, 2020a), where one can distinguish three main scenarios:

a. A very broad, possibly even prohibitively large set of patents selected on the basis of technical characteristics that make them potential candidates for essentiality. This scenario might be selected when the relevant standards come from SDOs that do not require parties to declare specific patents ('blanket disclosures').

b. The (smaller) set of all patents that have been declared potentially essential to the standard. Those might still run into thousands, up to 10,000 patents for a large standard.
c. A narrower set of patents that is selected by the respective owner for investigation, who will select these on its own believe these patents are actually essential.

The average difficulty of correctly assessing the patents in each set increases from scenario (a) to scenario (c). While the set of patents in scenario (a) will contain many patents that can easily be identified as non-essential, patents in scenario (c) are the most challenging ones since they were selected by their owners as being presumably essential. Still, we decided to investigate scenario (c) since, in terms of organizational feasibility, it was found in a recent study to be the most realistic one (Bekkers et al, 2020a). This choice implies that expected accuracy will be lower than had we chosen scenario (a) or (b). In turn, this means that our outcomes in terms of accuracy constitute a lower bound for any of the designs.

4.3 Reference assessments

To determine the accuracy of the assessments in this experiment, a reference was required for comparing our results. The ultimate, authorised decision concerning essentiality (and infringements) lies with competent courts. While some courts have indeed issued (public) verdicts on the essentiality of patents, the number of data points is very limited and may be based on different definitions of essentiality than assumed here (e.g., take only infringement into account). Moreover, we have little insight into the exact information used to reach that court verdict - making it hard to ensure our assessors would work from the same information. Moreover, there is the risk that our assessors were aware of these court verdicts, thereby creating a possible bias.

For our experiment, we used what we believe are the most accurate assessment points existing *outside* a litigation context: the assessment by patent pools. To comply with competition (antitrust) law, these pools have developed diligent and sophisticated procedures where patents submitted to the pool are scrutinised by external, independent parties (usually law firms or patent attorneys specialised in this task), as discussed in Section 3.2. While pool assessments cannot be regarded as ‘perfect’, they are considered by almost all stakeholders as the gold standard, and we believe they are appropriate as a reference point for our study. If an expert can replicate the pools’ assessments, this is a strong indicator of a high accuracy level. Furthermore, by collaborating with patent holders who submitted their patents to pools, we were able to ensure that the assessors in our experiment would receive no more information than the pools used, and are assessing patents against precisely the same version/release of the standard, etc.

4.4 Selection of assessment cases and associated data collection

To perform the experiment, we developed a sample of cases, where ‘case’ refers to a combination of a granted patent document and a (specific release of a) standard document (e.g. TS 25.211 V2.5.0). While ‘positive’ reference cases can be easily identified using public information by pools on which patents were determined essential, ‘negative’ reference cases required a different approach. Ideally, we wanted to know which patents were actually submitted to pools, but then rejected. To obtain such information, we sought collaboration with patent holders, and, after negotiations, several patent holders participating in pools were willing to share that (private) information. Moreover, we also found them willing to share the claim charts that they actually submitted to the pools, for both accepted and rejected patents, allowing us to provide exactly the same information to our assessors as provided to the pools. The very confidential nature of these claim charts did require non-disclosure agreements between all involved parties (it helped that companies had great confidence in the professionalism and confidentiality of the patent examiners who took part in our experiment).

Our experiment focused on patents essential to the ETSI/3GPP 3G and 4G standards, and used data from the following pools, that all have a licensing program for these patents: the ‘WCDMA’ patent pool, the Sisvel LTE/LTE-A patent pool, the Via Licensing LTE patent pool, and the Avanci patent pool. Given that we engaged patent examiners from European patent offices, we only included EPO patents in the experiment. While granted EPO patents always include patent claims in the English language, the other text in the document may be in German and French, which are also official EPO languages. Our selected patents also reflected that.

Our final data set has four categories, as shown in Table 1. Categories I and II are based on the data provided by patent owners, discussed above. There are two pertinent points concerning category II: firstly, because companies usually review their patents internally and only submit patents to pools which they believe likely to succeed, this set is smaller than Category 1. Secondly, because of this preselection, these patents may be relatively more difficult to assess. We will return to this in our analysis and conclusion. Categories III and IV are based on public data, and complement the above data to ensure we had the required number of patents for our factorial design. Category III are patents publicly disclosed by the pool as essential. We had no claim charts for these, so we used them for the cases where we did not plan to provide our assessors with claim charts. Category IV is the most challenging, because patent pools do not publish the identity of patents that were submitted but rejected. We had to reconstruct this category, by creating a set of patents very similar to the one in the actual pool, using a series of defined criteria.[12] By having to reconstruct this set, we acknowledge that the data quality may be lower than in the other three categories. (As discussed in Section 4.6, we paid special attention to any signs of potential issues with patents in this category.)

Table 1. Data sources

	Patent included in pool	Patent not included in pool
Based on data provided by patent owner	**Category I** - Data sources: pool acceptance information supplied by patent owner (verified by public pool information); claim charts supplied by patent owner - Data quality: very high - Assessment difficulty: average	**Category II** - Data sources: pool rejection information and claim charts supplied by patent owner - Data quality: very high - Assessment difficulty: high
Based on public data	**Category III** - Data source: pool inclusion information from pool publication - Data quality: very high - Assessment difficulty: average	**Category IV** - Data source: pool non-inclusion data reconstructed - Data quality: medium - Assessment difficulty: average

[12] These criteria were: (1) the patent owner is a member of the WCDMA pool, (2) the patent was declared to ETSI as potentially essential for the relevant standards, (3) the ETSI declaration included information on the specific standards documents for which the standard was potentially essential, (4) the ETSI declaration was within a time window when the declaring firm declared most of its patents that eventually became WCDMA pool patents, (5) the patent is not part of WCDMA pool patents nor of an INPADOC family containing other patents among WCDMA pool patents, and (6) the patent was applied for at the EPO and granted.

In total, the experiment involved 43 unique patents[13] and 48 unique standards documents.[14] Cases were randomly allocated to assessors, ensuring that an individual assessor did not receive more than one case on the same patent or on the same standard document (to prevent unobserved learning effects).

4.5 Assessors, case allocation, and instructions

Patent examiners employed by six different European patent offices participated as assessors. While the management of patent offices was closely involved in the study design and operationalisation, none of the ultimate assessors was given any of this information. In total, 20 patent examiners from six different patent offices participated, selected on the basis of their considerable expertise in the technical areas of our cases (ETSI/3GPP 3G and 4G standards). Each assessor received eight different cases, uniquely assigned to this participant and equally distributed in the combination of essential/non-essential and with or without claim charts, yielding 40 cases per combination (see Table 2). Assessors were instructed not to assume any particular distribution of cases. Since each assessment by our experts could be either consistent or inconsistent with the assessment of the pools, we assigned a value of 0/1 to each. Thus, each assessment is an independent, identically distributed observation from a Bernoulli distribution, so the final observation in each cell is of a Binomial distribution.

