Issues Note – Measuring the Digital Gender Divide through Official Statistics

DETF workshop, 18 February, Session 3

Introduction and issues for discussion

This issues note is intended as a guide for discussion during the DETF Workshop of 18 February (a draft agenda is available separately), which will include a session on the measurement of the digital gender divide. In 2018, the G20 endorsed a report on Bridging the Digital Gender Divide (OECD, 2018) and in 2020, a number of indicators related to gender were included in the report "A Roadmap Toward a Common Framework for Measuring the Digital Economy" (OECD, 2020). These reports mirror recent developments and approaches that provide hard evidence in support of policies aimed to narrow the gender divide, and in particular the digital gender divide. They point to a limited presence and role of women in many areas of the digital economy, including in ICT-related activities such as technological innovation, software development or entrepreneurship. In addition, the 2019 report on "The Role of Education and Skills in Bridging the Digital Gender Divide" (OECD, 2019a), highlights the need for women to acquire a number of skills needed to thrive in the digital economy. Despite the important advances made, gaps in the measurement of the digital gender divide nevertheless remain. These need to be addressed to inform the design and implementation of coordinated and effective gender-related policies.

The G20 has put a strong emphasis on gender divides in its work and the digital transformation provides "*new avenues for the economic empowerment of women and can thereby contribute to greater gender equality*" (OECD, 2018). However, for the digital transformation to hold its promises of greater inclusion, policy makers need to be able to better apprehend, monitor and steer the change. Designing and implementing policies able to narrow the digital gender divide should rely on solid data and clear targets. Measurement is key to help identify and address the causes and the many facets that the digital divide may take, including access, education and skills (or lack thereof), as well as time and safety concerns, especially those related to online violence.[1]

A gender breakdown is typically available in many statistics produced by National Statistical Offices (NSOs) in relation to the digital economy, such as those based on surveys of households of individuals, labour force surveys, statistics on educational achievement and, in a growing number of cases, employer-employee linked data and administrative sources. Other indicators, e.g. on women's role in innovation and entrepreneurship may draw on different types of administrative or private sources, and may not always be central to the work of NSOs. In what follows, the note first focuses on the statistics generally produced by National Statistical Offices and the way they can help substantiate the policy discussion on gender equality. It then briefly highlights the additional information that can be obtained from other administrative or private sources to inform policy discussions on the digital gender divide, areas where NSOs may consider increasing use and/or developing more guidance and/or statistics.

At the DETF workshop, delegates may wish to consider the following issues:

- ***Building on ongoing work across G20 countries, what additional indicators on the digital gender divide should receive priority? Should any priorities and/or targets be established?***
- ***How can the international comparability of indicators on the digital gender divide be improved? What steps could be taken?***
- ***What role can administrative and private sources play in measuring the digital gender divide? What role should NSOs play in coordinating, steering and advancing the development and production of statistics and indicators based on such additional sources?***
- ***How can the G20 best support the development of internationally comparable indicators on the digital gender divide? How can the G20 support statistics capacity building?***

Gender in official statistics on the digital economy

Existing gender-disaggregated statistics

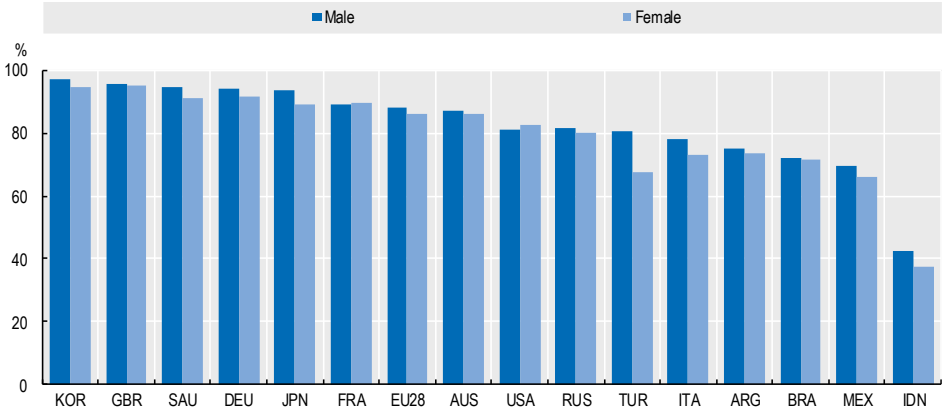
Several official statistics related to the digital economy also feature a gender dimension. This is the case, for example of indicators on ICT usage by households and individuals. The report "A Roadmap toward a Common Framework for Measuring the Digital Economy" (OECD, 2020) included a number of gender-related indicators drawing on official statistics, notably ICT

