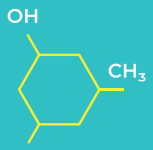
European Commission

SHE FIGURES 2021

HANDBOOK



Research and Innovation

She Figures Handbook 2021

European Commission
Directorate-General for Research and Innovation
Directorate D — People
Unit D.4 — Democracy and European Values
Contact Mina Stareva, Head of Sector D4.001 – Gender
Athanasia Moungou, Unit D4.001

Email RTD-GENDERINRESEARCH@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu

European Commission
B-1049 Brussels

Manuscript completed in September 2021

PDF ISBN 978-92-76-40596-2 doi: 10.2777/003736 KI-02-21-964-EN-N

Luxembourg: Publications Office of the European Union, 2021

EUROPEAN COMMISSION

She Figures Handbook 2021

2021

Directorate-General for Research and Innovation
Horizon 2020 Science with and for Society

CONTENTS

TABLES

ACRONYMS

ASJC	All Science Journal Classification
BES	Business Enterprise Sector
CAGR	Compound Annual Growth Rate
CI	Confidence Interval
DI	Dissimilarity Index
DG	Directorate-General
EFTA	European Free Trade Association
EGGE	Expert Group on Gender and Employment
EIGE	European Institute for Gender Equality
ENP	European Neighbourhood Policy
EPO	European Patent Office
ERA	European Research Area
ETER	European Tertiary Education Register
EU-LFS	Eurostat Labour Force Survey
EU MS	European Union Member States
FORD	Fields of Research and Development
FTE	Full-Time Equivalent
FWCI	Field-Weighted Citation Impact
GCI	Glass Ceiling Index
GOV	Government sector
HC	Head Count
HES	Higher Education Sector
HQP	Highly Qualified Personnel
HRST	Human Resources in Science and Technology
ICT	Information and Communication Technology
ILO	International Labour Organization
ILOSTAT	International Labour Organization Statistics Database
IPC	International Patent Classification (by WIPO)
ISCED	International Standard Classification of Education
ISCED-F	ISCED – Fields of Education and Training

ISCO	International Standard Classification of Occupations
JPO	Japan Patent Office
KIA	Knowledge-Intensive Activities
KIABI	Knowledge-Intensive Activities – Business Industries
MORE	Mobility and Career Paths of Researchers in Europe
NACE	Nomenclature générale des activités économiques dans les communautés européennes (Statistical Classification of economic activities in the European Community)
NPIs	Non-Profit Institutions
NPISH	Non-profit Institutions Serving Households
OECD	Organisation for Economic Co-operation and Development
PATSTAT	EPO Worldwide Patent Statistical Database
PhD	Doctor of Philosophy
PNP	Private Non-Profit
PPS	Purchasing Power Standards
R&D	Research and Development
R&I	Research and Innovation
RFOs	Research Funding Organisations
RPOs	Research Performing Organisations
S&E	Scientists and Engineers
S&T	Science and Technology
GDRIC	Gender Dimension in Research and Innovation Content
SNA	System of National Accounts
STEM	Science, Technology, Engineering and Mathematics
UIS	UNESCO Institute of Statistics
UNESCO	United Nations Educational, Scientific and Cultural Organisation
USPTO	United States Patent and Trademark Office
WIPO	World Intellectual Property Organization
WiS	Women in Science

COUNTRY CODES

Codes for aggregated data

EU-27	European Union – 27 countries (from 2020)
EU-28	European Union – 28 countries (2013-2020)

European Union Member States (and United Kingdom)

BE	Belgium
BG	Bulgaria
CZ	Czechia
DK	Denmark
DE	Germany
EE	Estonia
IE	Ireland
EL	Greece
ES	Spain
FR	France
HR	Croatia
IT	Italy
CY	Cyprus
LV	Latvia
LT	Lithuania
LU	Luxembourg
HU	Hungary
MT	Malta
NL	Netherlands
AT	Austria
PL	Poland
PT	Portugal
RO	Romania
SI	Slovenia
SK	Slovakia

FI	Finland
SE	Sweden

European Free Trade Association Countries

IS	Iceland
NO	Norway
CH	Switzerland

Candidate countries

ME	Montenegro
MK	North Macedonia
AL	Albania
RS	Serbia
TR	Turkey

Other countries

AM	Armenia
AR	Argentina
AU	Australia
BA	Bosnia and Herzegovina
BR	Brazil
CA	Canada
CN_X_HK	China (except Hong Kong)
FO	Faroe Islands
GE	Georgia
HK	Hong Kong
IL	Israel
IN	India
JP	Japan
KR	South Korea
MD	Moldova
MX	Mexico

RU	Russia
TN	Tunisia
UA	Ukraine
UK	United Kingdom
US	United States
ZA	South Africa

1. GENERAL INTRODUCTION

This Handbook of Indicators on Women in Science has been developed to accompany the She Figures publication. It contains methodological guidance on the collection of data and the calculation of all indicators in the She Figures. In doing so, it provides further guidelines and recommendations with regard to the collection, processing and use of data on gender equality in research, innovation and science, with the potential to inform organisations at both the national and European level.

She Figures

The She Figures provides pan-European, comparable statistics on the state of gender equality in research and innovation. It covers a wide range of themes, including the gender balance amongst Doctor of Philosophy (PhD) students and academic staff, the relative working conditions of male and female researchers and the steps taken by research institutions to promote gender equality internally. Released every three years since 2003, the report provides a crucial evidence base for policies in this area. It is produced in close collaboration with Member States, Associated Countries and Eurostat. It is recommended reading for policymakers, researchers and anybody with a general interest in these issues.

A large portion of the She Figures publication is dedicated to reporting back on a core set of well-established indicators, which serve as the foundation for exposing persistent gender inequalities in the fields of research and innovation (R&I). In addition, each She Figures publication builds on previous versions by introducing new indicators, which aim to bring additional and critical gender-based issues to the forefront of the science and technology debate.

This handbook serves as a resource detailing the relevant guidelines for the collection of data pertaining to all indicators of the She Figures 2021 main publication.

Upon future developments and new editions of the She Figures indicators, the handbook will be revised accordingly. As such, it is designed to reflect the state of the art in the mapping and monitoring of gender equality in science and research.

1.1. Aim and scope

Aim

This handbook aims to provide specific guidelines and recommendations concerning the necessary data and indicators for monitoring progress towards gender equality in science, research and innovation.

In particular, the handbook promotes cross-country uniformity in terms of data collection, indicator computation and data-validation procedures. Furthermore, it provides interested stakeholders with detailed information on the data needed to examine gender equality in research and innovation as well as the importance given to gender/sex issues in research content. It serves as a reference document and provides users with the methods needed to calculate the indicators, so as to increase the quality and consistency of gender-related indicators across countries and time periods.

Scope

The handbook is not intended to be specific to any version of the She Figures publications. Rather, it is intended to be used as the basis for the computation of indicators in current and future versions of She Figures and related publications.

Current version of the handbook

Although intended to act as a stand-alone document (i.e., untied to any of the specific versions of the She Figures publication), the current version of the handbook was created to accompany the 2021 edition of the publication and thus includes some details specific to that edition. In the 2021 version of She Figures, data are presented at the individual country level as well as the broader EU level for the current 27 EU Member States, plus United Kingdom, European Free Trade Association (EFTA) countries (Iceland, Norway, Switzerland), candidate countries (Albania, North Macedonia, Montenegro, Serbia and Turkey), Associated Countries (Armenia, Bosnia and Herzegovina, Faroe Islands, Georgia, Iceland, Israel, Moldova, Norway, Switzerland, Tunisia and Ukraine) and G20 countries (Argentina, Australia, Brazil, Canada, China (except Hong Kong), Hong Kong, India, Japan, Mexico, Russia, South Africa, South Korea and United States).

The handbook has been thoroughly cross-referenced and contains an indexed list of indicators aimed at improving accessibility and readability (see Annex 3).

1.2. History and background of the She Figures

History

In 1999, the Council of the EU recognised that women were under-represented in the fields of scientific and technical research, describing this as a 'common concern' at the national and European level.[1] At this time, there were virtually no pan-European statistics on what happened to women after they left university, despite fears that after graduating from their degrees, 'women frequently encounter[ed] obstacles in their career[s]', which contributed to their under-representation in scientific posts (Directorate-General (DG) Research, 2009).

Subsequently, the EU recognised the need for harmonised sex-disaggregated data on women in science and research if governments were to develop effective policies in this area.[2] Meeting in 1999, the Helsinki Group on Women and Science appointed a sub-group of Statistical Correspondents with responsibility for collecting national data and feeding into the creation of European statistics on these topics.

The end result of this process was the She Figures, first released in 2003 and updated every three years since. By presenting statistical indicators on a wide range of topics, the

[1] Council Resolution of 20 May 1999 on Women and Science (1999/C 201/01) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31999Y0716(01).

[2] For example, see European Commission (1999), 'Women and science: mobilising women to enrich European research', COM(1999)76; European Parliament Resolution on Commission's Communication COM(1999)76 (PE 284.656), 1999.

report enables readers to develop a comprehensive understanding of the state of gender equality in science, research and innovation.

Changes to the She Figures over time

Primarily, the She Figures publication serves as a tool for measuring the impact and effectiveness of gender equality policies in science, research and innovation. The majority of indicators in the She Figures present and explore the following themes:

- The presence of women in research across different sectors;
- Horizontal segregation by sex across different fields of study and occupations (in Research and Development (R&D) roles);
- Vertical segregation by sex in academia, i.e., the (under-)representation of women in the highest grades/posts of research and as heads of academic institutions.

Each edition also aims to further understanding of these issues by introducing additional indicators that explore new themes. She Figures 2006 developed new indicators to give a more detailed picture of the labour force as a whole and the patterns of employment for female and male researchers across different sectors, such as the business enterprise sector (BES). The 2009 edition introduced indicators on the gender pay gap and began to break down some data by age group (in addition to sex disaggregation). Amongst other things, the 2012 report added indicators on the mobility of researchers and the proportion of researchers with children.

Similarly, She Figures 2015 included new indicators to match emerging policy priorities. Some provide further insight into the working conditions of researchers, the proportion of Research Performing Organisations that have adopted Gender Equality Plans, while others measure the relative contribution of women and men to published research and inventorships and the degree to which researchers integrate a sex/gender analysis into their research papers. The 2018 edition, among other, developed indicators that measure the success of women and men in graduating from tertiary education as well as the research output and citation impact of women and men as authors and co-authors of scientific papers.

She Figures 2021 introduced nine new indicators. One of them measures the proportion of self-employed women within Information and Communication and Technology (ICT) and Science and Engineering (S&E) professionals. Two of the newly introduced indicators are based on the existing indicator that measures the proportion of women and men who work under precarious contracts in the higher education sector (HES), disaggregated further for family status and career stage respectively. Another indicator uses web scraping techniques to provide insight into the proportion of Research Performing Organisations (RPOs) promoting gender equality by integrating measures and actions to support it. Moreover, two new indicators measure the ratio of women to men among active and all authors respectively. Another indicator measures the average proportion of women among authors on publications that list among the author affiliations both a corporate entity and any other entity. Finally, two indicators that are introduced in this edition of She Figures assess the gender dimension in research and innovation content in Horizon 2020 projects, and the intersectional aspects in Horizon 2020 projects.

Data in the She Figures

Most of the She Figures indicators originate from Eurostat (the Statistical Office of the EU), which provides sex-disaggregated data on education, research and development, professional earnings and scientific employment. The Statistical Correspondents enrich this picture, by collecting primary data (broken down by sex) on senior academic staff, the heads of universities, funding applicants and beneficiaries and the membership of scientific and advisory boards.[3] Expansion of the She Figures since 2003 has resulted in the use of other sources (e.g., the MORE Survey on the Mobility of Researchers, the Scopus™ database) and of web-scraping techniques.