Table 2. Allocation of assessments

		Block		
	Claim chart	**Essential**	**Non-essential**	**Total**
Treatment	**With**	40	40	80
	Without	40	40	80
Total		80	80	160

Each case, as indicated above, refers to a combination of a granted patent document and a (specific release of a) standard document (e.g. TS 25.211 V2.5.0). Standards and patents (essential to them) have an *n:m* relationship, and also in our dataset, some cases shared a patent or standard. To avoid bias from learning, individual assessors considered every patent and every standard only once throughout the entire experiment. With the above restrictions in mind, both the allotment of the cases and the order in which the assessor processed them were randomised (and our logistics ensured they were indeed evaluated in that order).

Assessors were given an extensive set of instructions, developed together with the patent office management departments, and pre-tested for clarity (see Appendix 1). Assessors were instructed not to search for any information in addition to what we provided so that their assessment was solely based on the patent text and the standard document provided. They were not allowed to discuss cases with others for the entire duration of the experiment. Also, they were only allowed to

[13] The patents used in the cases have a filing date between 1994 and 2015 inclusive (the median filing year equals 2001 and the mode equals 2000). The EP patent publication document (EP…B1) was provided to the assessor in our experiment, except for nine patents, where the US publication document was provided.

[14] The following standards document were assigned the most (20 times or more): TS36.211, TS25.212, TS25.211, TS31.102 and TS25.223. Note that for each case the relevant version of the standards document was selected and provided to the assessor in our experiment. 27% of the cases came with a single standards document; 32% with two, 16% with three and the remaining 2% with four or five.

look up *technical* information from other sources if this was necessary to understand the technology described in the patent or standard (e.g. a technical handbook or standards document in the same 3GPP series). Patent documents were anonymised by removing patent number and assignee information, and assessors were instructed not to look up information on the specific patent (e.g. by searching under the title). For additional verification, assessors were asked to indicate if they recognized the identity of the patent owner and/or the patent. Finally, after the experiment was completed, all assessors received a feedback and debriefing form.

4.6 Data verification

Before carrying out the data analyses, we verified the assessment data for factors that could potentially have a confounding effect on the experiment. During debriefings, we realised that one group of participants had not respected all the elements in the instructions. Despite their good intentions, they did not realise this was at odds with our research design, and we had to exclude the associated observations from our quantitative analysis, but we still used their feedback in the qualitative analysis. Furthermore, in a few cases, participants reported they had seen the patent before (possibly as an examiner) and/or informed us they knew (or thought they knew) who the patent owner was. These cases were also discarded from the quantitative (but not the qualitative) analysis. Finally, studying the feedback we received from assessors, we identified 19 observations (all from data Category IV) for which assessors reported specific issues and where there might have been doubts about the cases or the reference assessment (for instance, a patent might not match the specific standards document provided even though that same set was previously provided to the pool). While we kept these cases in our dataset, we repeated the quantitative analysis excluding these cases, and this did not reveal a qualitative difference with our conclusions. Our final analysis included 109 valid observations.

5 Results

To present the results of our experiment, we address the accuracy level, then the impact of claim chart availability on this accuracy, and finally, the differences between the original and the alternative (novelty-based) definition of essentiality. For each, we discuss both the quantitative and qualitative outcomes.

5.1 Quantitative findings on overall accuracy of assessments

Table 3 shows the assessment outcomes of the experiment, compared to the reference points. In 74% of the cases, the outcome of the assessment was consistent with the reference. This percentage is higher for (according to the pools) essential patents (83%) than for non-essential patents (62%), indicating some difference between the false positives and the false negatives. Note that here we did not yet distinguish differences across assessments with or without a claim chart, nor across the different essentiality definitions.

Table 3: Discrimination between essential and non-essential patents

	Experiment outcome		
Essentiality status according to the reference point	**"Essential"**	**"Non-essential"**	**Total**
Essential	53 (83%)	11 (17%)	64 (100%)
Non-essential	17 (38%)	28 (62%)	45 (100%)
Total	70 (64%)	39 (36%)	109 (100%)

Note: Cells show the number of observations and percentage of row total.

The first test we perform is whether *participants can differentiate essential from non-essential patents (i.e. consistent with the reference).* To do so, we compare the assessments with a (hypothetical) set of random assessments with a probability of an assessment as 'essential', for all patents in the sample, equal to the overall share of assessments as 'essential'. The chi-squared test of proportions tests whether the two variables in the contingency table (in our case, the outcome of the experiment and whether the case was essential or not according to the reference point) are independent. In this case, it indicates that the assessors are significantly better than random in differentiating essential from non-essential patents ($\chi^2 = 23.32, p = 1.37 \cdot 10^{-6}$). We show below that the share of correctly assessed patents can be increased further.

5.2 Quantitative findings on impact of claim charts and essentiality definition on accuracy

Table 4 shows the result of the assessments depending on the availability of claim charts (52 of the 109 observations included a claim chart). The percentage of assessments inconsistent with the reference was twice as large without claim charts (33%) than with claim charts (17%).[15] The chi-squared test of proportions shows that availability of claim charts improves the outcome of the assessment significantly ($\chi^2 = 3.66, p = 0.056$). Importantly, providing claim charts increases the share of consistent assessments to 83%. However, claim charts do not seem to affect differently the essential and non-essential cases ($\chi^2 = 0.35, p = 0.34$).

Table 4: The effect of the availability of claim charts

	Experiment outcome compared to reference point		
Claim chart availability	**Consistent**	**Inconsistent**	**Total**
No claim chart	38 (67%)	19 (33%)	57 (100%)
Claim chart	43 (83%)	9 (17%)	52 (100%)
Total	81 (74%)	28 (26%)	109 (100%)

Note: Cells show the number of observations and percentage of row total

As mentioned in Section 4.1, most of the assessors applied the ETSI definition of essentiality, while others followed what we called a novelty-based test. So far, the results shown included the data points from both. To check this was indeed valid, we verified whether the results of the 'ETSI-based'

[15] Note again, that our non-essential cases with claim chart might have been the most difficult to assess, since patent holders would not create a claim chart if they did not believe the patent was essential. Thus, for the overall population of patents that might be candidates for an assessment procedure, the difference might be bigger than Table 4 indicates.

essentiality assessments and those based on the novelty-based tests are comparable.[16] Table 5 shows the results.

Table 5: Novelty-based vs. regular essentiality definition

	Experiment outcome compared to reference point		
Type of essentiality definition	**Consistent**	**Inconsistent**	**Total**
Regular	59 (73%)	22 (27%)	81 (100%)
Novelty-based	22 (79%)	6 (21%)	28 (100%)
Total	81 (74%)	28 (26%)	109 (100%)

Note: Cells show the number of observations and percentage of row total

A chi-squared test of proportions does not indicate a significant difference between the distributions in the rows of Table 5. ($\chi^2 = 0.358, p = 0.55$). If anything, we see that the outcome of novelty-based assessments appears slightly more often consistent with the reference (79%) than the regular assessments (73%). This result has an important implication. Even though most assessors felt qualified to perform the assessments (expressed in their open feedback), patent examiners are not always trained to perform infringement analyses, and infringement partly depends on the respective national law. However, since patent examiners are trained to perform novelty analyses, they can directly perform essentiality assessments under the novelty-based definition. This is relevant given that stakeholders expressed their confidence in the reputation of patent offices as trustworthy, independent third parties that were a qualified candidate to perform these tests on a large scale.