[1] OECD (2019a), The role of Education and Skills in Bridging the Digital Gender Divide – Evidence from APEC Economies, OECD, Paris.

usage surveys and labour force surveys.[2] Indicators by gender on individuals using the Internet and on ICT skills can be derived from ICT usage surveys for individuals and households (Figure 1; Figure 3), while indicators about ICT professionals and technicians by gender can be extracted from labour force surveys (Figure 2). These two surveys are rich sources of information to explore the digital gender divide and can be used to point to many gender-related differences in the use of digital technologies. For example, the ICT usage survey shows that more women use the Internet to look for health information than men (OECD, 2019b).[3]

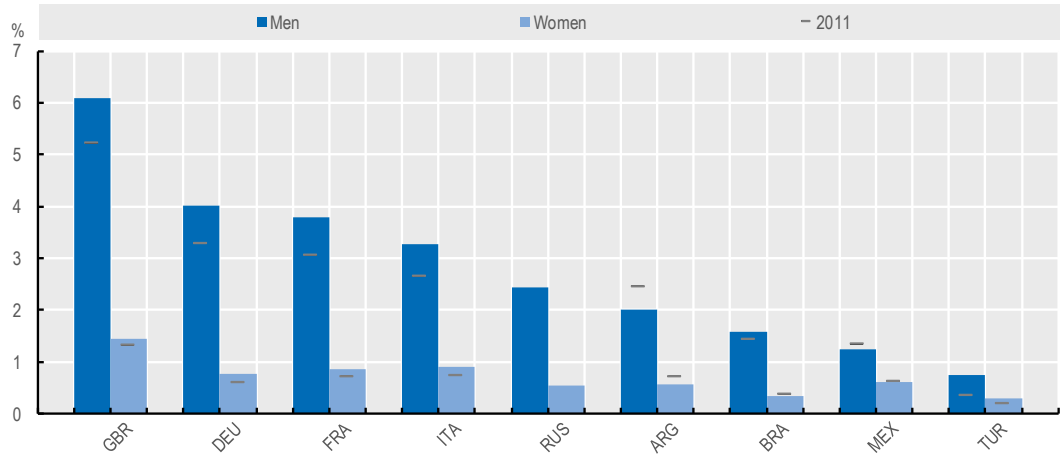
Education and skills-related statistics provide another important source of gender-disaggregated data. For example, the 2020 report included an indicator on the ICT task intensity by gender (Figure 3.2.1), drawing on the OECD's PIAAC survey; an indicator on students' reported ICT capabilities, by gender (Figure 3.3.2), drawing on the OECD's PISA survey; and an indicator on tertiary graduates in Natural Sciences and Engineering & ICT, by gender (Figure 4), drawing on education statistics. Linked employer-employee data provide another good source of information and can help point to women's role in firms and the economy more widely.

Figure 1. Individuals using the Internet, by gender, G20 economies, 2019
As a percentage of all men and women aged 16-74, respectively



Note: Unless otherwise stated, Internet users are defined as individuals who accessed the Internet within the 3 months prior to being surveyed. For Argentina, Indonesia, Japan, and Saudi Arabia the recall period is not specified in the ITU file. For Australia, data refer to individuals "who accessed the Internet for personal use in a typical week". For Brazil, Indonesia, Japan, Russian Federation, Mexico, and Saudi Arabia data relate to 2018 instead of 2019. For Argentina, Australia, and the United States data relate to 2017 instead of 2019.
Source: OECD, ICT Access and Usage by Households and Individuals Database, http://oe.cd/hhind and ITU, World Telecommunication/ICT indicators Database, February 2020.