1.3. Structure of the handbook

The Handbook of Indicators on Women in Science is made up of three sections and four annexes:

- The first (current) section provides a brief overview of the aim and scope of the handbook, as well as a background to the She Figures.

- The second describes all indicators used in the She Figures publication, including definitions, rationale as well as computation method (with the necessary data, data source, formulas and any calculation specifications or comments that may be of relevance).

- The third section details the general quality plan of the She Figures publication, focusing on the methodological principles employed in the verification and validation of data.

- There are four annexes. The first synthesises recent changes to international classification standards that were taken into account. The second provides an overview of how key terms are defined. The third one provides an index of the indicators. The last one provides the correspondence of All Science Journal Classification (ASJC) categories with Fields of Research and Development.

The sections and annexes are followed by the bibliography.

[3] This primary data makes up the Women in Science (WiS) database.

2. INDICATORS

The indicators presented in this handbook have been selected from a variety of different sources on the basis that they provide important information on gender inequalities in the field of research and innovation. The development of each new version of She Figures includes several in-depth consultations with key stakeholders to determine how the landscape of data on gender equalities has changed since the previous version of the publication was created and whether the inclusion of new indicators is merited. During this process, indicators from previous versions of the publication are also reassessed to determine whether they are still relevant and to ensure that they adhere to ever increasing quality and coverage standards.

The data required to compute the majority of indicators are drawn from Eurostat databases or from the Women in Science database of data collected by the Statistical Correspondents. Other data sources have been used to develop new indicators in recent years. For example, in the last two editions of She Figures (2018 and 2021), the Scopus™ has been used to produce scientometric indicators by sex (e.g., ratio of women to men authorship of scientific papers, proportion of a country's scientific production including a gender dimension), and the EPO Worldwide Patent Statistical Database (PATSTAT) has been used to produce a technometric indicator by sex (e.g., ratio of women to men inventorship). Moreover, web-scraping techniques were used to measure indicators on institutional change for the 2021 edition.

In addition, the UNESCO (United Nations Educational, Scientific and Cultural Organisation) Institute of Statistics (UIS), the Organisation for Economic Co-operation and Development (OECD) and the International Labour Organization (ILO) were used as supplementary data sources for countries not covered by Eurostat. The following Sections (2.1 to 2.9) present the She Figures indicators by data source and subject group. Each section is introduced by a general rationale for the selection of each group of indicators – based on a content perspective – as well as a broad description of the source.

2.1. Eurostat – Education statistics

Content-based rationale

Indicators computed from Eurostat education statistics aim to investigate the level of progress and the persistent barriers that exist for women in the pursuit of postgraduate education, as well as the differences in subject choice and fields of study by gender, particularly in regard to natural science and engineering, within the context of persistent gender stereotypes and the EU's policy agenda. Indicators falling into this category include the proportion of women ISCED 8 (Doctoral or equivalent) graduates by country, the compound annual growth of ISCED 8 graduates by sex, the proportion of women ISCED 8 graduates by field of study, the distribution of ISCED 8 graduates across the broad fields of study by sex, the ratio of ISCED 6 (Bachelor's or equivalent) graduates to ISCED 6 entrants, the ratio of ISCED 8 entrants to ISCED 7 (Master's or equivalent) graduates and the ratio of ISCED 8 graduates to ISCED 8 entrants, by sex and broad field of study.

Broad overview of the source

These data can be accessed through the Education and Training Statistics database on the Eurostat website (http://ec.europa.eu/eurostat/web/education-and-training/overview), UIS data centre (http://data.uis.unesco.org) for AL, RS, BA, GE, AM, FO, ME, MD, TN and UA and OECD (http://stats.oecd.org) for IL. Data for G20 countries can be accessed either through UIS data centre or through OECD, depending on the indicator and the data availability. The data on education and training statistics are concerned with student enrolment and the 'education expectancy, funding, and the characteristics (e.g., gender and age) of graduates and educational personnel' (European Commission, 2014a). Data are collected on an annual basis, based on the academic year (i.e., 2012 refers to the academic year 2011/12) (European Commission, 2015a). 'Education, vocational training and lifelong learning' are ongoing goals towards a sustainable Europe by 2030 (European Commission, 2019). Statistics that provide such information are publicly available, regularly updated and accompanied by extensive methodological notes.

The classification of education levels is based on the International Standard Classification of Education (ISCED) system. This classification was revised in 2011 and data shown in this publication follow the updated version except when it is referred otherwise. Further details on the ISCED revision and the correspondence of old and new levels of education are provided in Annex 1.

A complete list of indicators falling into this category can be found in Annex 3 and their detailed description follows below.

2.1.1. ***Proportion of women and men among tertiary education graduates***

2.1.1.1. *Definition of indicator*

This indicator presents the proportion of women and men graduates in tertiary education to the total graduates in tertiary, broken down by country.

2.1.1.2. *Rationale*

In line with its ambition to encourage more 'research-intensive' economies, the European Commission had called for more doctoral candidates and argued that efforts must be made to 'tackle stereotyping and dismantling the barriers still faced by women in reaching the highest levels in post-graduate education and research' (European Commission, 2011). Since that time, women in the EU have made significant developments in raising their level of education. According to the She Figures 2018, the presence of women among doctoral graduates has increased both at the EU-27 and EU-28 level and at country level. This indicator sheds light on the level of progress in increasing women's representation in the top levels of education and research, considering their success in ultimately graduating from doctoral degrees, as opposed to their entry as candidates.

2.1.1.3. *Computation method*

Data needed

(F_i) Number of women graduates at the i-th education level. **Unit: Number**.

(M_i) Number of men graduates at the i-th education level. **Unit: Number**.

(T_i) Number of total graduates at the i-th education level. **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

UIS data centre (*http://data.uis.unesco.org; Tertiary graduates by level of education*)

OECD (*http://stats.oecd.org; Graduates by age*).

Computation formula

Proportion of women among graduates at the i-th education level = F_i/T_i

Proportion of men among graduates at the i-th education level = M_i/T_i

where:

i denotes the education level according to the International Standard Classification of Education (ISCED-2011).

2.1.1.4. *Specifications*

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level.

- ISCED level 6 corresponds to studies at Bachelor's or equivalent level according to the International Standard Classification of Education (ISCED-2011).
- ISCED level 7 corresponds to studies at Master's or equivalent level and ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

- ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

2.1.1.5. *Comments and critical issues*

Data before 2012 were compiled using ISCED 97. Changes in the international classification standards are provided in Annex 1.

Proportion of women and men among total graduates at ISCED levels 6 and 7 are calculated only at the aggregated EU level, not at country level.

2.1.2. ***Proportion of women and men among tertiary education students***

2.1.2.1. *Definition of indicator*

This indicator presents the proportion of women and men students in tertiary education to the total students in tertiary.

2.1.2.2. *Rationale*

Tertiary education, provided by universities and other higher education institutions seems to play an essential role in society, by fostering innovation, increasing economic development and growth and improving more generally the well-being of citizens. In the EU-27 there were 17.5 million tertiary education students in 2018, with women accounting for 53.7% of them (Eurostat, 2020a).

According to She Figures 2018, in 2016, women accounted for 54% of ISCED 6 and 7 students but their presence drops to 48% among ISCED 8 students. This indicator sheds light on the level of progress in increasing women's representation in the top levels of education and research, considering their success in ultimately graduating from doctoral degrees, as opposed to their entry as candidates.

Data needed

(F_i) Number of women students at the i-th education level. **Unit: Number**.

(M_i) Number of men students at the i-th education level. **Unit: Number**.

(T_i) Number of total students at the i-th education level. **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_enrt03)*

Computation formula

Proportion of women among students at the i-th education level = F_i/T_i

Proportion of men among students at the i-th education level = M_i/T_i

where:

i denotes the education level according to the International Standard Classification of Education (ISCED-2011).

2.1.2.3. *Specifications*

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level.

- ISCED level 6 corresponds to studies at Bachelor's or equivalent level according to the International Standard Classification of Education (ISCED-2011).
- ISCED level 7 corresponds to studies at Master's or equivalent level and ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).
- ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

2.1.2.4. *Comments and critical issues*

Proportion of women and men among total students are calculated only at the aggregated EU level, not at country level.

2.1.3. ***Compound annual growth rate (CAGR) of ISCED 8 graduates, by sex***

2.1.3.1. *Definition of indicator*

This indicator presents the compound annual growth rate (CAGR) of graduates by sex, meaning the average percentage growth each year for women and men graduates in a given period for graduates at ISCED 8 level.

2.1.3.2. *Rationale*

In 2012, the European Commission warned that 'while the proportion of women at the first two levels of tertiary education is higher than that of men, the proportion of women at PhD level is lower (DG Research and Innovation, 2012). The situation is still similar as the share of women among those studying bachelor's and master's degrees are higher than the corresponding share of men in the EU-27 in 2018. However, for people following doctoral studies, men were slightly more than women (Eurostat, 2020a). This indicator demonstrates the level of progress over time in increasing women's presence amongst those taking doctoral degrees.

2.1.3.3. *Computation method*

Data needed

(F) Number of women ISCED 8 graduates in a start and an end year. **Unit: Number**.

(M) Number of men ISCED 8 graduates in a start and an end year. **Unit: Number**.

(N) Number of years in reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data codes: educ_uoe_grad02 and educ_grad5)*

UIS data centre (*http://data.uis.unesco.org; Tertiary graduates by level of education*)

OECD (*http://stats.oecd.org; Graduates by age*).

Computation formula

The CAGR shows the yearly average rate of growth for a given period. For women and men graduates, it is respectively computed as follows:

CAGR for women graduates = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

CAGR for men graduates = $(\mathrm{M_e}/\mathrm{M_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

N denotes the number of years in the reference period (i.e., e – s);

F_s denotes the number of women graduates in the start year;

F_e denotes the number of women graduates in the end year;

M_s denotes the number of men graduates in the start year;

M_e denotes the number of men graduates in the end year.

For example, if there were 100 women graduates in 2012 and 150 in 2016, the calculation would be:

CAGR for women graduates = $(150/100)^{1/4} - 1 = 10.7\%$

2.1.3.4. *Specifications*

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

2.1.3.5. *Comments and critical issues*

Data before 2012 were compiled using ISCED 97. Changes in the international classification standards are provided in Annex 1.

2.1.4. *Proportion of women among ISCED 8 graduates, by broad field of study*

2.1.4.1. *Definition of indicator*

This indicator presents the proportion of women ISCED 8 graduates, within various fields of study.

2.1.4.2. *Rationale*

Although there is some disagreement amongst experts,[4] it is generally accepted that differences in women and men's educational pathways may have some impact on the occupations they pursue at a later stage. More specifically, the EU Council stresses that gender segregation in education leads to inequality in terms of pay, pensions, lifelong earnings, working conditions and the working environment, reinforces gender stereotypes (EU Council, 2017). By breaking down PhD graduations into different fields of study, this indicator enables more in-depth analysis of the extent of gender difference in subject choice.

2.1.4.3. *Computation method*

Data needed

(T_b) Number of ISCED 8 graduates in broad field of study *b*. **Unit: Number**.

(F_b) Number of women ISCED 8 graduates in broad field of study *b*. **Unit: Number**.

(b) Denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

Proportion of women among students or graduates = F_b/T_b

2.1.4.4. *Specifications*

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

The broad fields of study according to the ISCED – Fields of Education and Training (ISCED-F) classification are the following:

[4] The debate relates to the level and nature of the impact on educational segregation on later segregation in the labour market. For an overview of the debate, consider EGGE (2009), pp. 42–45.