Finally, we looked at the combined effects of claim chart availability and different essentiality definitions. Table 6 shows that claim charts seem to improve the degree of consistency of regular assessments considerably (from 63% to 84%), while consistency of novelty-based assessments is unchanged (at 79%). However, the chi-squared test of proportions shows that the differences are not significant ($\chi^2 = 1.34, p = 0.25$).

Table 6: Separated results for claim chart availability and novelty-based assessments

		Experiment outcome compared to reference point		
Claim chart provided	**Type of assessment**	**Consistent**	**Inconsistent**	**Total**
No	Regular	27 (63%)	16 (37%)	43 (100%)
No	Novelty-based	11 (79%)	3 (21%)	14 (100%)
Yes	Regular	32 (84%)	6 (16%)	38 (100%)
Yes	Novelty-based	11 (79%)	3 (21%)	14 (100%)
All		**81 (74%)**	**28 (26%)**	**109 (100%)**

Note: Cells show the number of observations and percentage of row total

5.3 False positives vs. false negatives

An intriguing aspect of our experiment is the comparison between false positives and false negatives. As discussed above, the assessment accuracy might differ between presumed essential and presumed non-essential patents. We therefore tested whether *there is a difference between false positives (assessors claim essential, reference is not essential) and false negatives (expert claim*

[16] Note that we ran this analysis *before* the others, but only discuss this analysis now for readability reasons.

not essential, reference is essential). The contingency table for this analysis holds the same information as table Table 2, only reorganized so that it shows, in the rows, whether the case is essential according to the reference point, and in the columns, whether it was evaluated consistently to the reference point. The percentage of non-essential patents assessed as essential (inconsistent with the reference point) is 38%, more than twice the percentage of essential patents assessed as non-essential (17%). We compared this result with an expected outcome that would give the same likelihood of false positives and false negatives. The chi-squared test of proportions indicates that there are indeed significant differences ($\chi^2 = 5.87, p = 0.015$).

There are two possible interpretations of this result. Assessors might be more inclined to assess a patent as essential than non-essential, as a form of confirmation bias. Or, our non-essential cases (according to the reference) might have been 'more difficult' to assess than the essential cases. The non-essential patents with claim charts are, as discussed in Section 4.4, cases for which the patent holders thought they were close enough to being essential that it was worth investing the effort and cost of creating a claim chart. Thus, some of the non-essential cases might be more difficult to assess. Relevant to both explanations is that in their qualitative feedback, assessors indicated that proving a patent's essentiality typically took less effort than proving non-essentiality, particularly when no claim charts were provided.

5.4 Qualitative findings

We also gathered extensive qualitative feedback from the assessors, both for each individual assessment they performed, and after having finalized all their work. They expressed their strong belief that the task they were given required a thorough knowledge of the standard documents. Such knowledge, they felt, could be gained in practice by specialising in essentiality assessments. For example, in one case, the patent referred to a feature that was not available at all in the specific standards document provided to the assessor but may have been elsewhere in the standard (3GPP standards together cover hundreds of separate documents). In another case, an assessor suspected that a patent would be essential for a newer release of the relevant standards document, whereas the provided version did not require the use of the patented technology. Assessors indicated that improved searching tools could help a less experienced assessor, especially where the claimed essential features were spread over a combination of standard documents. Moreover, assessors felt they would have benefited from access to additional information about the patent, such as written opinions from patent offices, claim trees, external knowledge, and interaction with stakeholders. Additionally, while the text in the standard (and, where applicable, the claim charts) we provided was always in English, like the claims in the patent publication, we had cases where the rest of the text in the patent publication was in German or French. In response, some examiners expressed that they felt less confident about their interpretation of the patent coverage. Implementing the above suggestions should help to increase the percentage of consistent assessments.

On average, assessors reported spending 7.7 hours per case – slightly less than the eight hours we communicated that they could spend on the assessment. For nine observations, participants reported spending 'much more' time than anticipated, for five observations 'much less' and for other observations only a little more or less than expected. Using the hourly rate currently used by commercial patent attorneys working for patent pools,[17] this would translate to a cost of

[17]Bekkers et al. (2020a), find pools spend between € 5000 and € 10,000 per evaluation, taking 2 to 3 days on average, i.e., average fees are between €312 and €416 per hour. The pool assessments are performed by attorneys, chosen to match the applicable jurisdiction of the patent. Their fees greatly differ between jurisdictions.

approximately € 2,400 - € 3,200 per assessment. Yet, the assessments proposed in this paper are performed by patent examiners, not attorneys, and costs/fees might be quite different.

Participants appeared relatively confident in their evaluations, labelling them as 'very certain' (25 observations) or 'quite certain' (101). In the remaining observations, participants felt 'undecided' (16), 'quite uncertain' (12) or 'very uncertain' (6). However, the level of self-reported confidence does not seem to indicate a higher accuracy of the evaluation, with respect to the reference point, as indicated by a chi-squared test of proportions between the two-level self-reported confidence and the accuracy of the assessment ($\chi^2 = 1.01, p = 0.29$). Participants generally (in qualitative feedback) indicated they felt qualified to perform the assessment, even though it was a new task, and a few cases were noted as outside the assessor's regular field of expertise (at a lower-level technical 'layer'). In 137 cases, participants reported that their skill level increased 'slightly' or more, which may indicate the presence of a learning effect. See (Bekkers et al, 2020a) which elaborates on the learning effect observed in the earlier experiment.

Participants indicated (on a closed question form) that the claim chart was 'very helpful' (31 observations) or even 'extremely helpful' (28 observations).[18] In their open feedback, assessors indicated that claim charts were useful for two reasons: they saved time and made them feel more confident about the outcome. They also commented that in the absence of claim charts, the procedure of reading the patent description, isolating the parts truly reflected in the claims, and then doing the same for the standard document and matching both parts, were a lot more effort.

6 Conclusion and discussion

The purpose of this study was to investigate whether essentiality assessments can be made sufficiently efficient (in terms of time and costs) as well as sufficiently accurate, to set up a large-scale system of essentiality assessment, and thus overcome important inefficiencies in the market for SEP licensing.

In our experiment, where assessors were instructed to spend an average of eight hours on each case, we found that 74% of the outcomes are consistent with the essentiality assessments of patent pools (which we used as reference point). When the assessors were provided with claim charts as additional input and used the 'regular' essentiality definition, consistency increased to 84%. Given our choice to study the organisationally most feasibly, but in terms of accuracy most challenging assessment scenario (see Section 4.2), we believe these results are encouraging, and can be considered as a lower bound for other designs which are likely to have much higher accuracy scores.