Figure 2. ICT professionals and technicians by gender, 2018
As a percentage of all male and female workers

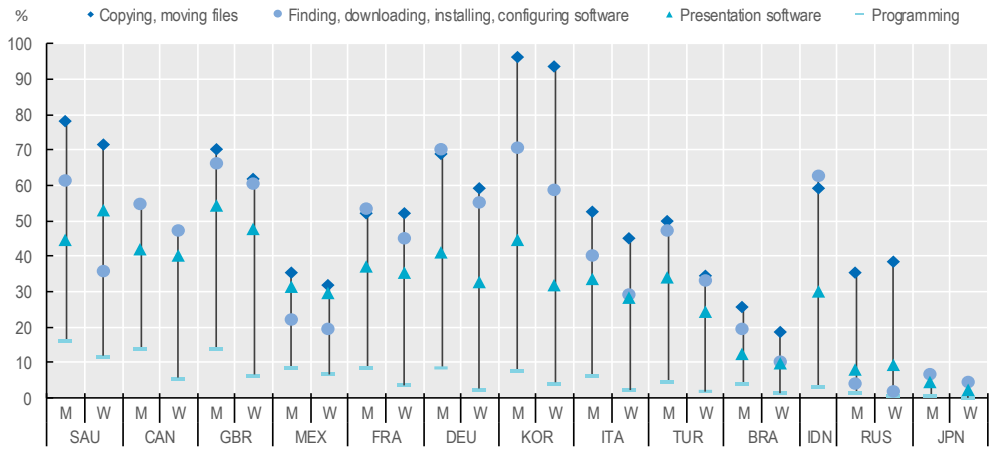


Source: International Labour Organization (ILO) estimates based on Australian, European, Korean and South African labour force surveys, Census of Japan 2015, the United States Current Population Survey, alongside ILO data.

[2] Some countries administer the ICT survey as a module *within* the Labour Force Survey, while others administer it as part of a broader multipurpose survey or as a stand-alone survey, which is deemed the best option in terms of quality and response rates.
[3] OECD (2019b), Measuring the Digital Transformation – A Roadmap for the Future, OECD, Paris.

