E

Committee on Development and Intellectual Property (CDIP)

Twenty-Sixth Session
Geneva, July 26 to 30, 2021

REPORT ON WOMEN AND IP: COMPILATION AND SHARING OF DATA

Prepared by the Secretariat

1.   The Committee on Development and Intellectual Property (CDIP) at its twenty-second session, held from November 19 to 23, 2018, while discussing the topic "Women and IP" under the agenda item IP and Development, took the following decision:

> *"The Committee considered the document CDIP/22/16 Rev. 2 and adopted a proposal by Mexico on "Women and IP" as contained in the Annex to this document. The Annex will be forwarded to the WIPO General Assembly of 2019."* (para. 9 of the Summary by the Chair)

2.   The mentioned Proposal by Mexico, *inter alia*, requested the Secretariat to undertake a series of actions and decided *"to revisit the issue of "Women and IP" at its 26th session for the way forward".*

3.   This document, together with document CDIP/26/8, seeks to facilitate the Committee's revisiting of the issue of "Women and IP" at the present session. While this document responds to requests 1 and 2 of the Proposal by Mexico related to the compilation and sharing of sex-disaggregated data in the field of IP, document CDIP/26/8 reports on the actions taken by the Secretariat to implement requests 3, 4 and 5. The reporting period covers the years 2019 and 2020.

I. SECRETARIAT'S IMPLEMENTATION OF PROPOSAL BY MEXICO

4.   In the following, a description of the Secretariat's actions in the implementation of requests 1 and 2 of the Proposal by Mexico is provided.

Request 1: *Compile comparable and disaggregated international data on the gender of IP rights owners and creators.*

5.   International patent applications do not include information on the gender of the applicant. However, this information is valuable for the IP community, for example, to detect and understand potential innovation gaps. Therefore, since CDIP/22, the Secretariat has continued to work to develop methodologies to further understand the intersection of gender equality and IP and, based on these, to compile comparable international sex-disaggregated data. This work has been channeled, in particular, through the Department of Economics and Data Analytics (DEDA) in the IP and Innovation Ecosystems Sector (IPIES). In particular:

6.   The Statistics and Data Analytics Division published PCT statistics disaggregated by sex in its main flagship publications and web data center during the same period. Namely:

6.1. PCT yearly review (2019)

6.2. PCT yearly review (2020)

6.3. World IP Indicators (2018)

6.4. World IP Indicators (2019)

6.5. World IP Indicators (2020)

6.6. IP Statistics Data Center (regular updates during the period)

7.   The Innovation Economy Section continued to perform methodological and economic research on gender equality and IP. The research during the reporting period has had two main axes: (i) further improve gender disambiguation methodologies; and, (ii) improve and expand the IP collections and data with sex disaggregation.

8.   During the reporting period, the Innovation Economy Section further improved gender disambiguation methodologies by:

8.1. Creating more accessible versions of the World Gender-Name Dictionary (WGND) by publishing it in the Harvard Dataverse (WGND 1.0) and in GitHub (WGND 1.1). See more on this in paragraph 10 below.

8.2. Improving the international coverage of the WGND by including more sources from new countries (WGND 2.0). This new version is published on WIPO website, Harvard Dataverse (WGND 2.0), and in GitHub (WGND 2.1).

8.3. Analyzing the robustness of using gender dictionaries for given names from certain Asian languages. This analysis indicates the problems of Romanized Chinese and Korean given names. The main implications from the analysis are that future gender analysis will benefit from: (i) increasing the coverage of IP data in original Asian characters; and (ii) extending the coverage of Asian names in original characters in future versions of the WGND.

8.4. Analyzing alternative methods for gender disambiguation, such as machine learning imputation. A preliminary exploration was done in collaboration with the Advanced Technology Applications Center (ATAC) in the Infrastructure and Platforms Sector.

8.5. Conceiving a gender parity forecast series of methods. These methods allow to assess different scenarios of when gender parity can be attained. The different forecast scenarios can be applied to any new gender disambiguated data. It has been applied regularly by DEDA to PCT inventors and to the list of WIPO General Assemblies delegates upon request of WIPO's Gender and Diversity Specialist.

9. In the same period, the Innovation Economy Section performed research on improving, expanding, and analyzing the sex disaggregated IP data by:

9.1. Automating the PCT inventors' gender disambiguation using the WGND. A method was conceived, designed and fully implemented to be applied to the historical and new PCT unit-record data based on WGND 1.0.

9.2. Automating the PCT inventors' gender disambiguation for non-Romanized inventor's names based on complementary original Asian character data. A method was conceived, designed and fully implemented to apply the WGND 1.0. The method requires additional data containing inventors' names in original Asian characters which are provided regularly by the Operations and Support Section in the PCT Information Systems Division in the Patents and Technology Sector.