Further improvements should be achievable when implementing our approach in practice. Firstly, we introduced several limitations to ensure a proper research design. Participating assessors were not allowed to work in teams or exchange information, nor look up patent prosecution history. Apart from the information provided by us, the only other source they could look up was general technical background information (such as the usual meaning of a specific term. In a practical implementation of the approach, such limitations would not be imposed, thus likely increasing performance. Second, even though the assessors were selected on expertise with the relevant ETSI/3GPP 3G and 4G standards, this is still a relatively broad area. In practice, a larger assessment team would include specialists in relevant subfields (switching, radio protocols, etc.), and patents could be allocated to assessors according to their key technological competences, thus improving performance. Thirdly, our data set included a significant number of relatively 'difficult' cases, such as patents previously submitted to a pool but subsequently rejected. Patent pool assessments are

[18] These numbers include qualitative feedback on cases that we had to exclude from the quantitative analysis.

costly, and we must assume that the patent owner had reasonable expectations that the patent was essential. By contrast, in a large-scale implementation of our approach, assessments may start from *all* patents declared as potentially essential to SDOs, implying that there will be many more obvious (i.e. 'easy' to assess) cases. Again, performance should increase. Fourth, we see many learning opportunities, both for individuals (progressing experience and knowledge) and in a team setting. Indeed, the participating assessors commented that the eight cases were not enough to generate learning effects. Finally, a practically implemented system could allow the patent owner involved as well as third parties (such as implementers) to appeal the result of the assessment. This measure should improve the accuracy of the assessments, though we recognize such procedures need to be designed carefully to avoid potential misuse.

Given the outcome of our experiment and the above opportunities to improve performance, we believe that a large-scale system of essentiality assessments based on our approach can achieve a good degree of accuracy at a comparatively low cost (compared to the currently known processes with high accuracy), and overcome important inefficiencies in the market for SEP licensing.

We note that a large-scale assessment system, as studied in this paper, is not the only way of creating more transparency in terms of essentiality. An alternative approach would be a two-stage disclosure procedure within SDOs, where companies at an early stage disclose patents that are potentially essential, and later, once the standard is frozen and the potentially essential patent is granted, companies re-assess their patent and make an additional disclosure whether or not they believe the patent actually to be essential. While such two-stage procedures have been proposed by some (Qualcomm, 2006), no SDO has reached any consensus to include such an obligation in its IPR policy and given the convenance structure of SDOs (see Baron et al., 2019), it is unlikely that this will happen any time soon. Moreover, such a two-stage procedure will still be a self-declaration by the patent owner, not an assessment by an independent, impartial third party that applies consistent standards.

Our experiment has several limitations. First, in our assessment, we used patents declared at ETSI as cases and the ETSI essentiality definition. Performing such assessments using other SDOs' definitions of essentiality could be more challenging, especially if these definitions are based, for instance, on "commercial" rather than technical essentiality. Second, while ETSI requires parties to disclose which specific patents they believe to be potentially essential, other SDOs, including ITU, IEEE, and ISO/IEC, allow parties to submit 'blanket' declarations that do not indicate specific patents. While a large-scale essentiality test mechanism does not necessarily rely on declarations made at SDOs (it may also start by patent owners proposing their patents for assessment), this may limit system design options. Third, the availability of input claim charts, where we observed the highest degree of consistency, will depend on the willingness of patent owners to make such information available – and, in turn, on the incentives patent owners see to do so. As indicated above, such questions concerning 'institutional feasibility' are beyond the scope of this paper but are addressed in complementary work presented in (Bekkers et al., 2020a). Fourth, differing patent prosecution procedures, examination details, linguistic translation variations, third party interventions and substantive legal rules across jurisdictions may result in differences in the scope of issued patent claims from one jurisdiction to another, even for patents in the same family and originating from the same international (Patent Cooperation Treaty) application. Determining, even with a high degree of certainty, that a particular member of a patent family (e.g., a European patent) is essential to a given standard may only provide approximate information about patents issued in different countries (e.g., in the U.S., China or Japan). (Note that EPO patents, however, are identical across all countries for which the patent owner chose for national validation.) Finally, while we believe that patent pool assessments are an appropriate reference point for this study, they do not represent an absolute reference point, and as mentioned above, such a reference point does not exist. Therefore, our findings are necessarily limited to observing consistency, not accuracy.

There are ample opportunities for future research in this area, especially since the European Commission announced in its November 2020 IP Action Plan that it will explore the creation of an independent system of third-party essentiality checks with a view to improving legal certainty and reducing litigation (European Commission, 2020:13). One of these opportunities is research on understanding how AI-based systems, while not replacing human assessors (see Section 3.3), can complement human assessments.

References

Akerlof, G., 1970. The market for ‘lemons’: Quality uncertainty and the market mechanism. Quarterly Journal of Economics 84(3), 488-500.

Aoki, R., Arai, Y. (2018). Strategic Declaration of Standard Essential Patents. RIETI Discussion Paper Series 18-E-035. Retrieved from https://www.rieti.go.jp/jp/publications/dp/18e035.pdf

Arora, A., Ceccagnoli, M., 2006. Patent protection, complementary assets, and firms’ incentives for technology licensing. Management Science 52 (2), 293–308.

Arora, A., Fosfuri, A., & Gambardella, A. (2001a). Markets for technology and their implications for corporate strategy. Industrial and corporate change, 10(2), 419-451.

Arora, A., Fosfuri, A., Gambardella, A., 2001b. Markets for Technology: The Economics of Innovation and Corporate Strategy. MIT Press, Cambridge, MA.

Article One Partners. 2011. LTE Standard Essential Patents now and in the future. Retrieved January 30, 2019, from http://newsletters.articleonepartners.com/news_f1317eac-ee13-5a66-d0f5-38ea99a4c1eeLTE-Standard-Essential-Patents-Now-and-in-the-Future.pdf

Arts, S., Cassiman, B. & Gomez, J. 2017. Text matching to measure patent similarity. Strategic Management Journal 39(1), 62-84.

Baron, J., Contreras, J., Husovec, M., Larouche, P., 2019. Making the Rules: The Governance of Standard Development Organizations and their Policies on Intellectual Property Rights, Thumm, N. editor, EUR 29655 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-00022-8, doi:10.2760/207268, JRC115004.

Baumol, W., 2010. The Microtheory of Innovative Entrepreneurship. Princeton University Press, Princeton, NJ.

Beggs, A., 1992. The licensing of patents under asymmetric information. International Journal of Industrial Organization 10(2), 171-191.

Bekkers, R., Catalini, C., Martinelli, A., Righi, C., Simcoe, T., 2017. Disclosure rules and declared essential patents. NBER Working Paper 23627, available at http://www.nber.org/papers/w23627. National Bureau of Economic Research, Cambridge (MA).

Bekkers, R., Raiteri, E., Martinelli, A., Tur, E., 2020 Landscape Study of Potentially Essential Patents Disclosed to ETSI. Nikolaus Thumm (ed.). Publications Office of the European Union, Luxembourg, JRC121411, doi:10.2760/313626, ISBN 978-92-76-20786-3.

Bekkers, R., Henkel, J., Tur, E. M., van der Vorst, T., Driesse, M., Kang, B., Martinelli, A., Maas, W., Nijhof, B., Raiteri, E., Teubner, L., 2020a. Pilot study for essentiality assessment of Standard Essential

Patents. Thumm, N. (ed.), EUR 30111 EN, Publications Office of the European Union, Luxembourg, JRC119894.

Bekkers, R., Updegrove, A. 2013. A study of IPR policies and practices of a representative group of Standards Setting Organizations Worldwide. Washington: National Academies of Science. Retrieved from https://www.nap.edu/resource/18510/Bekkers-Updegrove%20Paper_092013.pdf

Bekkers, R., West, J. 2009. The limits to IPR standardization policies as evidenced by strategic patenting in UMTS. Telecommunications Policy, 33(1-2), 80-97.