Figure 3. Selected ICT skills, by gender, 2018

As a percentage of individuals; M = men, W = women

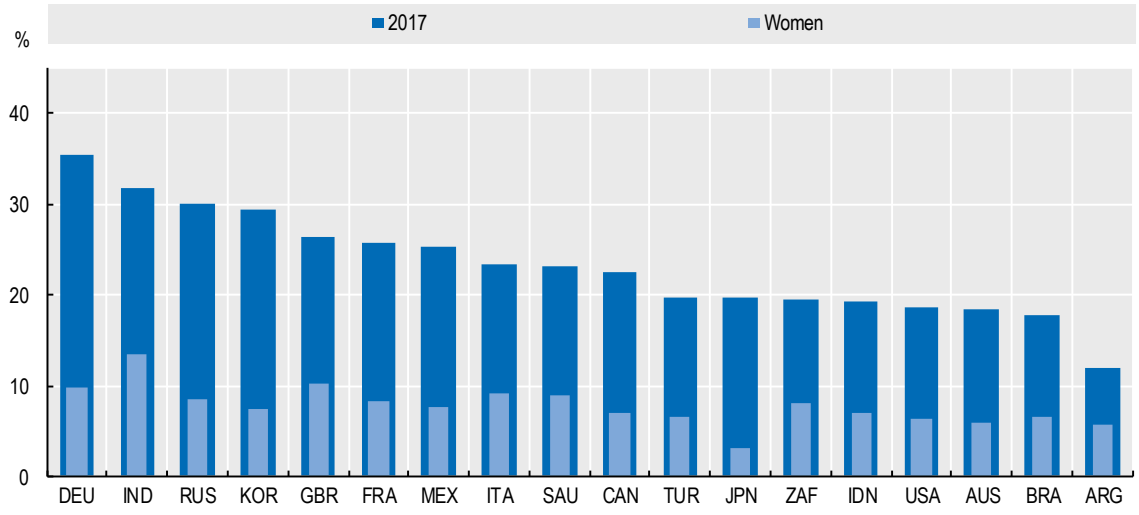


Notes: Germany and Turkey relate to 2019 rather than 2018. France, Indonesia, and United Kingdom to 2017, Italy to 2016. For France, the gender breakdown for copying and moving files is unavailable so the overall share is used and other series are sourced from the OECD database, which may use slightly different definitions. Data for Canada, as well as Turkey software downloads also from OECD. Population in scope varies by country: age 5+ for Indonesia; 6+ for Italy, Japan and Mexico; age 10+ for Brazil and Germany; 10-74 years for Saudi Arabia; 15-72 years for Russian Federation and 16-74 years for France, Korea, Turkey and United Kingdom.

Source: ITU World Telecommunication/ICT Indicators database and OECD ICT Access and Usage by individuals database, March 2020

Figure 4. Tertiary graduates in NSE & ICT, by gender, 2017

As a percentage of individuals graduating at the tertiary level



Source: OECD calculations based on OECD Education database, March 2020.

Table 1 provides an overview of the indicators included in the 2020 report. In general, it is recommended that gender and other key personal characteristics are duly considered when further developing data sources and indicators and that data collections are designed to deliver breakdowns along these lines. Also, to monitor progress over time, it is important to identify priority statistics to be collected or developed from a gender perspective on a recurrent basis, rather than in an ad-hoc fashion.

Table 1: Indicators on Jobs, Skills, and Growth in the Digital Economy with a Gender Dimension included in the 2020 Roadmap report

Section	Indicator name	Data source(s)	Underpinning data source
Digital economy (Chapter 1)	**3.2.2a.** Individuals using the Internet, by gender	ITU World Telecommunication/ICT Indicators database and OECD ICT Access and Usage by individuals database	Household and Individuals ICT usage surveys / modules in LFS
	3.2.2b. Patenting activity by women inventors	Patent statistics	Worldwide Statistical Patent Database (PATSTAT)
Jobs (Chapter 3)	**2.2.2** ICT professionals and technicians by gender	International Labour Organization (ILO) based on national Labour Force Surveys	LFS
Further indicators for development	**3.1.1** Selected ICT skills by gender	ITU World Telecommunication/ICT Indicators database and OECD ICT Access and Usage by individuals database	Household and Individuals ICT usage surveys / modules in LFS
	3.2.1 ICT task intensity of jobs by gender	PIAAC database	PIAAC skills survey module
Skills (Chapter 3)	**3.3.2** Students' reported ICT capabilities, by gender	OECD PISA Database	PISA assessments
	3.4.2 Tertiary graduates in NSE & ICT, by gender	OECD Education Database	Administrative registers and/or survey sources

How to step up collection and use of gender-disaggregated statistics

Previous work for the G20 already included a number of recommendations on how to improve digital economy statistics, with several of these also relevant for improved measurement of the digital gender divide. For example, the 2018 G20 Toolkit for Measuring the Digital Economy (OECD, 2018) lists the following actions for G20 members to make statistical systems more flexible and responsive to the new and rapidly evolving digital era:

i. experiment with concepts and data gathering within existing measurement frameworks,
ii. exploit the potential of existing survey and administrative data,
iii. add questions to existing surveys,
iv. periodically augment existing surveys with topic-specific modules,
v. develop short turnaround surveys to meet specific needs,
vi. define policy needs and, in cooperation with other stakeholders, set priorities for internationally comparable measurement; and
vii. work with stakeholders, including IOs, to harness the potential of big data for developing indicators

The 2020 Roadmap provided the following recommendations to improve statistics on the digital economy and capitalise upon the work set out in the roadmap:

1. Sharing of experiences and best practices between G20 countries.
2. Fully utilising and working to improve existing statistical infrastructures. A specific example are ICT usage surveys that can look in detail at the impacts of digital transformation at work, on well-being, and on key groups within society.
3. Developing alternative (non-survey) sources of data, including private sector data sources, and data gathered from the Internet to measure certain indicators describing more qualitative phenomena – including on skills, inclusiveness, and business activities.
4. Developing a multi stakeholder approach, enabling dialogue between businesses, government and other actors from civil society to strengthen the evidence base and complement current statistics to develop new interdisciplinary approaches to data collection.
5. Further developing public-private partnerships, including with owners of Internet-based platforms.
6. Recognising and acting upon the "*need to support statistical operations and capacity building in G20 members, [and also] in developing countries*" (*G20 Digital Economy Ministerial Declaration, 2018*).
7. Making the resulting indicators, and where possible datasets, publicly available, including through international statistical collections coordinated by IOs.

8. Regular reassessment and updating of the indicators and methods in use.

Finally, the United Nations (UN, 2016), set out the key principles for integrating a gender perspective in statistics. Key elements include (UN 2016):

- Taking a gender perspective into account from the stage of planning the data collection and setting out the objectives of the survey or census. This is important in terms of topics covered and their specific implementation, as reflected in approaches to data collection.
- Within the topics covered, identifying relevant gender issues and taking into account gender-specific conceptual and measurement issues in the questionnaire design, interviewer's manual and training of the interviewers and supervisors.
- Ensuring that the language, terms or phrasing of the questions do not induce gender biases.
- Samples used should cover all groups of population, households or economic units known to have distinct gender patterns. The sample design should also ensure that reliable statistics are produced for both women and men in sufficient detail and allow disaggregation by other characteristics as required for meaningful gender analysis.
- Addressing gender-related measurement issues and gender stereo-types in the training for interviewers and supervisors.
- Avoiding the introduction of gender bias into the data at the stage of data coding and data editing.
- Considering other gender-specific issues, depending on the type of data collection (e.g. labour force surveys or time-use surveys).

Official statistics on the digital gender divide tend to be relatively well developed for certain issues, notably ICT use by gender, and some aspects of educational achievements and skills by gender. However, official statistics are relatively less available for other key dimensions of the digital gender divide, notably its economic aspects, e.g. innovation and entrepreneurship, where the gender dimension is also of great importance. Other sources can help in these areas and are discussed below.

Other key measurements of the digital gender divide – some highlights

While surveys of individual and household use represent important sources of information in measuring the digital gender divide, other administrative and private data can provide important complementary information on the challenges related to the digital gender divide. This section briefly explores some other sources of data that have been used to measure key aspects of the digital gender divide.

For example, patent records provide a wide array of information on the invention itself, its development, the technology area(s) the invention contributes to, as well as about the ownership of the invention and who, i.e. the inventors, contributed to its development. Women inventors can be identified using country-specific gender-name dictionaries applied to the names of the inventors listed in patent documents (OECD, 2018).[4] Figure 5 shows that among G20 economies, over 2012-15, India, Mexico and Brazil reported the highest shares of patents invented by women (OECD, 2020). High proportions of patents invented by women are also observed in Korea and in China, though these figures are not fully comparable with other countries so are not shown here (OECD, 2018)[5]. In G20 economies, over the years 2012-15 female inventors appeared generally less active than men in all forms of ICT-related patents. Available indicators for scientific publications in AI by gender confirm similar patterns. Such figures need to be considered with care, as countries differ substantially in the overall number of IP5 patent families filed and the total number of publications; therefore, in absolute numbers, figures may (and) look different from those emerging when ratios are considered. Many factors may contribute to explain the figures at hand, including education and industry specific characteristics, as well as selection effects determined by culture or social norms.

OECD (2018) included a number of other indicators drawn from administrative and private sources. For example, experimental indicators using information about a popular open-source programming language for data analysis, R, showed that about three-quarters (i.e. 77%) of the 12 000 R-based software packages created during the period 2012-17 were produced by teams composed of only men. Women-only teams accounted for a mere 6% of such packages, whereas the remaining 17% came out of mixed teams of software developers.