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services.

2.1.4.5. *Comments and critical issues*

Data before 2012 were compiled using ISCED 97. Changes in the international classification standards are provided in Annex 1.

2.1.5. ***Distribution of ISCED 8 graduates across broad fields of study, by sex***

2.1.5.1. *Definition of indicator*

This indicator presents the distribution of ISCED 8 graduates by sex and broad field of study.

2.1.5.2. *Rationale*

As mentioned above, experts generally consider that the differences in women and men's educational pathways may have an impact on the occupations they pursue at a later stage.[5] This association between education and employment is a core part of the EU policy agenda. For instance, the European Commission promotes gender inclusive Science, Technology, Engineering and Mathematics (STEM) education and communication, i.e., encouraging more girls to take science subjects, with a view to considering a career in this area (e.g., Hypatia[6] project funded by the European Union's Horizon 2020 Framework Programme for Research and Innovation) and nurture the future women innovators (e.g., shemakes.eu[7] and equals-eu[8] projects funded under SwafS-26-2020 call: Innovators of the future: bridging the gender gap).

[5] There was, however, debate about the level and nature of this impact. For an overview of the debate, consider EGGE (2009), pp.42–45.

[6] https://cordis.europa.eu/project/id/665566

[7] https://cordis.europa.eu/project/id/101006203

[8] https://cordis.europa.eu/project/id/101006396

This indicator gives a picture of the overarching differences in women's and men's fields of study at ISCED 8 level. It is slightly different from the indicator 'Proportion of women ISCED 8 graduates by broad field of study' in that it breaks down the fields of study for women ISCED 8 graduates and men ISCED 8 graduates respectively.

2.1.5.3. *Computation method*

Data needed

(F) Number of women ISCED 8 graduates (all broad fields of study). **Unit: Number**.

(M) Number of men ISCED 8 graduates (all broad fields of study). **Unit: Number**.

(F_b) Number of women ISCED 8 graduates in broad field of study *b*. **Unit: Number**.

(M_b) Number of men ISCED 8 graduates inbroad field of study *b*. **Unit: Number**.

(b) Denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

For each sex, this indicator presents the proportion of graduates in each broad field of study, in order to show how women/men graduates (at ISCED 8) are spread out across different subjects.

For each broad field of study, the formula for this indicator is:

Distribution of women graduates across fields of study = $\left(\frac{F_b}{F}\right)$ for each field of study

Distribution of men graduates across fields of study = $\left(\frac{M_b}{M}\right)$ for each field of study

The proportions for each field are shown alongside each other, with a sum total of 100 % for each sex.

For example, suppose there are 100 women ISCED 8 graduates and, of these, 23 are in education, 16 in arts and humanities, 5 in social sciences, journalism and information, 5 in business, administration and law, 10 in natural science, mathematics and statistics, 8 in Information and Communication Technologies, 11 in engineering, manufacturing and construction, 12 in agriculture, forestry, fisheries and veterinary, and 10 in health and welfare. The proportion of women ISCED 8 graduates in each field (out of all fields) would be as follows:

Education: 23 / 100 = 23 %

Arts and Humanities: 16 / 100 = 16 %

Social sciences, journalism and information: 5 / 100 = 5 %

Business, administration and law: 5 / 100 = 5 %

Natural science, mathematics and statistics: 10 / 100 = 10 %

Information and Communication Technologies: 8 / 100 = 8 %

Engineering, manufacturing and construction: 11 / 100 = 11 %

Agriculture, forestry, fisheries and veterinary: 12 / 100 = 12 %

Health and welfare: 10 / 100 = 10 %

Sum total of 100 %.

2.1.5.4. *Specifications*

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

The broad fields of study according to the ISCED-F classification of fields of education and training are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services.

– The field 'Education' includes both teacher training and education science.

2.1.5.5. *Comments and critical issues*

Data before 2012 were compiled using ISCED 97. Changes in the international classification standards are provided in Annex 1.

2.1.6. *Proportion of women among ISCED 8 graduates by narrow field of study in natural science, ICT and engineering*

2.1.6.1. *Definition of indicator*

This indicator presents the proportion of women ISCED 8 graduates within the eight subfields of natural science, ICT and engineering, falling under the broad fields 'natural sciences, mathematics and statistics', 'Information and Communication Technologies' and 'Engineering, manufacturing and construction'.

2.1.6.2. *Rationale*

The EU recognises the existence of horizontal segregation, whereby women and men at the same level of education or employment are concentrated in different fields (full definition available in Annex 2). For example, according to the Gendered Innovations project, 'in both the United States and European Union, women are slightly underrepresented with respect to overall doctoral (ISCED 8) degrees, but substantially underrepresented with respect to S&E doctorates' (Stanford University, 'Disparities between women and men').

This indicator allows one to measure such segregation at ISCED 8, by presenting the proportion of women graduates in certain subfields. By breaking down the graduations by subfield, one can assess variations within broader fields of study.

2.1.6.3. *Computation method*

Data needed

(F_n) Number of women ISCED 8 graduates in each narrow field of study *n* in natural science, ICT and engineering. **Unit: Number**.

(T_n) Total number of ISCED 8 graduates in each narrow field of study *n* in natural science, ICT and engineering. **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

Proportion of women among graduates in each narrow field = F_n/T_n

where:

n refers to a particular narrow field of study.

2.1.6.4. *Specifications*

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

- The narrow fields of study in natural science and engineering according to the ISCED-F classification of fields of education and training are the following:
- 051 Biological and related sciences
- 052 Environment
- 053 Physical sciences
- 054 Mathematics and statistics
- 061 Information and Communication Technologies
- 071 Engineering and engineering trades
- 072 Manufacturing and processing
- 073 Architecture and construction.

2.1.6.5. *Comments and critical issues*

Data before 2012 were compiled using ISCED 97. Changes in the international classification standards are provided in Annex 1.

2.1.7. ***Compound annual growth rates (CAGR) of ISCED 8 graduates by narrow field of study in natural science, ICT and engineering, by sex***

2.1.7.1. *Definition of indicator*

This indicator presents the compound annual growth rate of the number of men and women ISCED 8 graduates within the eight subfields of natural science, ICT and engineering, falling under the broad fields 'Natural sciences, mathematics and statistics', 'Information and Communication Technologies' and 'Engineering, manufacturing and construction'.

2.1.7.2. *Rationale*

The EU recognises the persistent differences in the educational choices of women and men. In 2020, the European Commission mentions that in the European Education Area, the education and training systems should consider 'Developing a better gender sensitivity[9] in education processes and institutions […], challenging and dissolving gender stereotypes, especially those that constrain the choices of boys and girls for their field of study […] and working towards a proper gender balance in leadership positions' (European Commission, 2020b).

This indicator allows one to gauge the extent of such segregation at ISCED 8 level, by calculating the changes in women and men's representation over time. By breaking down the graduations by subfield, one can assess variations within broader fields of study. Please note that the results of this indicator can be compared with those of the similar

[9] According to the EIGE, gender sensitive policies and programmes are those that take into account the particularities pertaining to the lives of both women and men, while aiming to eliminate inequalities and promote gender equality, including an equal distribution of resources, therefore addressing and taking into account the gender dimension

indicator showing the proportion of women ISCED 8 graduates by narrow field of study in natural science, ICT and engineering.

2.1.7.3. *Computation method*

Data needed

(F_n) Number of women ISCED 8 graduates in each narrow field of study *n* in a start and an end year. **Unit: Number**.

(M_n) Number of men ISCED 8 graduates in each narrow field of study *n* in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

The compound annual growth rate (CAGR) shows the average rate of growth per year for a given period. In this indicator, it shows the average percentage growth of women and men ISCED 8 graduates in narrow fields of natural science and engineering.

It is respectively computed as follows:

CAGR for women graduates in a narrow field = $\left(F_{n,e}/F_{n,s}\right)^{1/N} - 1$

CAGR for men graduates in a narrow field = $\left(M_{n,e}/M_{n,s}\right)^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

$F_{n,s}$ denotes the number of ISCED 8 women graduates in narrow field *n* in the start year *s*;

$F_{n,e}$ denotes the number of ISCED 8 women graduates in narrow field *n* in the end year *e*;

$M_{n,s}$ denotes the number of ISCED 8 men graduates in narrow field *n* in the start year *s*;

$M_{n,e}$ denotes the number of ISCED 8 men graduates in narrow field *n* in the end year *e*.

For example, if there were 100 women PhD graduates from physical science in 2012 and 150 in 2016, the calculation would be:

CAGR for women PhD graduates in physical science = $(150/100)^{1/4} - 1 = 10.7\ \%$.

2.1.7.4. *Specifications*

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

The narrow fields of study in natural science and engineering according to the ISCED-F classification of fields of education and training are the following:

- 051 Biological and related sciences
- 052 Environment
- 053 Physical sciences
- 054 Mathematics and statistics
- 061 Information and Communication Technologies
- 071 Engineering and engineering trades
- 072 Manufacturing and processing
- 073 Architecture and construction.

2.1.7.5. *Comments and critical issues*

Data before 2012 were compiled using ISCED 97. Changes in the international classification standards are provided in Annex 1.

2.1.8. ***Ratio of ISCED 6 graduates to ISCED 6 entrants, by sex and broad field of study***

2.1.8.1. *Definition of indicator*

This indicator is the ratio of ISCED 6 graduates to ISCED 6 entrants, broken down by sex, broad field of study and country.

2.1.8.2. *Rationale*

The segregation between female and male scientists is already connected to early segregation in education pathways chosen by young women and men. This indicator shows the level of progress in increasing women's representation in the higher levels of education and research, considering their success in graduation at ISCED level 6.

2.1.8.3. *Computation method*

Data needed

$(F_{g,b})$ Number of women ISCED 6 graduates in broad field of studies *b*. **Unit: Number**.

$(F_{e,b})$ Number of women ISCED 6 entrants in broad field of studies *b*. **Unit: Number**.

$(M_{g,b})$ Number of men ISCED 6 graduates in broad field of studies *b*. **Unit: Number**.

$(M_{e,b})$ Number of men ISCED 6 entrants in broad field of studies *b*. **Unit: Number**.

Source of data
For entrants:

Eurostat – Education and Training Statistics *(online data code:* educ_uoe_ent02*)*

OECD (*http://stats.oecd.org; New entrants by field*).

For graduates:

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

Ratio of women ISCED 6 graduates to ISCED 6 entrants for a given field of study = $(F_{g,b}) \Big/ (F_{e,b})$

Ratio of men ISCED 6 graduates to ISCED 6 entrants for a given broad field of study = $(M_{g,b}) \Big/ (M_{e,b})$

where:

b denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

2.1.8.4. *Specifications*

ISCED level 6 corresponds to studies at Bachelor's or equivalent level according to the International Standard Classification of Education (ISCED-2011). The implementation of different ISCED versions should be considered when assessing the indicator overtime.

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

New entrants to a level of education are students who, during the course of the reference school or academic year, enter for the first time any programme in a given level of education, irrespective of whether they enter the programme at its beginning or at an advanced stage of it.

The broad fields of study according to the ISCED-F classification of fields of education and training are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services.

2.1.8.5. *Comments and critical issues*

The indicator compares the same reference year's entrants and graduates. These are two different groups of persons. Ideally one would need to follow up each year's entrants and count those that graduate. The proposed formulation is considered as a proxy.