Bessen, J., Meurer, M., 2008. Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk. Princeton University Press, Princeton, PA.

Brachtendorf, L., Gaessler, F., Harhoff, D. 2020. Truly standard-essential patents? A semantics-based analysis. CEPR Discussion Paper No. DP14726. https://ssrn.com/abstract=3603956.

Bresnahan, T., Trajtenberg, M., 1995. General purpose technologies 'Engines of growth'? Journal of econometrics, 65(1), 83-108.

Caves, R., Crookell, H., Killing, J., 1983. The imperfect market for technology licenses. Oxford Bulletin of Economics and Statistics, 45(3), 249-267.

Charles River Associates (CRA). 2016. Transparency, predictability, and efficiency of SSO-based standardization and SEP licensing. A Report for the European Commission. Retrieved from https://ec.europa.eu/growth/content/study-transparency-predictability-and-efficiency-sso-based-standardization-and-sep-0_en

Contreras, J., 2013. Fixing FRAND: A Pseudo-Pool Approach to Standards-Based Patent Licensing, Antitrust Law Journal, 79(1), 47-97.

Contreras, J., 2016. When A Stranger Calls: Standards Outsiders and Unencumbered Patents, Journal of Competition Law and Economics, 12(3), 507-540.

Contreras, J. (ed). 2007. Standards Development Patent Policy Manual. American Bar Association (ABA), Committee on Technical Standardization, Section of Science & Technology Law. Chicago: ABA Publishing.

Contreras, J., 2017. Essentiality and Standards-Essential Patents, In J. Contreras (Ed.), The Cambridge Handbook of Technical Standardization Law: Competition, Antitrust, and Patents (pp. 209-230). New York: Cambridge University Press. doi:10.1017/9781316416723.016

Contreras, J., 2019. Technical Standards, Standards-Setting Organizations and Intellectual Property: A Survey of the Literature (With an Emphasis on Empirical Approaches), in Menell, P. and Schwartz, D. (eds.), Research Handbook on the Economics of Intellectual Property Law, Vol 2 - Analytical Methods (185-235). Cheltenham and Northampton: Edward Elgar.

Cyber Creative Institute. 2011. Evaluation of LTE essential patents declared to ETSI. Retrieved January 30, 2019, from http://www.cybersoken.com/file/lte01EN.pdf.

Den Uijl, S., Bekkers, R., de Vries, H. (2013). Managing Intellectual Property Using Patent Pools: Lessons from Three Generations of Pools in the Optical Disc Industry. California Management Review, 55(4), 31-50. doi:10.1525/cmr.2013.55.4.31

Dewatripont, M., Legros, P. 2013. ‘Essential’ patents, FRAND royalties and technological standards. Journal of Industrial Economics 61(4), 913-937.

European Commission (2017) Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee: Setting out the EU approach to Standard Essential Patents. COM(2017) 712. https://ec.europa.eu/docsroom/documents/26583.

European Commission (2020). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Making the most of the EU’s innovative potential: An intellectual property action plan to support the EU’s recovery and resilience. Brussels, 25.11.2020 COM(2020) 760 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0760

European Telecom. Standards Inst. (ETSI), 2020. Rules of Procedure of the European Telecommunications Standards Institute, Version approved by the extraordinary SCM of the General Assembly on 3 September 2020. ETSI’ Sofia Antipolis.

Fairfield Resources International. 2007. Analysis of Patents Declared as Essential to GSM as of June 6, 2007. Retrieved January 30, 2019, from http://www.frlicense.com/recent.html

Fairfield Resources International. 2009. Review of Patents Declared as Essential to WCDMA through December, 2008. Retrieved January 30, 2019, from http://www.frlicense.com/recent.html

Fairfield Resources International. 2010. Review of patents declared as essential to LTE and SAE (4G Wireless Standards) through June 30, 2009. Retrieved January 30, 2019, from http://www.frlicense.com/recent.html

Farrell, J., & Shapiro, C. 2008) How strong are weak patents? American Economic Review, 98(4), 1347–1369.

Gallini, N., Wright, B., 1990. Technology transfer under asymmetric information. RAND Journal of Economics 21(1), 147-160.

Gambardella, A., Giuri, P., Luzzi, A., 2007. The market for patents in Europe. Research Policy 36, 1163–1183.

Gans, J., Hsu, D., Stern, S., 2008. The impact of uncertain intellectual property rights on the market for ideas: Evidence from Patent Grant Delays. Management Science 54 (5) 982-997.

Goldstein, L., and Kearsey, B. 2004. Technology Patent Licensing. Aspatore Books.

Goodman, D., Myers, R., 2005. 3GPP cellular standards and patents. IEEE WirelessCom 2005.

Graham, S., Mowery, D., 2003. Intellectual property protection in the U.S. software industry. In: Cohen, W., Merrill, S. (eds.), Patents in the Knowledge-Based Economy. National Academies Press, Washington, DC, 219–258.

Henkel, J. 2021. How to license SEPs to promote innovation and entrepreneurship in the IoT. Working paper, TUM School of Management. https://ssrn.com/abstract=3808987.

Henkel, J., Zischka, H. 2019. How many patents are truly valid? Extent, causes, and remedies for latent patent invalidity. European Journal of Law and Economics, 48(2), 195-239.

IPLytics (2021). AI may be the solution to skyrocketing numbers of SEP declarations, available at https://www.iplytics.com/report/ai-solution-skyrocketing-numbers-sep-declarations/

iRunway. 2012. Patent & landscape analysis of 4G-LTE technology. Retrieved January 30, 2019, from https://www.i-runway.com/images/pdf/iRunway - Patent & Landscape Analysis of 4G-LTE.pdf

Jaffe, A., & Lerner, J. (2006). Innovation and its discontents. NBER Innovation Policy and the Economy, 6(1), 27–65.

Jefferies & Company. 2011. Research in Motion (RIMM) Limited patent value; cut target to salvage value of $21 as we wait for QNX. Retrieved January 30, 2019, from https://ipcloseup.files.wordpress.com/2011/10/rimm.pdf

Lamoreaux, N., Sokoloff, K., 1999. Inventive Activity and the Market for Technology in the United States, 1840–1920. NBER Working Paper 7107. National Bureau of Economics Research, Inc., Cambridge, MA.

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. doi:10.1016/j.ijindorg.2010.08.006

Lemley, M., Shapiro, C., 2007. Patent holdup and royalty stacking. Texas Law Review 85, 1991–2048.

Lemley, M., Simcoe, T. 2019. How Essential Are Standard-Essential Patents? Cornell Law Review 104(3): 607-642.

Lerner, J. (1994). The importance of patent scope: An empirical analysis. RAND J. Econ. 25 (2), 319–333.

Macho-Stadler, I., Martinez-Giralt, X., Perez-Castrillo, J., 1996. The role of information in licensing contract design. Research Policy 25(1), 43-57.

Madiès, T., Guellec, D., Prager, J. (eds.), 2014. Patent Markets in the Global Knowledge Economy: Theory, Empirics and Public Policy Implications. Cambridge University Press.

Magliocca, G., 2007. Blackberries and barnyards: patent trolls and the perils of innovation. Notre Dame Law Review 82 (5), 1809–1838.