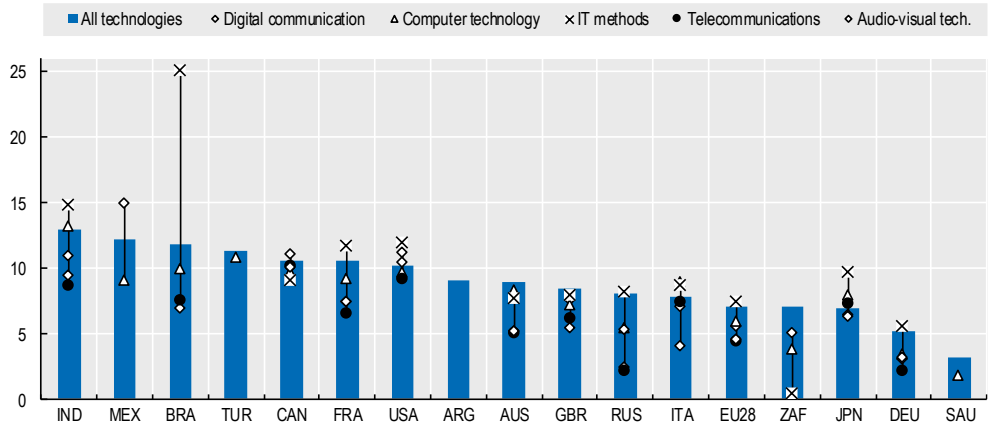
When it comes to AI skills, data from private sources such as LinkedIn enable cross-country comparisons of AI skills endowments by country, industry, gender and over time. For instance, as shown on the OECD's AI Observatory, women in

[4] See Martínez, Raffo and Saito (2016) for more detail.

[5] This is due to the difficulty to clearly attribute gender based on inventors' name, given naming traditions and use of the same name for both men and women.

India, the United States and Korea were on average more likely than women in other countries to state they had AI skills on their LinkedIn profiles in the period 2015-2020.[6]

Figure 5. Patenting activity by women inventors, 2012-15
As a percentage of IP5 patent families, by technology and inventors' country



Source: OECD based on STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2017.

Another dimension is the creation of high-technology start-ups by gender. Despite the role of entrepreneurship as an engine of social inclusion, the gender gap in entrepreneurship is striking and persistent, with men three times more likely than women to own a business with employees across OECD countries (Piacentini, 2013). For example, OECD analysis shows that among the start-ups covered by the private database Crunchbase (Lassebie, et al., 2019), less than 6% are founded by women only, while only 15% have at least one woman among the founders. Moreover, start-ups with at least one female founder are significantly less likely to receive VC funding than start-ups founded exclusively by men. In addition, conditional on receiving venture capital, the amount of funding received by startups with at least one female founder is a third less than that for male-led start-ups. Such data can be derived as the Crunchbase database covers detailed information on the socio-demographic characteristics of the founders of start-up companies.

As with the development of other indicators, some of the indicators that can be derived from administrative data (e.g. patent data) or private data may eventually be complemented by or included in official statistics. At the same time, other sources and academic and private studies are likely to remain important for the timely tracking of the digital gender divide, in particular for rapidly evolving areas like artificial intelligence. Statistical offices and governments may therefore wish to consider whether they should play a role in guiding the development of statistics and indicators based on sources other than those already managed by NSOs, e.g. by developing statistical quality frameworks or by working to harmonise indicators and enhance their comparability across country and over time.

Conclusions

This note has shown that official statistics on the digital gender divide tend to be relatively well developed for certain issues, notably ICT use by gender, and some aspects of educational achievements and skills by gender. However, only few official statistics are available for other key dimensions of the digital gender divide, notably its economic aspects, e.g. innovation, entrepreneurship or trade, all areas where the gender dimension is also of great importance. Other sources can help in these areas and some were discussed in the previous section.