2.1.9. ***Ratio of ISCED 8 entrants to ISCED 7 graduates, by sex and field of study (broad and narrow)***

2.1.9.1. *Definition of indicator*

This indicator is the ratio of ISCED 8 entrants to ISCED 7 graduates, broken down by sex, broad field of study and country.

2.1.9.2. *Rationale*

The segregation between female and male scientists is already connected to early segregation in education pathways chosen by young women and men. The indicator helps assess the propensity of women and men who graduate from ISCED level 7 to continue to ISCED level 8 studies.

2.1.9.3. *Computation method*

Data needed

$(F_{g,b})$ Number of women ISCED 7 graduates in broad field of studies *b*. **Unit: Number**.

$(F_{g,n})$ Number of women ISCED 7 graduates in narrow field of studies *n*. **Unit: Number**.

$(F_{e,b})$ Number of women ISCED 8 entrants in broad field of studies *b*. **Unit: Number**.

$(F_{e,n})$ Number of women ISCED 8 entrants in narrow field of studies *n*. **Unit: Number**.

$(M_{g,b})$ Number of men ISCED 7 graduates in broad field of studies *b*. **Unit: Number**.

$(M_{g,n})$ Number of men ISCED 7 graduates in narrow field of studies *n*. **Unit: Number**.

$(M_{e,b})$ Number of men ISCED 8 entrants in broad field of studies *b*. **Unit: Number**.

$(M_{e,n})$ Number of men ISCED 8 entrants in narrow field of studies *n*. **Unit: Number**.

Source of data

For entrants:

Eurostat – Education and Training Statistics *(online data code: educ_uoe_ent02)*

OECD (*http://stats.oecd.org; New entrants by field*).

For graduates:

Main data source: Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

Ratio of women ISCED 8 entrants to ISCED 7 graduates for a given **broad** field of study

$$= (F_{e,b}) \Big/ (F_{g,b})$$

Ratio of women ISCED 8 entrants to ISCED 7 graduates for a given **narrow** field of study

$$= (F_{e,n}) \Big/ (F_{g,n})$$

Ratio of men ISCED 8 entrants to ISCED 7 graduates for a given **broad** field of study

$$= (M_{e,b}) \Big/ (M_{g,b})$$

Ratio of men ISCED 8 entrants to ISCED 7 graduates for a given **narrow** field of study

$$= (M_{e,n}) \Big/ (M_{g,n})$$

where:

b denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields;

n denotes the narrow fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

2.1.9.4. *Specifications*

ISCED level 7 corresponds to studies at Master's or equivalent level and ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

New entrants to a level of education are students who, during the course of the reference school or academic year, enter for the first time any programme in a given level of education, irrespective of whether they enter the programme at its beginning or at an advanced stage of it.

The broad fields of study according to the ISCED-F classification of fields of education and training are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services.

The narrow fields of study in natural science and engineering according to the ISCED-F classification of fields of education and training are the following:

051 Biological and related sciences

052 Environment

053 Physical science

054 Mathematics and statistics

061 Information and Communication Technologies

071 Engineering and engineering trades

072 Manufacturing and processing

073 Architecture and construction.

2.1.9.5. *Comments and critical issues*

The indicator compares the same reference year's ISCED 8 entrants and ISCED 7 graduates. Ideally one would need to follow up each year's ISCED 7 graduates and count those that start ISCED 8 studies at some point in the future. The proposed formulation is considered as a proxy.

2.1.10. ***Ratio of ISCED 8 graduates to ISCED 8 entrants, by sex and broad field of study***

2.1.10.1. *Definition of indicator*

This indicator is the ratio of ISCED 8 graduates to ISCED 8 entrants, broken down by sex, broad field of study and country.

2.1.10.2. *Rationale*

The indicator shows the level of progress in increasing women's representation in the top levels of education and research, considering their success, as well as that of men, in graduation at ISCED level 8.

2.1.10.3. *Computation method*

The indicator is calculated from the following data.

Data needed

$(F_{g,b})$ Number of women ISCED 8 graduates in broad field of studies *b*. **Unit: Number**.

$(F_{e,b})$ Number of women ISCED 8 entrants in broad field of studies *b*. **Unit: Number**.

$(M_{g,b})$ Number of men ISCED 8 graduates in broad field of studies *b*. **Unit: Number**.

$(M_{e,b})$ Number of men ISCED 8 entrants in broad field of studies *b*. **Unit: Number**.

Source of data

For entrants:

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; New entrants by field*).

For graduates:

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

Ratio of women ISCED 8 graduates to ISCED 8 entrants for a given broad field of study

$$= {(F_{g,b})}\Big/{(F_{e,b})}$$

Ratio of men ISCED 8 graduates to ISCED 8 entrants for a given broad field of study

$$= {(M_{g,b})}\Big/{(M_{e,b})}$$

where:

b denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

2.1.10.4. *Specifications*

ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

New entrants to a level of education are students who, during the course of the reference school or academic year, enter for the first time any programme in a given level of education, irrespective of whether they enter the programme at its beginning or at an advanced stage of it.

The broad fields of study according to the ISCED-F classification of fields of education and training are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services.

2.1.10.5. *Comments and critical issues*

The indicator compares the same reference year's entrants and graduates. These are two different groups of persons. Ideally one would need to follow up each year's entrants and count those that graduate. The proposed formulation is considered as a proxy.

2.2. Eurostat – Human resources in science and technology

Content-based rationale

The European Commission has warned that 'gender segregation, or the tendency for men and women to take different jobs, is pervasive across Europe' (European Commission, 2014b, p. 26). Historically, women have been under-represented in scientific and technical fields. She Figures indicators based on human resources in science and technology (HRST) data explore this situation further. Many are designed to consider the extent to which available human resources in science and technology are being fully utilised, and whether differences by sex persist. These include the following: the proportion of tertiary-educated women employed as professionals or technicians, and the proportion of scientists and engineers in the total labour force, by sex.

Broad overview of the source

These data can be accessed through the Human Resources in Science and Technology (HRST) database on the Eurostat website: http://ec.europa.eu/eurostat/web/science-technology-innovation/data/database.

The **Human Resources in Science and Technology database** presents data on '**stocks**' and '**flows**'. Specifically, the data cover the 'demand for and supply of' highly qualified personnel (HQP) in the field of science and technology (S&T) and deal with 'stock', i.e., the current state of the labour force in S&T, and 'flow', i.e., the movement of HQP from job to job and from the academic sector to the public and private sectors. Data are disseminated on a yearly basis (European Commission, 2014c), and are used by both scientists and policymakers (European Commission, 2015d). Data from Eurostat is publicly available, regularly updated and accompanied by extensive methodological notes.

Many data breakdowns are available through the HRST database: sex, age, region, sector of economic activity, occupation, educational attainment and fields of education (however, not all combinations are possible). The HRST database uses some international classifications, including:

- the International Standard Classification of Education (ISCED 2011)
- the International Standard Classification of Occupations (ISCO-08)
- the Statistical classification of economic activities in the European Community (NACE Rev. 2).

In She Figures, indicators based on HRST data consider women and men's employment, including S&T occupations in general, and as scientists and engineers in particular. Additional data are required for these indicators from the Eurostat Labour Force Survey (EU-LFS) database (http://ec.europa.eu/eurostat/web/lfs/data/database) as indicated in the following subsections. For countries not covered by the EU-LFS (EU candidate countries, associated and G20 countries), data were assessed through the International Labour Organization Statistics Database (https://ilostat.ilo.org/).

A complete list of indicators falling into this category can be found in Annex 3 and their detailed description follows below.

2.2.1. ***Proportion of women among total employment in the EU***

2.2.1.1. *Definition of indicator*

This indicator presents the proportion of women in total employment as a starting point for considering their participation in different fields and sectors of the labour market.

2.2.1.2. *Rationale*

Even though, significant progress has been made, gender equality in total employment is far from reality and women are still under-represented in the labour market. (European Commission, 2017a). Boosting women's participation in employment is also fundamental to the EU's strategy for sustainable growth. (European Commission, 2019) This indicator considers the current representation of women in the labour market in general.

2.2.1.3. *Computation method*

Data needed

(F) Number of women in employment (aged 25–64). **Unit: Number**.

(T) Total number of people in employment (aged 25–64). **Unit: Number**.

Note that the numbers here are in thousand units.

Source of data

Eurostat – Labour Market Statistics *(online data code: lfsa_egan)*

Computation formula

The formula for this indicator is as follows:

Proportion of women among total employment = F/T

2.2.1.4. *Specifications*

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

2.2.1.5. *Comments and critical issues*

In the body of the She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E), and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

It is important to ensure the same age range when calculating the indicators in this figure. The age range 25–64 is available only through the detailed Labour Force Survey results, at data code *lfsa_egan*. There are minor differences between the detailed LFS results and the general LFS results (online data code: *lfsi_emp_a*).

2.2.2. *Compound annual growth rate (CAGR) of people in employment in the EU, by sex*

2.2.2.1. *Definition of indicator*

This indicator presents the average yearly growth in the number of women and men in total employment.

2.2.2.2. *Rationale*

Even though, significant progress has been made, gender equality in total employment is far from reality and women are still under-represented in the labour market. (European Commission, 2017a). Boosting women's participation in employment is also fundamental to the EU's strategy for sustainable growth. (European Commission, 2019) This indicator considers the current representation of women in the labour market in general.

2.2.2.3. *Computation method*

Data needed

(F) Number of women in employment (aged 25–64) in a start and an end year. **Unit: Number**.

(M) Number of men in employment (aged 25–64) in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Note that the numbers here are in thousand units. This does not affect the calculation of the compound annual growth rates.

Source of data

Eurostat – Labour Market Statistics *(online data code: lfsa_egan)*

Computation formula

The compound annual growth rate shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of women employees and men employees in a given period. For women and men, it is respectively computed as follows:

CAGR of women in employment = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

CAGR of men in employment = $(\mathrm{M_e}/\mathrm{M_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

N denotes the number of years in the reference period (in other words, $e - s$);

F_s denotes the number of women in employment in the start year;

F_e denotes the number of women in employment in the end year;

M_s denotes the number of men in employment in the start year;

M_e denotes the number of men in employment in the end year.

For example, if there were 1000 men in employment in 2006 and 1500 in 2016, the calculation would be:

$$CAGR\ for\ men\ in\ employment = (1500/1000)^{\frac{1}{10}} - 1 = 4.14\ \%$$

2.2.2.4. *Specifications*

The EU Labour Force Survey (LFS) defines **employed persons** as 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

2.2.2.5. *Comments and critical issues*

In the body of the She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E), and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

It is important to ensure the same age range when calculating the indicators in this figure. The age range 25-64 is available only through the detailed Labour Force Survey results, at data code *lfsa_egan*. There are minor differences between the detailed LFS results and the general LFS results (online data code: *lfsi_emp_a*).

2.2.3. ***Proportion of women among tertiary-educated and employed as professionals or technicians (HRSTC) in the EU***

2.2.3.1. *Definition of indicator*

This indicator presents the proportion of women within the Human Resources in Science and Technology Core group. This category covers those who have completed tertiary education (in any subject) and are employed in a science and technology (S&T) occupation (either as professionals or technicians).

2.2.3.2. *Rationale*

Fostering greater investment in science and technology is a core part of the European vision for growth. The EU's main funding instruments for research and innovation (R&I), the Horizon 2020 and the Horizon Europe (2021-2027) programmes, recognise the economic benefits that science and technology can deliver (DG Research and Innovation, 2014 and DG Research and Innovation, 2019a), whilst the Europe 2030 strategy sees this as a priority growth area.

This indicator considers the extent to which the available human resources in science and technology are being fully utilised, broken down by sex.