Marco, A., Sarnoff, J., de Grazia, C. (2019). Patent claims and patent scope. Research Policy 48(9). https://doi.org/10.1016/j.respol.2019.04.014

Martin, D., De Meyer, C. 2006. Patent counting, a misleading index of patent value: A critique of Goodman & Myers and its uses. Available at SSRN: https://ssrn.com/abstract=949439

Merges, R., Mattioli, M. (2017). Measuring the Costs and Benefits of Patent Pools. Ohio State Law Journal 78(2), 282-347.

Milgrom, P., Roberts, J., 1992. Economics, Organization and Management. Englewood Cliffs, NJ: Prentice-Hall.

Miller, S. (2013). Where's the innovation? An analysis of the quantity and qualities of anticipated and obvious patents. Virginia Journal of Law and Technology, 18(01), 1–58.

Natterer, M. 2016. Ähnlichkeit von Patenten: Entwicklung, empirische Validierung und ökonomische Anwendung eines textbasierten Ähnlichkeitsmaßes. Verlag für Nationalökonomie, Management und Politikberatung.

PA Consulting Group. 2015. LTE Essential IPR. PA's 3GPP-LTE Database and Report.

Reitzig, M., Henkel, J., Heath, C., 2007. On sharks, trolls, and their patent prey – unrealistic damage awards and firms' strategies of 'being infringed'. Research Policy 36 (1), 134–154.

Serrano, C., 2010. The dynamics of the transfer and renewal of patents. RAND Journal of Economics 41 (4), 686–708.

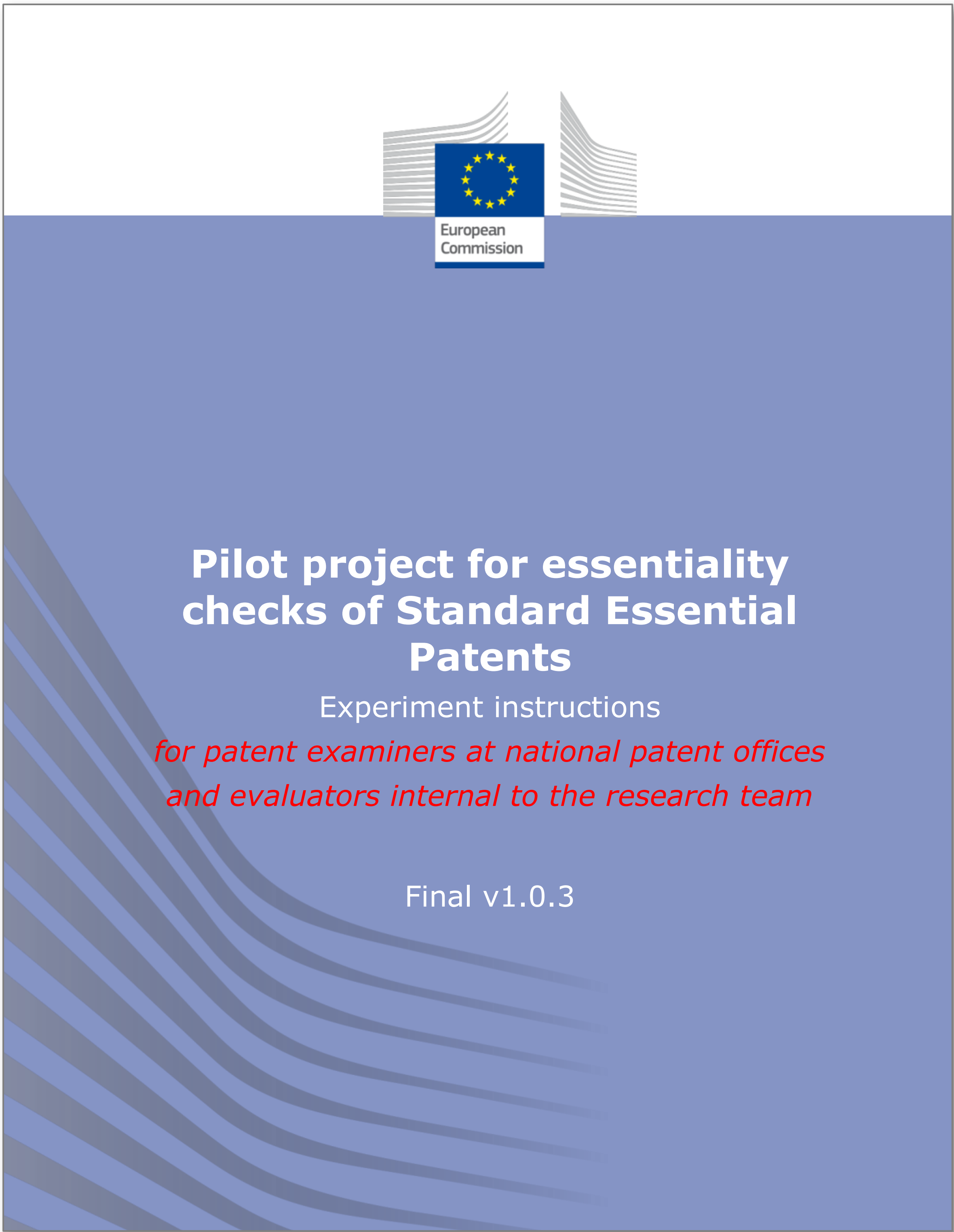
Siebrasse, N., Cotter, T. (2017). Judicially determined FRAND royalties. In Contreras, J. (ed.), The Cambridge Handbook of Technical Standardization Law: Competition, Antitrust, and Patents, 365-388. Cambridge University Press. https://doi.org/10.1017/9781316416723.026

Younge, K., Kuhn, J.. 2016. Patent-to-patent similarity: A vector space model. Social Science Research Network. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2709238

Qualcomm, 2006. Proposal for Defining Articulated Actual Problems and Solutions (DAAPS), document submitted to ETSI as Document ETSI GA/IPRR04(06)13, dated 10 May 2006.

Appendix 1: Instructions for assessors

The instructions below were provided to assessors conducting regular (ETSI-definition based) essentiality tests.


European Commission

Pilot project for essentiality checks of Standard Essential Patents

Experiment instructions

for patent examiners at national patent offices and evaluators internal to the research team

Final v1.0.3

Table of Contents

1 Introduction

1.1 About the study

The European Commission Joint Research Centre (JRC) in cooperation with DG GROW has requested our research consortium to perform a feasibility study on essentiality checking mechanisms for standard-essential patents (SEPs). The objective of the project is to assess the feasibility of essentiality testing mechanisms both a technical as well as institutional point of view. The study may prepares proposals for how to develop such mechanisms. To this end, the study will consist of an analysis of various existing essentiality testing mechanisms, among which those employed by patent pools, the Japanese patent office (Hantei advisory opinion) and in legal cases. From these insights, various possible mechanisms are defined, whose technical merits are tested in an experiment.

In this experiment, a large number of evaluators (among which are engineers, attorneys, and patent examiners) are asked to perform essentiality tests on various example cases (based on real-world patents and essentiality information). This experiment allows the researchers to find out which mechanisms are suitable, and which are not, and how they are influenced by various parameters.

1.2 About this document

This document is intended for patent offices who are participating in the experiment described above and have agreed to have a number of their patent examiners carry out essentiality tests.