9.3. Expanding the gender disambiguation to individual applicants and legal representatives in the PCT collection. A method was conceived, designed, and fully implemented to be applied to the historical and new PCT unit-record data based on WGND 1.0. The results from this method are being further analyzed.

9.4. Expanding the gender disambiguation to patent and utility model inventors in national IP collections. A method was conceived, designed, and experimentally implemented to be applied to the historical data available in DOCDB/PATSTAT unit-record data using WGND 1.0. Preliminary results of this research were published in the World IP Indicators series.

9.5. Analyzing the potential of gender disambiguation of Hague industrial design unit-record data. Data was collected and analyzed for the Hague system applying the WGND 1.0. As a preliminary result from this and the following analyses, it was concluded to expand the international coverage of WGND sources (see paragraph 8.2). The same data were reanalyzed using WGND 2.0. The further analysis of the Hague System unit-record data indicated two main constraints: (i) the entry of new Hague members at different moments of time generate large fluctuations in the sample; and, (ii) several Hague member states do not require the declaration of creators in their design law. These two constraints limit the interpretation of global gender equality trends, including several national ones. During the analysis, it was decided that further work is required before publishing the results.

9.6. Analyzing the potential of gender disambiguation of Madrid trademark unit-record data. As for Hague, data was collected and analyzed for the Madrid system applying the WGND 1.0 and 2.0. The analysis of the Madrid System unit-record data indicated one main limitation: trademark data only captures information from applicants, which can be physical and non-physical persons. Only the gender of physical persons (i.e. individual trademark applicants) can be disambiguated using gender dictionaries. The gender analysis of individual applicants offers an incomplete characterization of gender parity of the Madrid System. However, it may be a potential characterization of smaller entities

using this system, such as entrepreneurs, micro or small enterprises. This latter interpretation requires further analysis before publishing the results.

Request 2: *Share methods and procedures for the collection of sex-disaggregated data, the use of indicators, the monitoring and evaluation methodologies, and the economic analysis of gender gaps related to IP.*

10. During the same period, the Secretariat has undertaken several initiatives to share the analyses and methodologies of women and IP-related topics. In particular, the Innovation Economy Section explored several ways to share publicly the first version of the gender disambiguation methodology (based on WGND 1.0), such as:

10.1. Direct contact with academic networks, UN entities (such as UN Women and UNESCO), and the interested public from private companies.

10.2. Using the Harvard Dataverse Repository, which is a free data repository open to all researchers from any discipline, both inside and outside of the Harvard community (Dataverse Repository).

10.3. Using a GitHub Repository to share the WGND data with a wider audience and build tools upon (GitHub Repository).

11. From the above sharing experience, the Secretariat notes that these channels work quite well for a technical audience, but have some limits to reach a more general one. In particular, the direct sharing channel had an impact in several IP offices (such as the US Patent and Trademark Office and the Canadian IP Office, among others) and scholars, but most of these were based in developed countries. The experience of the digital sharing suggests that the dissemination is much wider, with more than 1,400 downloads on WIPO's website and 1,353 downloads in the Harvard Dataverse.[1]

12. In order to support a wider and faster diffusion of the gender methodologies and analyses, the Innovation Economy Section has prepared a new *Guidelines* draft document focusing on how to measure gender in IP. These guidelines share the basics of IP data gender disambiguation and related gender indicators.

13. In complement to these guidelines, the Innovation Economy Section created a series of dedicated web contents on "Innovation, Creativity and Gender" explaining the Secretariat's economic and methodological research to a more general audience.[2] The preliminary feedback indicates that this could be used not only to promote WIPO's work on the topic but also to highlight work produced by other national IP offices – such as those from the US, UK, Canada or Chile – and other partners, such as UN Women or UNESCO. The targeted audience is, among others, government officials from Member States and international government agencies, worldwide scholars and students, and anyone interested in the topic.

---

[1] To date, it is too soon to assess the GitHub engagements.
[2] At the time of drafting the current document, the contents were in the process of being added to WIPO's website.

II. THE WAY FORWARD

14.   In the Secretariat's experience during the reporting period, there are a series of concrete steps to be taken into consideration.

15.   First, the Secretariat needs to further ease the access to gender-related methodological content to a less technical audience. This will permit the diffusion of the methodologies to a wider set of Member States and generate local capacity to produce analyses that are gender-responsive and sex-disaggregated. These dissemination efforts will include the publication of research papers and web articles on "Innovation, Creativity and Gender", and the organization of tailored capacity building workshops and information sharing sessions. The latter within the framework of the sharing sessions promoted by the CDIP decision on "Women in IP".

16.   Second, the Secretariat will continue to produce research on methodologies extending and analyzing comparable and disaggregated international data on the gender of IP rights owners and creators. In particular, those aiming at extending the analysis to other forms of IP, such as utility models, industrial designs, or trademarks.

17.   *The Committee is invited to take note of the information contained in this document.*

[End of document]