2.2.3.3. *Computation method*

Data needed

(F) Number of tertiary-educated women aged 25–64 who are employed as professionals or technicians (Human Resources in Science and Technology – Core (HRSTC)). **Unit: Number**.

(T) Total number of tertiary-educated people aged 25–64 who are employed as professionals or technicians (Human Resources in Science and Technology – Core (HRSTC)). **Unit: Number**.

Note that the numbers are in thousand units.

Source of data

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

Specifications

The formula for this indicator is:

Proportion of women among the HRSTC group = F / T

2.2.3.4. *Specifications*

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

HRST are persons that have successfully completed tertiary education (at least ISCED2011 level 7) or persons employed in S&T occupations even if not with tertiary education. Thus, the three categories of HRST are:

HRSTO: Persons employed in an S&T occupation;

HRSTE: Persons with tertiary education in any field of study;

HRSTC: Persons with tertiary education in any field of study and employed in an S&T occupation (the intersection of the former two groups).

S&T occupations are all occupations classified into major group 2, 'Professionals', or 3, 'Technicians and Associate Professionals', of the ISCO-08 International Standard Classification of Occupations.

- ISCO Major Group 2 (**Professionals**) – Occupations whose main tasks usually include: conducting analysis and research and developing concepts, theories and operational methods; advising on or applying existing knowledge related to physical sciences, mathematics, engineering and technology, life sciences, medical and health services, social sciences and humanities; teaching the theory and practice of one or more disciplines at different educational levels; teaching and educating persons with learning difficulties or special needs; providing various business, legal and social services; creating and performing works of art; providing spiritual guidance; preparing scientific papers and reports.

- ISCO Major Group 3 (**Technicians and Associate Professionals**) – Occupations whose main tasks usually include: undertaking and carrying out technical work connected with research and the application of concepts and operational methods in the fields of physical sciences including engineering and technology, life sciences

including the medical profession, and social sciences and humanities; initiating and carrying out various technical services related to trade, finance and administration including administration of government laws and regulations and to social work; providing technical support for the arts and entertainment; participating in sporting activities; executing some religious tasks.

2.2.3.5. *Comments and critical issues*

In the body of the She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E), and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

Since 2011, the ISCO-08 edition has been used for Eurostat statistics on human resources in science and technology. The new version of ISCO affects the precise population covered by HRSTO, due to changes in the definition of 'Professionals' and 'Technicians and Associate Professionals'. This has an impact on comparability across She Figures editions (which use the older ISCO-88 classifications up to 2011) and affects data interpretation. For more information, see Annex 1.

2.2.4. ***Compound annual growth rate (CAGR) of tertiary-educated people who are employed as professionals or technicians (HRSTC) in the EU, by sex***

2.2.4.1. *Definition of indicator*

This indicator presents the average percentage growth each year in the number of women and men in the Human Resources in Science and Technology – Core (HRSTC) group. This covers those who have completed tertiary education (in any subject) and are employed in a science and technology (S&T) occupation (either as professionals or technicians).

2.2.4.2. *Rationale*

Fostering greater investment in science and technology is a core part of the European vision for growth. The EU's main funding instruments for research and innovation (R&I), the Horizon 2020 and the Horizon Europe (2021-2027) programmes, recognise the economic benefits that science and technology can deliver (DG Research and Innovation, 2014 and DG Research and Innovation, 2019a), whilst the Europe 2030 strategy sees this as a priority growth area.

This indicator considers whether there have been any changes to the use of available human resources in science and technology over time (broken down by sex).

2.2.4.3. *Computation method*

<u>Data needed</u>

(F) Number of tertiary-educated women, aged 25–64 who are employed as professionals or technicians (HRSTC) in a start and an end year. **Unit: Number**.

(M) Number of tertiary-educated men, aged 25–64 who are employed as professionals or technicians (HRSTC) in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Note that the numbers here are in thousand units. However, this does not affect the calculation of the compound annual growth rates.

Source of data

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

Computation formula

The CAGR shows the yearly average rate of growth for a given period. In this case, it shows the average percentage growth per year in the number of tertiary-educated women and men employed in S&T occupations. For women and men, it is respectively computed as follows:

CAGR of women = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

CAGR of men = $(\mathrm{M_e}/\mathrm{M_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

N denotes the number of years in the reference period;

F_s denotes the number of women in the HRSTC category in the start year;

F_e denotes the number of women in the HRSTC category in the end year;

M_s denotes the number of men in the HRSTC category in the start year;

M_e denotes the number of men in the HRSTC category in the end year.

For example, if there were 1 000 tertiary-educated women employed in S&T occupations in 2012, and 1 500 in 2016, the calculation would be: $(1500/1000)^{1/4} - 1 = 10.7\,\%$.

2.2.4.4. *Specifications*

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

HRST are persons that have successfully completed tertiary education (at least ISCED2011 level 7) or persons employed in S&T occupations even if not with tertiary education. Thus, the three categories of HRST are:

HRSTO: Persons employed in an S&T occupation;

HRSTE: Persons with tertiary education in any field of study;

HRSTC: Persons with tertiary education in any field of study and employed in an S&T occupation (the intersection of the former two groups).

S&T occupations are all occupations classified into major group 2, 'Professionals', or 3, 'Technicians and Associate Professionals', of the ISCO-08 International Standard Classification of Occupations.

- ISCO Major Group 2 (**Professionals**) – Occupations whose main tasks usually include: conducting analysis and research and developing concepts, theories and operational methods; advising on or applying existing knowledge related to physical sciences, mathematics, engineering and technology, life sciences, medical and health services, social sciences and humanities; teaching the theory and practice of one or more disciplines at different educational levels; teaching and educating persons with learning difficulties or special needs; providing various business, legal and social services; creating and performing works of art; providing spiritual guidance; preparing scientific papers and reports.

- ISCO Major Group 3 (**Technicians and Associate Professionals**) – Occupations whose main tasks usually include: undertaking and carrying out technical work connected with research and the application of concepts and operational methods in the fields of physical sciences including engineering and technology, life sciences including the medical profession, and social sciences and humanities; initiating and carrying out various technical services related to trade, finance and administration including administration of government laws and regulations and to social work; providing technical support for the arts and entertainment; participating in sporting activities; executing some religious tasks.

2.2.4.5. *Comments and critical issues*

In the body of the She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E), and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

Since 2011, the ISCO-08 edition has been used for Eurostat statistics on human resources in science and technology. The new version of ISCO affects the precise population covered by HRSTO, due to changes in the definition of 'Professionals' and 'Technicians and Associate Professionals'. This has an impact on comparability across She Figures editions (which use the older ISCO-88 classifications up to 2011) and affects data interpretation. For more information, see Annex 1.

2.2.5. ***Proportion of women among scientists and engineers (S&E) in the EU***

2.2.5.1. *Definition of indicator*

This indicator presents the proportion of women within the total number of scientists and engineers in employment.

2.2.5.2. *Rationale*

According to a report by the European Commission, 'gender segregation, or the tendency for men and women to take different jobs, is pervasive across Europe. Only 16 % of all employees work in mixed occupations (i.e., where the proportions of men and women are between 40 % and 60 %)' (European Commission, 2014b, p. 26). Traditionally, women are severely under-represented in the population of scientists and engineers. This is of particular significance given that such professionals are 'often the innovators at the centre of technology-led development' (European Commission, 2015c). By considering the sex breakdown for employed engineers and scientists, this indicator enables one to

see whether there have any advances in equalising the representation of women and men in this area.

For a full explanation of 'gender segregation' and other terms, see Annex 2.

2.2.5.3. *Computation method*

Data needed

(F) Number of women employed as scientists and engineers, aged 25–64. **Unit: Number**.

(T) Total number of people employed as scientists and engineers, aged 25–64. **Unit: Number**.

Note that the numbers are in thousand units.

Source of data

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

Computation formula

Proportion of women among persons employed as S&E = F/T

2.2.5.4. *Specifications*

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

S&E (Scientists and engineers) are people who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge relating to the fields which are covered by one of the following occupations defined in the ISCO-08:

science and engineering professionals (ISCO-08 code: 21)

health professionals (ISCO-08 code: 22)

information and communications technology professionals (ISCO-08 code: 25).

- ISCO Sub-major Group 21 (**Science and engineering professionals**) – Occupations whose main tasks usually include: conducting research, enlarging, advising on or applying scientific knowledge obtained through the study of structures and properties of physical matter and phenomena, chemical characteristics and processes of various substances, materials and products, all forms of human, animal and plant life and of mathematical and statistical concepts and methods; advising on, designing and directing construction of buildings, towns and traffic systems, or civil engineering and industrial structures, as well as machines and other equipment; advising on and applying mining methods and ensuring their optimum use; surveying land and see and making maps; studying and advising on technological aspects of particular materials, products and processes, and on efficiency of production and work organisation; preparing scientific papers and reports.

- ISCO Sub-major Group 22 (**Health professionals**) – Occupations whose main tasks usually include: conducting research and obtaining scientific knowledge through the study of human and animal disorders and illnesses and ways of

treating them; advising on or applying preventive and curative measures, or promoting health; preparing scientific papers and reports.

- ISCO Sub-major Group 25 (**Information and communications technology professionals**) – Occupations whose main tasks usually include: researching information technology use in business functions; identifying areas for improvement and researching the theoretical aspects and operational methods for the use of computers; evaluating, planning and designing hardware or software configurations for specific applications including for Internet, Intranet and multimedia systems; designing, writing, testing and maintaining computer programs; designing and developing database architecture and database management systems; developing and implementing security plans and data administration policy, and administering computer networks and related computing environments; analysing, developing, interpreting and evaluating complex system design and architecture specifications, data models and diagrams in the development, configuration and integration of computer systems.

2.2.5.5. *Comments and critical issues*

In the body of the She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E), and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

Since 2011, the ISCO-08 edition has been used for Eurostat statistics on human resources in science and technology. The new version of ISCO affects the precise population covered by HRSTO, due to changes in the definition of 'Professionals' and 'Technicians and Associate Professionals'. This has an impact on comparability across She Figures editions (which use the older ISCO-88 classifications up to 2011) and affects data interpretation. For more information, see Annex 1.

2.2.6. ***Compound annual growth rate (CAGR) of scientists and engineers (S&E) in the EU, by sex***

2.2.6.1. *Definition of indicator*

This indicator presents the average percentage growth each year in the number of people employed as scientists and engineers, broken down by sex.

2.2.6.2. *Rationale*

The European Commission has reported on the persistence of 'gender segregation' in the labour market (the concentration of women and men in particular fields and at particular levels) (European Commission, 2014b, p. 26). Traditionally, women are severely under-represented in the population of scientists and engineers. This indicator enables one to see the rate at which women and men's employment as scientists and engineers has been growing over time. To reduce the gender imbalance, it is likely that women's representation will need to be growing at a faster rate than that for men.

For a full explanation of 'gender segregation' and other terms, see Annex 2.

2.2.6.3. *Computation method*

Data needed

(F) Number of women, aged 25–64, employed as scientists and engineers in a start and an end year. **Unit: Number**.

(M) Number of men, aged 25–64, employed as scientists and engineers in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Note that the numbers here are in thousand units. However, this does not affect the calculation of the compound annual growth rates.

Source of data

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

Computation formula

The CAGR shows the yearly average rate of growth for a given period. In this case, it shows the average percentage growth of women and men employed as scientists and engineers. For women and men, it is respectively computed as follows:

CAGR of women scientists and engineers = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

CAGR of men scientists and engineers = $(\mathrm{M_e}/\mathrm{M_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

N denotes the number of years in the reference period;

$\mathrm{F_s}$ denotes the number of women S&E in the start year;

$\mathrm{F_e}$ denotes the number of women S&E in the end year;

$\mathrm{M_s}$ denotes the number of men S&E in the start year;

$\mathrm{M_e}$ denotes the number of men S&E in the end year.