The document is written as a guide for the evaluators as well as the coordinators at each patent office.

1.3 Rules of the game

Confidentiality

For methodological reasons, it is crucial to adhere to the following confidentiality rules (some of which will be formalized through NDAs):

- For methodological reasons, it is crucial that the contents of this document shall remain confidential between the European Commission/JRC, the research team, and the involved persons from the patent office.
- Evaluators shall not share any materials provided to them, nor the outcomes of their evaluations, with others. Evaluators shall remove all data provided to them directly after finishing an evaluation.
- Evaluators shall not discuss the exercise with others (within or outside the PO) before the final deadline of the evaluation period (31/7). Any questions should be directed to the research team directly.

Identifiability

This exercise is *not* a race. We will not be comparing evaluators or POs, nor sharing results identifiable to any evaluator or office. The exercise is about learning about the mechanisms, not about who is the best evaluator or PO, nor about the quality of certain SEPs portfolios.

Therefore:

- The research team will not share outcomes in a way that is deducible to particular organisations, evaluators or patents with anyone outside the research team. Any outcomes that are shared will be anonymous with respect to individual organisations, evaluators and patents.
- SEP owners in particular will *not* get any feedback on the evaluation results for cases referencing their patents.
- Evaluators shall not attempt to find out more information about the cases provided (e.g. by looking up the patent or similar patents, looking up patent pool information, et cetera) for evaluation during the evaluation period.

Methodological considerations

- Instructions should be closely followed.
- Evaluators cannot re-assign work to other evaluators or POs. An evaluator should stick to the patent set specifically assigned to him/her by the researchers. Should there be any reason to re-assign work, please contact the researchers.
- Evaluators should perform the tasks alone and not accept any interference by others.

1.4 Planning and logistics

The experiment is planned to take place between June 6 and July 31, 2019. Evaluators

This means that the researchers will provide the necessary (final) instructions as well as any data required from June 6 onwards. Evaluators can perform the essentiality testing tasks at any time before the deadline of July 31, before which we expect the evaluators to have finished all cases assigned to them.

Around June 6, we will send each evaluator a link by e-mail, pointing to the first case to be evaluated. After completion of an evaluation, the evaluator will receive a new link for the next case. This link can be saved/bookmarked.

In case an evaluator experiences difficulties, forgets a link, et cetera: please contact the researchers.

Technical requirements

The cases for evaluation will be provided through an online form. Any documents are linked from the form as PDF documents, possibly in a ZIP file. Evaluators fill in their evaluation through the online form.

Evaluators need to have a PDF reader (Google Chrome will do), an unzip tool such as WinZip or 7-Zip, internet access, and a recent web browser.

1.5 Point of contact

- In case of any question or issue related to the experiment, please contact Tommy van der Vorst (vandervorst@dialogic.nl, +31302150593) or Rudi Bekkers (r.n.a.bekkers@tue.nl).
- Please do not contact anyone else, not even fellow evaluators, in case of an issues or question.

2 Instructions for evaluators

2.1 Objectives

The European Commission Joint Research Centre (JRC) in cooperation with DG GROW has requested our research consortium to perform a feasibility study on essentiality checking mechanisms for standard-essential patents (SEPs). The objective of the project is to assess the feasibility of essentiality testing mechanisms both from a technical as well as from an institutional point of view. To this end, the study will consist of an analysis of various existing essentiality testing mechanisms, among which those employed by patent pools, the Japanese patent office (Hantei advisory opinion) and in legal cases. From these insights, various possible mechanisms are defined, whose technical merits are tested in an experiment.

In this experiment, a large number of evaluators (among which are engineers, attorneys, and patent examiners) are asked to perform essentiality tests on various example cases (based on real-world patents and essentiality information). This experiment allows the researchers to find out which mechanisms are suitable, and which are not, and how they are influenced by various parameters.

This exercise is *not* a race. We will not be comparing evaluators or POs, nor sharing results identifiable to any evaluator or office. The exercise is about learning about the mechanisms, not about who is the best evaluator or PO, nor about the quality of certain SEPs portfolios. Please provide honest answers over the course of the experiment.

2.2 Definitions

For the purposes of this exercise, the following definitions are used.

A patent is *essential* with respect to a particular standard if it is not possible to comply with the standard without infringing that patent.[1]

More specifically:

- Essentiality should take into account normal technical practice and the state of the art generally available *at the time of standardization.*
- A patent is essential even if it would only be infringed when implementing *optional* features of the standard.
- The costs of alternative (non-infringing) implementations should not be taken into account when deciding on essentiality (i.e. ‘commercial essentiality’ is not considered here).

There may be a very specific situation where the following applies:

- If an unpatented alternative is available that complies with the standard, then the patent cannot be essential. If there exist alternatives, but all are patented, then they are all considered essential.

[1] This definition is inspired by the definition adopted by ETSI.

Note that this exercise does not take into account the validity of a patent, nor the enforceability of a patent (e.g. whether the patent has expired, has been declared invalid by a court, et cetera).[2]

A 'linkage' is defined as a specific combination of (1) a single patent claim, and (2) one or more paragraphs, figures and/or other elements in the standard document, related to that claim, positively determining *essentiality* of the patent with respect to the standard. A claim chart lists one or more linkage candidate(s).

2.3 Evaluation procedure

Who will perform the evaluation?

- Evaluators are assigned cases by the researchers in a specific order.
- Evaluators cannot share or redistribute the cases assigned to them. In case reassignment is necessary, evaluators shall contact the researchers.
- Evaluators shall not discuss the evaluations with anyone.

How should an evaluator perform the evaluation procedure?

For each case to be evaluated, you as an evaluator:

1. Obtain the input documents (patent, standard, possibly a claim chart)
2. Consider the essentiality definition provided (below).
3. Take a look at the score sheet (below).
4. Given the definition, provide an assessment whether the evaluator believes the patent to be essential to the standard.
5. Fill in the answers and an evaluation in the form.

You will be provided with a link to an online form for the first case by e-mail. After completion of the first case the system will provide you with a link to subsequent cases.

Note: due to the experimental methodology, it is possible that all of the cases provided to you are essential patents, none of the patents are essential, or anything in between.

Which documents should the evaluator use?

- Evaluators shall base their evaluation of essentiality on the documents provided (per case this is one or more standards documents, one anonymized patent grant publication, and possibly a claim chart document).

- Evaluators shall not use other versions of the documents provided.

- Evaluators may look up *technical* information from other sources (technical handbooks, academic papers, web sites, et cetera) in order to aid their understanding of the technology described in the patent and/or the standard.

- Evaluators may look up other 3GPP standards documents, but only if they are part of the same 3GPP release.

[2] Nor are we considering whether a specific standards-based product infringes the patent. After all, such a specific product may only implement a part of the standard (e.g. only the part for mobile terminals) or might not implement all optional features.

- Evaluators shall NOT look up (or do internet searches that may result in) information related to the *patent* or the *patent owner* (evaluators should not perform a web search on the patent title or text, for instance, but can look up technical terminology).

Which parts of the patent provided should be considered?

If the provided patent includes text in multiple languages, then the evaluator shall only consider the English texts.

If no claim chart was provided:

- The evaluator may consider any element of the standard document provided.