A few conclusions can be drawn from this brief overview:

- Household, labour force and ICT usage surveys provide a first source of information on the digital gender divide in many G20 countries and indicators drawing on such surveys are available for most G20 countries. The design of such surveys is important, however, as noted by the UN (UN, 2016), to ensure that these sources provide the best possible and unbiased information. Moreover, further efforts may be needed to improve the international comparability of indicators being derived from such sources.

[6] See: https://www.oecd.ai/data-from-partners?selectedTab=AIJobsAndSkills

- Other official statistical sources can provide complementary information on the digital gender divide, e.g. education statistics or linked employer-employee datasets (LEEDs).
- Administrative sources are also increasingly being used by statistical offices and others to develop new insights, including at the international level.
- Private sources of information may also provide useful insights. Statistical offices may wish to consider whether they should play a role in guiding the development of statistics and indicators based on such sources, e.g. by developing statistical quality frameworks or supporting the comparability of such indicators.

As noted already in 2018, data by gender are generally insufficient to monitor and evaluate gender-related policies. Moreover, the brief overview shows that not all aspects of the gender divide are equally well covered by statistics and indicators. While many official indicators are available on access to and use of digital technologies by gender, far fewer official data are available on engagement in the digital economy by gender, including on key dimensions of innovation and entrepreneurship, but also on topics such as the involvement in trade and e-commerce by gender. Discussions on the future development of statistics on the digital gender divide may wish to consider this in setting priorities.

In order to collect data allowing for the construction of gender-related indicators, and fine-tune existing guidelines for the collection of gender and ICT-related data, countries may wish to consider complementing existing data collections and collecting data on a recurrent basis, enabling progress to be monitored. Moreover, G20 countries may wish to consider whether new international standards on measurement of the digital gender divide might be required (UNCTAD, 2014).

Finally, as in other areas, development of new indicators and statistics requires strong capabilities within statistical offices. As noted in 2018, G20 countries may wish "to support statistical operations and capacity building in G20 members, [and also] in developing countries" (G20 Digital Economy Ministerial Declaration, 2018).

References

G20 (2018), *G20 Toolkit for Measuring the Digital Economy*, https://www.oecd.org/g20/summits/buenos-aires/G20-Toolkit-for-measuring-digital-economy.pdf

Martínez, G.L., J. Raffo and K. Saito (2016), "Identifying the gender of PCT inventors", *World Intellectual Property Organization Economic Research Working Paper*, Vol. 33, No. 2016:941, October

OECD (2018), *Bridging the Digital Gender Divide*, OECD Publishing, Paris, http://www.oecd.org/digital/bridging-the-digital-gender-divide.pdf

OECD (2019a), *The Role of Education and Skills in Bridging the Digital Gender Divide*, OECD Publishing, Paris, https://www.oecd.org/sti/education-and-skills-in-bridging-the-digital-gender-divide-evidence-from-apec.pdf

OECD (2019b), *Measuring the Digital Transformation: A Roadmap for the Future*, OECD Publishing, Paris, https://doi.org/10.1787/9789264311992-en

OECD (2019c), Levelling the Playing Field – Dissecting the Gender Gap in the Funding of Start-ups, *OECD Science, Technology and Industry Policy Papers*, June 2019 No. 73, OECD Publishing, Paris, **No.** 73, https://doi.org/10.1787/7ddddd07-en

OECD (2020), *A*__Roadmap *Toward a Common Framework for Measuring the Digital Economy*, http://www.oecd.org/sti/roadmap-toward-a-common-framework-for-measuring-the-digital-economy.pdf

Piacentini, M. (2013), "Women entrepreneurs in the OECD: Key evidence and policy challenges", OECD Social, Employment and Migration Working Papers, No. 147, OECD Publishing, Paris, http://dx.doi.org/10.1787/5k43bvtkmb8v-en

UNCTAD (2014), *Measuring ICT and Gender – An Assessment*, United Nations, New York and Geneva.

United Nations (2016), *Integrating a Gender Perspective into Statistics*, Studies in Methods, Series F No. 111, Statistics Division, Department of Economic and Social Affairs, United Nations, New York, https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/gender/Integrating-a-Gender-Perspective-into-Statistics-E.pdf