For example, if there were 100 women S&E in 2012, and 150 in 2016, the calculation would be: $CAGR\ for\ women\ S\&E = (150/100)^{1/4} - 1 = 10.7\ \%$.

2.2.6.4. *Specifications*

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

S&E (Scientists and engineers) are people who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge relating to the fields which are covered by one of the following occupations defined in the ISCO-08:

science and engineering professionals (ISCO-08 code: 21)

health professionals (ISCO-08 code: 22)

information and communications technology professionals (ISCO-08 code: 25).

- ISCO Sub-major Group 21 (**Science and engineering professionals**) – Occupations whose main tasks usually include: conducting research, enlarging, advising on or applying scientific knowledge obtained through the study of structures and properties of physical matter and phenomena, chemical characteristics and processes of various substances, materials and products, all forms of human, animal and plant life and of mathematical and statistical concepts and methods; advising on, designing and directing construction of buildings, towns and traffic systems, or civil engineering and industrial structures, as well as machines and other equipment; advising on and applying mining methods and ensuring their optimum use; surveying land and see and making maps; studying and advising on technological aspects of particular materials, products and processes, and on efficiency of production and work organisation; preparing scientific papers and reports.

- ISCO Sub-major Group 22 (**Health professionals**) – Occupations whose main tasks usually include: conducting research and obtaining scientific knowledge through the study of human and animal disorders and illnesses and ways of treating them; advising on or applying preventive and curative measures, or promoting health; preparing scientific papers and reports.

- ISCO Sub-major Group 25 (**Information and communications technology professionals**) – Occupations whose main tasks usually include: researching information technology use in business functions; identifying areas for improvement and researching the theoretical aspects and operational methods for the use of computers; evaluating, planning and designing hardware or software configurations for specific applications including for Internet, Intranet and multimedia systems; designing, writing, testing and maintaining computer programs; designing and developing database architecture and database management systems; developing and implementing security plans and data administration policy, and administering computer networks and related computing environments; analysing, developing, interpreting and evaluating complex system design and architecture specifications, data models and diagrams in the development, configuration and integration of computer systems.

2.2.6.5. *Comments and critical issues*

In the body of the She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E), and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

Since 2011, the ISCO-08 edition has been used for Eurostat statistics on human resources in science and technology. The new version of ISCO affects the precise population covered by HRSTO, due to changes in the definition of 'Professionals' and 'Technicians and Associate Professionals'. This has an impact on comparability across She Figures editions

(which use the older ISCO-88 classifications up to 2011) and affects data interpretation. For more information, see Annex 1.

2.2.7. *Proportion of tertiary educated and employed as professionals and technicians (HRSTC) among tertiary-educated population (HRSTE), by sex*

2.2.7.1. *Definition of indicator*

This indicator identifies the proportions of highly educated men and women who are employed as professionals or technicians. Specifically, it presents the proportions of women and men who are tertiary educated *and* working in a science and technology occupation, out of the number of women and men who are tertiary educated. Those in science and technology occupations are those working as 'Professionals' or 'Technicians and Associate Professionals'.

2.2.7.2. *Rationale*

Fostering greater investment in science and technology is a core part of the European vision for growth. The EU's main funding instruments for research and innovation (R&I), the Horizon 2020 and the Horizon Europe (2021-2027) programmes, recognise the economic benefits that science and technology can deliver (DG Research and Innovation, 2014 and DG Research and Innovation, 2019a), whilst the Europe 2030 strategy sees this as a priority growth area.

2.2.7.3. *Computation method*

Data needed

(F) Number of tertiary-educated women, aged 25–64 (HRSTE). **Unit: Number**.

(F_c) Number of tertiary-educated women, aged 25–64, who are also employed in S&T occupations (HRSTC). **Unit: Number**.

(M) Number of tertiary-educated men, aged 25–64 (HRSTE). **Unit: Number**.

(M_c) Number of tertiary-educated men, aged 25–64, who are also employed in S&T occupations (HRSTC). **Unit: Number**.

Note that the numbers here are in thousand units.

Source of data

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

Computation formula

The formula for this indicator is:

Percentage of tertiary educated women working in S&T occupations = F_c / F

Percentage of tertiary educated men working in S&T occupations = M_c / M

where:

c denotes tertiary-educated people who are also working in an S&T occupation.

2.2.7.4. *Specifications*

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

HRST are persons that have successfully completed tertiary education (at least ISCED2011 level 7) or persons employed in S&T occupations even if not with tertiary education. Thus, the three categories of HRST are:

HRSTO: Persons employed in an S&T occupation;

HRSTE: Persons with tertiary education in any field of study;

HRSTC: Persons with tertiary education in any field of study and employed in an S&T occupation (the intersection of the former two groups).

S&T occupations are all occupations classified into major group 2, 'Professionals', or 3, 'Technicians and Associate Professionals', of the ISCO-08 International Standard Classification of Occupations.

- ISCO Major Group 2 (**Professionals**) – Occupations whose main tasks usually include: conducting analysis and research and developing concepts, theories and operational methods; advising on or applying existing knowledge related to physical sciences, mathematics, engineering and technology, life sciences, medical and health services, social sciences and humanities; teaching the theory and practice of one or more disciplines at different educational levels; teaching and educating persons with learning difficulties or special needs; providing various business, legal and social services; creating and performing works of art; providing spiritual guidance; preparing scientific papers and reports.

- ISCO Major Group 3 (**Technicians and Associate Professionals**) – Occupations whose main tasks usually include: undertaking and carrying out technical work connected with research and the application of concepts and operational methods in the fields of physical sciences including engineering and technology, life sciences including the medical profession, and social sciences and humanities; initiating and carrying out various technical services related to trade, finance and administration including administration of government laws and regulations and to social work; providing technical support for the arts and entertainment; participating in sporting activities; executing some religious tasks.

2.2.7.5. *Comments and critical issues*

Since 2011, the ISCO-08 edition has been used for Eurostat statistics on human resources in science and technology. The new version of ISCO affects the precise population covered by HRSTO, due to changes in the definition of 'Professionals' and 'Technicians and Associate Professionals'. This has an impact on comparability across She Figures editions (which use the older ISCO-88 classifications up to 2011) and affects data interpretation. For more information, see Annex 1.

2.2.8. *Proportion of scientists and engineers (S&E) among the total labour force, by sex*

2.2.8.1. *Definition of indicator*

This indicator presents the proportion of scientists and engineers in the labour force, by sex.

2.2.8.2. *Rationale*

According to a report by the European Commission, 'gender segregation, or the tendency for men and women to take different jobs, is pervasive across Europe. Only 16 % of all employees work in mixed occupations (i.e., where the proportions of men and women are between 40 % and 60 %)' (European Commission, 2014b, p. 26). Traditionally, women are severely under-represented in the population of scientists and engineers. This is of particular significance given that such professionals are 'often the innovators at the centre of technology-led development' (European Commission, 2015c). By comparing the proportion of women and men engineers and scientists in the entire labour force, this indicator offers one measure of the level of segregation in this area (which is sometimes seen as connected to earlier segregation in the education pathways chosen by young women and men).

2.2.8.3. *Computation method*

Data needed

(F) Number of women working as scientists and engineers aged 25–64. **Unit: Number**.

(M) Number of men working as scientists and engineers aged 25–64. **Unit: Number**.

(T) Total number of people (both men and women) in the labour force, aged 25–64 (definition provided below). **Unit: Number**.

Note that the number of people in the labour force (T) is in thousand units.

Source of data

For F and M:

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

For T:

Eurostat – Labour Market Statistics *(online data code: lfsa_agan)*

Computation formula

Proportion of women scientists and engineers among the total labour force = F/T

Proportion of men scientists and engineers among the total labour force = M/T

2.2.8.4. *Specifications*

According to the EU Labour Force Survey (LFS), **the labour force** (also termed 'active population') is defined as the sum of employed and unemployed persons. **Employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'. **Unemployed persons** are 'all persons aged 15 to 74 who were not employed during the reference week, were available to start work within the two weeks following

the reference week and had been actively seeking work in the four weeks preceding the reference week or had already found a job to start within the next three months'.

S&E (Scientists and engineers) are people who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge relating to the fields which are covered by one of the following occupations defined in the ISCO-08:

science and engineering professionals (ISCO-08 code: 21)

health professionals (ISCO-08 code: 22)

information and communications technology professionals (ISCO-08 code: 25).

- ISCO Sub-major Group 21 (**Science and engineering professionals**) – Occupations whose main tasks usually include: conducting research, enlarging, advising on or applying scientific knowledge obtained through the study of structures and properties of physical matter and phenomena, chemical characteristics and processes of various substances, materials and products, all forms of human, animal and plant life and of mathematical and statistical concepts and methods; advising on, designing and directing construction of buildings, towns and traffic systems, or civil engineering and industrial structures, as well as machines and other equipment; advising on and applying mining methods and ensuring their optimum use; surveying land and see and making maps; studying and advising on technological aspects of particular materials, products and processes, and on efficiency of production and work organisation; preparing scientific papers and reports.

- ISCO Sub-major Group 22 (**Health professionals**) – Occupations whose main tasks usually include: conducting research and obtaining scientific knowledge through the study of human and animal disorders and illnesses and ways of treating them; advising on or applying preventive and curative measures, or promoting health; preparing scientific papers and reports.

- ISCO Sub-major Group 25 (**Information and communications technology professionals**) – Occupations whose main tasks usually include: researching information technology use in business functions; identifying areas for improvement and researching the theoretical aspects and operational methods for the use of computers; evaluating, planning and designing hardware or software configurations for specific applications including for Internet, Intranet and multimedia systems; designing, writing, testing and maintaining computer programs; designing and developing database architecture and database management systems; developing and implementing security plans and data administration policy, and administering computer networks and related computing environments; analysing, developing, interpreting and evaluating complex system design and architecture specifications, data models and diagrams in the development, configuration and integration of computer systems.

2.2.8.5. *Comments and critical issues*

Since 2011, the ISCO-08 edition has been used for Eurostat statistics on human resources in science and technology. The new version of ISCO affects the precise population covered by HRSTO, due to changes in the definition of 'Professionals' and 'Technicians and Associate Professionals'. This has an impact on comparability across She Figures editions (which use the older ISCO-88 classifications up to 2011) and affects data interpretation. For more information, see Annex 1.

Note that this indicator is calculated differently to the 'Percentage of active population' unit at the data code *hrst_st_ncat*. In She Figures, the denominator for this indicator is

the total labour force (women and men combined), whereas on Eurostat the denominator is restricted to either women or men.

2.2.9. ***Unemployment rate of tertiary educated people, by sex***

2.2.9.1. *Definition of indicator*

This indicator presents the rate of unemployment for persons who have completed tertiary education, by sex.

2.2.9.2. *Rationale*

The education level does have significant influence to the employment level in most EU countries, but the differences of this influence vary crucially among the countries due to some historical reasons, labour market structural differences and unemployment insurance system peculiarities (Snieska V. et al, 2015). The indicator helps to compare the employment outlook of female and male science and technology resources and to reveal potential disadvantages of one of the two sexes.

2.2.9.3. *Computation method*

Data needed

(F_u) Number of unemployed women, aged 25-64, with tertiary education. **Unit: Number**.

(F) Number of women in the labour force, aged 25-64, with tertiary education. **Unit: Number**.

(M_u) Number of unemployed men, aged 25-64, with tertiary education. **Unit: Number**.

(M) Number of men in the labour force, aged 25-64, with tertiary education. **Unit: Number**.