If a claim chart was provided:

- The evaluator verifies essentiality by verifying only the candidate linkages between patent and standard text described in the claim chart (i.e. only looking at the parts of the standard and the patent that are referenced). The evaluator shall not consider any candidate linkages other than the ones in the claim chart.

- Where necessary, the evaluator *may* consult other (not referenced) parts of the provided standard document (e.g. to check whether the composition of features is actually as intended in the standard).

Which parts of the patent provided should be considered?

If no claim chart was provided:

- The determination of essentiality shall exclusively be based on the claims.
- Other parts of the patent may be used for understanding the claims.

If a claim chart was provided:

- The evaluator verifies essentiality by verifying only the linkages between patent and standard text described in the claim chart (i.e. only looking at the parts of the patent that are referenced). The evaluator shall not consider any linkages other than the ones in the claim chart.
- Other parts of the patent may be used for understanding the claims.

An example claim chart template is provided at the end of this document.

When should an evaluator consider the patent essential?

The patent shall be considered essential by the evaluator when the evaluator is confident in (at least) one of the linkages. After confirming one linkage, the evaluator does not need to consider other candidate linkages.

If no claim chart was provided:

- The evaluator will attempt to construct a linkage. For the convenience of the evaluator, a template is provided in Annex 1: Empty claim chart template).

If a claim chart was provided:

- The linkage shall be one of the candidate linkages from the claim chart.

How much time should an evaluator spend?

- The evaluator shall decide on the amount of time spent per patent (i.e. will continue until the evaluator is sufficiently confident).

2.4 Example evaluations

To illustrate more clearly what kind of evaluation we are looking for in the experiment, we provide example evaluations based on a fictitious standard. First, the example standard is defined as follows:

> *Standard TS99.888: "A UE shall include function A, function B and function C. Optionally, a mobile terminal may include function D. A base station shall include function A, function B and function E.*

Standard case of essentiality

Below is an annotated claim chart for an example patent A. The colors indicate how the patent relates to the standard (note that 'real' cases in the experiment may or may not have claim charts, and the level of detail of the claim charts provided may differ between cases).

Patent A	**Standard TS99.888**
Claim 1: A mobile telecommunications device comprising functions A, B and C	A UE shall include function A, function B and function C.

Essentiality assessment: the standard requires that a UE includes A, B and C, and thus it is not possible to make a UE conforming to the standard without necessarily infringing patent A. Hence, patent A is essential to the standard.

Patent-side scoping

Patent B	**Standard TS99.888**
Claim 1: A base station device comprising functions A, B and C	A UE shall include function A, function B and function C.

Essentiality assessment: the standard requires that a UE includes A, B and C. Yet, the scope of the patent only covers *base stations* that comprising functions A, B and C. Hence, the patent is not necessarily infringed by implementing (this part of) the standard. Hence, patent B is ***not*** essential to the standard.

3 Feedback form

Note: do not fill in the form below. Evaluators will receive a link through which they can access their cases and will be provided with an online version of this form, that should be filled in.

Score page

The following documents are provided (these are examples): instruction leaflet, patent, standard document, empty claim chart template.

1 Is this patent essential with respect to the standard provided?

- ☐ Yes
- ☐ No

2 How confident are you in your evaluation of essentiality?

	very uncertain	quite uncertain	Undecided	quite certain	very certain
I am:	○	○	○	○	○

Please proceed to the explanation page.

* mandatory question

Explanation of the essentiality

The questions in this block should only be answered if the question *"Is this patent essential with respect to the standard provided? "* has been answered with "Yes" .

3 Specify the number of the claim that is essential.

4 Specify at least one associated relevant section in the standard document.

5 Specify the device categories (if any) for which the patent would be essential for implementation of the mobile communication standard.

- Terminals (e.g. UE)
- Base station (e.g. BS / NodeB / eNB)
- Core network element (e.g. RNC, CN, SGSN, GGSN, SAE-GW, EPC)
- Other (e.g. SIM, eSIM)

6 Would you consider this patent ONLY to cover an optional feature(s)?

- Yes
- No
- Not sure

7 How much time did you spend evaluating (including filling in this page)?

Choose...

Choose...

8 Have you seen this patent before and/or did you recognize a specific applicant?

- Yes
- No

9 Did you encounter any issues evaluating this patent (i.e. did you have to assume anything, was information missing)?

10 Please proceed to the feedback page.

* mandatory question

Feedback page

Please report the feedback honestly, we are not comparing nor sharing individual performance data.

11 Is anything missing (in terms of data, knowledge, tools, training, et cetera) to perform a proper evaluation of essentiality? What additional resources would improve the ability to perform a proper evaluation?

12 Do you have any recommendations to the researchers and/or the European Commission with respect to systematic essentiality testing?

13 Did you spend more or less time than expected?

	Much less	A little less	Not more, not less	A little more	Much more
I spent:	○	○	○	○	○

14 Please elaborate on the time you spent.

15 Do you feel you became more skilled in evaluating essentiality over the course of the different cases?

	Not at all	Slightly	Moderately	Considerably	A great deal
My skills improved:	○	○	○	○	○

16 How helpful did you find the provided claim chart?

[This question is only visible if a claim chart was provided]

Not at all helpful	Slightly helpful	Moderately helpful	Very helpful	Extremely helpful

The claim chart was: ○ ○ ○ ○ ○

17 Do you feel qualified to perform essentiality testing in the way done it is requested in this experiment? Does the task align well with your regular activities, experience in the technical domain, experience in standard-setting processes, et cetera? Which qualifications do you think are required?

18 Do you have any further comments or suggestions with respect to this experiment, the project, your experience in it and/or essentiality testing in general?

The answers provided will be shown in future cases, so they can be updated at any time.

* mandatory question

Delete documents page

Please delete all documents now.

You have finalized the evaluation procedure for this case. You are now obliged to delete all case related documents and notes.

☐ [illegible]

Thank you. Please proceed to submit this case.

Annex 1: Empty claim chart template

Claim number	Standard document and version	Relevant section(s) in the standard document
Example:		
Claim 1	TS99.888 V9.3.1	§4.3.1, §4.3.2, Figure 2-1

Colorized text mapping

Claim text	Standard document text
Example:	
Claim 1	TS99.888 V9.3.1
Claim 1: A mobile telecommunications device comprising functions A, B and C	A UE shall include function A, function B and function C.

Instructions for assessors using the novelty-based test

Assessors at the patent office who decided to use the novelty-based test instead of the regular essentiality assessment received the same set of instructions, except Section 2.2, which reads as follows:

2.2 Definitions

For the purposes of this experiment, we define 'novelty-based essentiality test' as follows:

> *Evaluation of whether the patent meets the novelty requirement in the (imaginary) hypothetical situation where the relevant standard document already would have been in the public domain before the filing date of the patent.*

Note that this exercise does not take into account the validity of a patent, nor the enforceability of a patent (e.g. whether the patent has expired, has been declared invalid by a court, et cetera).[1]

A 'linkage' is defined as a specific combination of (1) a single patent claim, and (2) one or more paragraphs, figures and/or other elements in the standard document, related to that claim, positively determining *essentiality* of the patent with respect to the standard. A claim chart lists one or more linkage candidate(s).