(T_u) Number of unemployed persons, aged 25-64, with tertiary education. **Unit: Number**.

(T) Number of persons in the labour force, aged 25-64, with tertiary education. **Unit: Number**.

Source of data

Eurostat - Human resources in Science and Technology (*online data code:* hrst_st_nunesex)

International Labour Organization (https://ilo.org/global/statistics-and-databases; Unemployment by sex, age and education).

Computation formula

Unemployment rate of tertiary educated women = F_u/F

Unemployment rate of tertiary educated men = M_u/M

Unemployment rate of tertiary educated persons = T_u/T

2.2.9.4. *Specifications*

According to the EU Labour Force Survey (LFS), **unemployed persons** are 'all persons aged 15 to 74 who were not employed during the reference week, were available to start work within the two weeks following the reference week and had been actively seeking work in the four weeks preceding the reference week or had already found a job to start within the next three months'. **The labour force** (also termed 'active population') is defined as the sum of employed and unemployed persons.

2.2.9.5. *Comments and critical issues*

Data from Eurostat do not need computation and can be extracted from the above-mentioned dataset. The indicator based on data from International Labour Organisation (ILO) needs to be computed with the formula shown above. Definition of unemployed persons does not change in ILO.

2.2.10. ***Proportion of women among self-employed individuals within Information and Communication Technology (ICT) and Science and Engineering (S&E) professionals***

2.2.10.1. *Definition of indicator*

This indicator calculates the proportion of women who are self-employed within the occupations of Information and Communication Technology Professionals (ISCO-08 Division 25) and Science and Engineering Professionals (ISCO-08 Division 21).

2.2.10.2. *Rationale*

The European Commission's Entrepreneurship Action Plan 2020 points to the gender gap in the share of women who are self-employed in the EU (European Commission, 2013a). Furthermore, the European Commission's 'Women in Digital' policy aims to foster women's labour market participation in technology-oriented occupations and in knowledge-intensive sectors including ICT (European Commission, 2020a). Taking these priorities into account, this indicator sheds light on women's share of self-employment specifically within the ICT sector.

2.2.10.3. *Computation method*

Data needed

(F_i) Number of self-employed women with and without employees. **Unit: Number**.

(T_i) Total number of self-employed individuals with and without employees. **Unit: Number**.

(i) Denotes the occupation type

Source of data

Eurostat –European Labour Force Survey (EU-LFS).

Computation formula

Proportion of self-employed women among ICT and S&E professionals = F_i/T_i

2.2.10.4. *Specifications*

According to the EU Labour Force Survey, **self-employed persons with employees** refer to persons who work in their own business or professional practice for the purpose of earning a profit, and who employ at least one other person. Similarly, **self-employed persons without employees** refer to persons who work in their own business or professional practice for the purpose of earning a profit, and who do not employ any other person.

ISCO Sub-major Group 25 (**Information and communications technology professionals**) – Occupations whose main tasks usually include: researching information technology use in business functions; identifying areas for improvement and researching the theoretical aspects and operational methods for the use of computers; evaluating, planning and designing hardware or software configurations for specific applications including for Internet, Intranet and multimedia systems; designing, writing, testing and maintaining computer programs; designing and developing database architecture and database management systems; developing and implementing security plans and data administration policy, and administering computer networks and related computing environments; analysing, developing, interpreting and evaluating complex system design and architecture specifications, data models and diagrams in the development, configuration and integration of computer systems.

ISCO Sub-major Group 21 (**Science and Engineering professionals**) – Occupations whose main tasks usually include: conducting research, enlarging, advising on or applying scientific knowledge obtained through the study of structures and properties of physical matter and phenomena, chemical characteristics and processes of various substances, materials and products, all forms of human, animal and plant life and of mathematical and statistical concepts and methods; advising on, designing and directing construction of buildings, towns and traffic systems, or civil engineering and industrial structures, as well as machines and other equipment; advising on and applying mining methods and ensuring their optimum use; surveying land and see and making maps; studying and advising on technological aspects of particular materials, products and processes, and on efficiency of production and work organisation; preparing scientific papers and reports.

2.2.10.5. *Comments and critical issues*

Since 2011, the ISCO-08 edition has been used for Eurostat statistics.

2.3. Eurostat – High-tech industry and knowledge-intensive services

Content-based rationale

These indicators were originally developed in line with Europe's 2020 vision of 'smart growth' to determine the extent to which women's full educational capacities are being exploited and as a way to gauge the EU's use of available human capital and women's role within a priority area of the economy. The European Commission is building a European Education Area by 2025 in order to harness the full potential of education, training and culture as drivers for job creation, economic growth and social fairness (European Commission, 2020b). The fast-changing labour market and societal transitions require people with high-level skills and advanced knowledge.

The indicators for high-tech industry and knowledge-intensive services include employment in knowledge-intensive activities (KIA), and employment in knowledge-intensive activities – business industries (KIABI).

Broad overview of the source

These data can be accessed from the Science, Technology and Innovation Statistics database on Eurostat's website (http://ec.europa.eu/eurostat/web/science-technology-innovation/overview).

'Statistics on high-tech industry and knowledge-intensive services' (sometimes referred to simply as 'high-tech statistics') cover statistics concerning employment, economic indicators, patents and products in the high-tech categories of the manufacturing sector, as well as the knowledge-intensive service sector (European Commission, 2015b). Data from Eurostat is publicly available, regularly updated and accompanied by extensive methodological notes.

A complete list of indicators falling into this category can be found in Annex 3 and their detailed description follows below.

2.3.1. *Proportion of employment in knowledge-intensive activities (KIA) among total employment, by sex*

2.3.1.1. *Definition of indicator*

This indicator presents the relative presence of employed women and men in KIA,[10] covering all sectors of the economy.

2.3.1.2. *Rationale*

The Europe's strategy for sustainable growth aims to foster an economy based on people with advanced knowledge and high-level skills (European Commission, 2019). This indicator reveals the extent to which women's full educational capacities are being utilised, by measuring the relative proportion of women and men in KIA.

2.3.1.3. *Computation method*

Data needed

(F) Number of women employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{F_k})$ Number of women employed specifically in KIA. **Unit: Number**.

(M) Number of men employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{M_k})$ Number of men employed specifically in KIA. **Unit: Number**.

(T) Total number of people employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{T_k})$ Total number of people employed specifically in KIA. **Unit: Number**.

Source of data

Eurostat – High-tech industry and knowledge-intensive services *(online data code: htec_kia_emp2)*

Note that this data code (*htec_kia_emp2*) provides the KIA employment rates and the numerators of the computation formula but not the denominators.

Computation formula

Respectively, the formulas for this indicator are:

Proportion of women employed in knowledge intensive activities = $\mathrm{F_k} / F$

Proportion of men employed in knowledge intensive activities = $\mathrm{M_k} / M$

Proportion of persons employed in knowledge intensive activities = $\mathrm{T_k} / T$

where:

k denotes knowledge-intensive-activities specifically.

[10] Activities where more than one third of the workforce is tertiary educated.

2.3.1.4. *Specifications*

The **International Standard Classification of Education (ISCED 2011)** categorises education programmes by level. Tertiary-educated people are those who have graduated from the following stages:

- The first stage, which includes largely theory-based programmes to provide sufficient qualifications to gain entry to advanced research programmes and professions with high skills requirements and programmes which are practically, technically or occupationally specific (ISCED 5, 6 and 7).

- The second stage, which leads to the award of an advanced research qualification (e.g., PhD, non-PhD programmes with an advanced research component, etc.). The programmes are devoted to advanced study and original research (ISCED 8).

An activity is classified as '**knowledge-intensive'** if tertiary-educated people employed in this activity represent more than 33% of the total employment in the activity. The definition is based on the average number of employed persons aged 25–64 at aggregated EU-27 level.

- Women may be over-represented in some knowledge-intensive activities that are not related to science and research. This indicator does not enable one to analyse differences in representation across the activities.

The activities come from the NACE Rev. 2 categories (2-digit level),[11] based on EU Labour Force Survey (LFS) data. NACE refers to the European Community's statistical classification of economic activities.

In this indicator, there is one aggregate in use based on the following classification: total knowledge-intensive activities (KIA). The lists of activities included in each aggregate, according to NACE Rev. 2 (2-digit level), are presented in the table below:

Table 1 Total knowledge-intensive activities (KIA), NACE Rev. 2

Codes	Description
09	Mining support service activities
19	Manufacture of coke and refined petroleum products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
26	Manufacture of computer, electronic and optical products
51	Air transport
58	Publishing activities
59	Motion picture, video and television programme production, sound recording
60	Programming and broadcasting activities
61	Telecommunications

[11] Due to the revision of the NACE from NACE Rev. 1.1 to NACE Rev. 2, the definition of high-technology industries and knowledge-intensive services changed in 2008. See references: Eurostat, 2008

Codes	Description
62	Computer programming, consultancy and related activities
63	Information service activities
64	Financial service activities, except insurance and pension funding
65	Insurance, reinsurance and pension funding, except compulsory social security
66	Activities auxiliary to financial services and insurance activities
69	Legal and accounting activities
70	Activities of head offices; management consultancy activities
71	Architectural and engineering activities; technical testing and analysis
72	Scientific research and development
73	Advertising and market research
74	Other professional, scientific and technical activities
75	Veterinary activities
78	Employment activities
79	Travel agency, tour operator reservation service and related activities
84	Public administration and defence; compulsory social security
85	Education
86	Human health activities
90	Creative, arts and entertainment activities
91	Libraries, archives, museums and other cultural activities
94	Activities of membership organisations
99	Activities of extraterritorial organisations and bodies

2.3.2. *Proportion of employment in knowledge-intensive activities – Business industries (KIABI) out of total employment, by sex*

2.3.2.1. *Definition of indicator*

Similar to the previous indicator, this indicator shows the relative proportion of employed women and men in knowledge-intensive activities (KIA) (activities where more than one third of the workforce is tertiary educated), although it is restricted to business industries only.

2.3.2.2. *Rationale*

KIA are key to the EU's vision of fostering a knowledge-based economy. The term itself encompasses a wide range of activities, although this indicator restricts the focus to business industries (KIABI). This is a particularly important sector of the economy to examine, given that the EU considers 'innovation through business activities' to represent a strength of national research and innovation systems.[12] Assessing the relative proportion of women and men's employment in KIABI is thus a key way of gauging the EU's use of available human capital, as well as the foundation for considering women's role within a priority area of the economy.

2.3.2.3. *Computation method*

Data needed

(F) Number of women employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{F_b})$ Number of women employed specifically in KIABI. **Unit: Number**.

(M) Number of men employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{M_b})$ Number of men employed specifically in KIABI. **Unit: Number**.

(T) Total number of people employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{T_b})$ Total number of people employed in specifically in KIABI. **Unit: Number**.

Source of data

Eurostat – High-tech industry and knowledge-intensive services *(online data code: htec_kia_emp2)*

Note that this data code (*htec_kia_emp2*) provides the numerators ($\mathrm{F_b}$, $\mathrm{M_b}$, $\mathrm{T_b}$) but not the denominators (T, F or M) of the above computation formula. However, it provides the KIA employment rates in this indicator.

[12] This is evident in, for example, European Commission (2021), *Innovation Union Scoreboard 2021.*

Computation formula

Respectively, the formulas for this indicator are:

Proportion of women employed in KIABI = F_b / F

Proportion of men employed in KIABI = M_b / M

Proportion of persons employed in KIABI = T_b / T

2.3.2.4. *Specifications*

The **International Standard Classification of Education (ISCED 2011)** categorises education programmes by level. Tertiary-educated people are those who have graduated from the following stages:

- The first stage, which includes largely theory-based programmes to provide sufficient qualifications to gain entry to advanced research programmes and professions with high skills requirements and programmes which are practically, technically or occupationally specific (ISCED 5, 6 and 7).

- The second stage, which leads to the award of an advanced research qualification (e.g., PhD, non-PhD programmes with an advanced research component, etc.). The programmes are devoted to advanced study and original research (ISCED 8).

An activity is classified as '**knowledge-intensive'** if tertiary-educated people employed in this activity represent more than 33% of the total employment in the activity. The definition is based on the average number of employed persons aged 25–64 at aggregated EU-27 level.

The activities come from the NACE Rev. 2 categories (2-digit level),[13] based on EU Labour Force Survey (LFS) data. NACE refers to the European Community's statistical classification of economic activities.

In this indicator, there is one aggregate in use based on the following classification: Knowledge-intensive activities – Business industries (KIABI). The list of activities included in this aggregate, according to NACE Rev. 2 (2-digit level), is given below:

Table 2 Knowledge-intensive activities – Business industries (KIABI), NACE Rev. 2

Codes	Description
09	Mining support service activities
19	Manufacture of coke and refined petroleum products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
26	Manufacture of computer, electronic and optical products
51	Air transport

[13] Due to the revision of the NACE from NACE Rev. 1.1 to NACE Rev. 2, the definition of high-technology industries and knowledge-intensive services changed in 2008.

58	Publishing activities
59	Motion picture, video and television programme production, sound recording
60	Programming and broadcasting activities
61	Telecommunications
62	Computer programming, consultancy and related activities
63	Information service activities
64	Financial service activities, except insurance and pension funding
65	Insurance, reinsurance and pension funding, except compulsory social security
66	Activities auxiliary to financial services and insurance activities
69	Legal and accounting activities
70	Activities of head offices; management consultancy activities
71	Architectural and engineering activities; technical testing and analysis
72	Scientific research and development
73	Advertising and market research
74	Other professional, scientific and technical activities
75	Veterinary activities
78	Employment activities
79	Travel agency, tour operator reservation service and related activities
90	Creative, arts and entertainment activities

2.4. Eurostat – Research and development statistics

Content-based rationale

Research and development (R&D) is central to the EU's vision for growth, as demonstrated by the Horizon 2020 programme, the Horizon Europe programme, the Europe 2020 and Europe 2030 strategy. Although, on average, a higher proportion of women in the EU are completing degrees than ever before, there are signs that they continue to lag behind men when it comes to their representation amongst the researcher population. This situation persists across all sectors, particularly in the business enterprise sector (BES). She Figures indicators based on R&D statistics explore women's presence in R&D personnel and especially as researchers, broken down by sector, as well as by field of Research and Development and age group.

In addition, some indicators consider R&D expenditure, in order to provide an insight into whether there are any correlations between spending levels and other factors.

Broad overview of the source

These data can be accessed through the Research and Development (R&D) database on the Eurostat website, through the 'science and technology' tab here: http://ec.europa.eu/eurostat/data/database

Eurostat's Statistics on Research and Development provide data on R&D spending and R&D personnel working in the main sectors of the economy: the business enterprise (BES), government (GOV), higher education (HES), and the private non-profit (PNP) sectors. R&D personnel data can be viewed in full-time equivalent (FTE), in head count (HC), as a percentage of employment and as a percentage of the labour force. Amongst other things, the data are disaggregated by occupation, qualification, sex, citizenship, age group, fields of Research and Development and economic activity (NACE Rev. 2).

Data from Eurostat is publicly available, regularly updated and accompanied by extensive methodological notes. The data are collected through samples, census surveys or administrative registers – or through a combination of sources.

A complete list of indicators falling into this category can be found in Annex 3 and their detailed description follows below.

2.4.1. ***Proportion of women among researchers***

2.4.1.1. *Definition of indicator*

This indicator presents the proportion of female researchers, broken down by country, out of the researcher population in all sectors of the economy.

2.4.1.2. *Rationale*

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the research and innovation (R&I) workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society (EIGE, 2016). This indicator aims to shed light on whether there have been any improvements in the gender balance amongst this group.

2.4.1.3. *Computation method*

Data needed

(F) Number of female researchers in all sectors of the economy. **Unit: Head count/Full-time Equivalent**.

(T) Number of researchers in all sectors of the economy. **Unit: Head count/Full-time Equivalent**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment and sex)*

Computation formula

Proportion of women among researchers = F/T

2.4.1.4. *Specifications*

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

Full-time equivalent (FTE) of R&D personnel is defined as the ratio of working hours actually spent on R&D during a specific reference period (usually a calendar year) divided by the total number of hours conventionally worked in the same period by an individual or by a group. (§5.49, Frascati Manual, OECD, 2015).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

2.4.1.5. *Comments and critical issues*

For the countries that have data available in Eurostat, the indicator can be found calculated in the data code *rd_p_femres.*

The proportion of women among researchers is calculated both in Head Count (HC) and in Full-time Equivalent (FTE) but they are presented in different figures in the main publication. The proportion of women among researchers that is presented in different tables and figures of Chapter 4 of main publication is calculated in HC while the same proportion presented in Chapter 5 in combination with R&D Expenditure per capita researcher is calculated in FTE.

2.4.2. ***Compound annual growth rate (CAGR) of researchers, by sex***

2.4.2.1. *Definition of indicator*

This indicator compares the average annual percentage change in the proportion of women and men in the researcher population over a particular period.

2.4.2.2. *Rationale*

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the research and innovation (R&I) workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society (EIGE, 2016). This indicator aims to capture the relative changes in women's and men's participation in the researcher population.

2.4.2.3. *Computation method*

Data needed

(F) Number of female researchers in a start and an end year. **Unit: Head count**.

(M) Number of male researchers in a start and an end year. **Unit: Head count**.

(N) Number of years in reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number of years**.

Source of data

Eurostat – Statistics on research and *development (online data code: rd_p_persocc)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment and age)*

Computation formula

The CAGR shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of female researchers and male researchers each year in a given period. It is calculated in the following way:

CAGR of female researchers = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

CAGR of male researchers = $(\mathrm{M_e}/\mathrm{M_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

F_s denotes the number of female researchers in the start year;

F_e denotes the number of female researchers in the end year;

M_s denotes the number of male researchers in the start year;

M_e denotes the number of male researchers in the end year.

For example, if there were 200 female researchers in 2012 and 150 in 2016, the calculation would be:

CAGR of women researchers = $(150/200)^{1/4} - 1 = -6.9\,\%$.

2.4.2.4. *Specifications*

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

2.4.2.5. *Comments and critical issues*

Ensure the same reference period when calculating the CAGR for women and men respectively.

Numbers (HC) for male researchers are not given directly and need to be calculated as the difference of female researchers from total researchers

In the EU, men are more likely to be employed as researchers than women. For a reduction of the gender gap in employment rates, the CAGR needs to be higher for women than it is for men.

2.4.3. ***Researchers per thousand labour force, by sex***

2.4.3.1. *Definition of indicator*

This indicator presents the number of researchers for every thousand people in the labour force in a given country, broken down by sex.

2.4.3.2. *Rationale*

This indicator is another measure of the level of gender balance amongst the researcher population, given the historic tendency for this field to be dominated by men. Fostering equality in the representation of women and men amongst researchers demonstrates the EU's wider desire to 'reduce gender segregation at all levels in education and employment, as it contributes to inequalities in terms of the economic independence of women and men' (Council of the European Union, 2014).

2.4.3.3. *Computation method*

Data needed

(F) Number of female researchers. **Unit: Head count**.

(M) Number of male researchers. **Unit: Head count**.

($\mathrm{F_i}$) Number of women in the labour force (definition below), aged 15 and over. **Unit: Number**.

($\mathrm{M_i}$) Number of men in the labour force (definition below), aged 15 and over. **Unit: Number**.

Source of data

For F and M:

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

For $\mathrm{F_i}$ and $\mathrm{M_i}$:

Eurostat – Labour Force Survey *(online data code: lfsa_agan)*

Note that the numbers from the Labour Force Survey are in thousand units.

Computation formula

The formula for this indicator is:

Female researchers per thousand female labour force = $\mathrm{F} / \mathrm{F_i}$

Male researchers per thousand male labour force = $\mathrm{M} / \mathrm{M_i}$

2.4.3.4. *Specifications*

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

- According to the EU Labour Force Survey (LFS) **the labour force** (also termed 'active population') is defined as the sum of employed and unemployed persons.

Employed persons are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

Unemployed persons are 'all persons aged 15 to 74 who were not employed during the reference week, were available to start work within the two weeks following the reference week and had been actively seeking work in the four weeks preceding the reference week or had already found a job to start within the next three months'.

2.4.3.5. *Comments and critical issues*

F and *M* are in head count, whereas F_i and M_i are in thousand units. It is for this reason that the indicator states 'per thousand labour force'.

2.4.4. ***Proportion of women among researchers, by sector***

2.4.4.1. *Definition of indicator*

This indicator presents the proportion of female researchers in three broad economic sectors: the higher education sector (HES), the government sector (GOV) and the business enterprise sector (BES).

2.4.4.2. *Rationale*

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates (Eurostat, 2020a). Despite this, the EU's researcher population has continued to be dominated by men (DG Research and Innovation, 2019b). According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the research and innovation (R&I) workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society (EIGE, 2016). This indicator enables greater analysis by considering the situation for researchers within different sectors.

2.4.4.3. *Computation method*

Data needed

(F) Number of female researchers, in the higher education sector, the government sector and in the business enterprise sector. **Unit: Head count**.

(T) Total number of researchers, in the higher education sector, the government sector and in the business enterprise sector. **Unit: Head count**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment and sex)*

Computation formula

Proportion of women among researchers in a particular sector = F_i / T_i

where:

i denotes the sector (either HES, GOV or BES).

2.4.4.4. *Specifications*

The Frascati Manual (OECD, 2015) identifies and defines five **sectors of the economy**: the higher education sector (HES), the government sector (GOV), the business enterprise sector (BES), the private non-profit sector (PNP) and the Rest of the world. The definitions for the first three of these (included in this indicator) are:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market non-profit institutes (NPIs) that are controlled by government units that are not part of the Higher education sector'.

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e., units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

2.4.5. ***Distribution of researchers across institutional sectors, by sex***

2.4.5.1. *Definition of indicator*

This indicator presents the distribution of male and female researchers across four institutional sectors: the higher education sector (HES), the government sector (GOV), the business enterprise sector (BES) and the private non-profit sector (PNP).

2.4.5.2. *Rationale*

This indicator enables one to compare the sectors in which male and female researchers work. There are many reasons why this may be of interest, partly arising from the economic changes currently affecting much of the EU. In 2017, almost half (45.65 %) of the research population works in the business enterprise sector[14]. At the same time, the

[14] See Eurostat, total R&D personnel by sectors of performance, occupation and sex [*rd_p_persocc*].

share of R&D expenditure in the same sector is 66 % (Eurostat, 2020b). However, women have higher shares among researchers in the government or higher education sector (DG Research and Innovation, 2019b).

2.4.5.3. *Computation method*

Data needed

(F) The number of female researchers in each of the four economic sectors: the higher education sector (HES), the government sector (GOV), the business enterprise sector (BES) and the private non-profit sector (PNP). **Unit: Head count**.

(M) The number of male researchers in each of the four economic sectors: the higher education sector (HES), the government sector (GOV), the business enterprise sector (BES) and the private non-profit sector (PNP). **Unit: Head count**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment and sex)*

Computation formula

This indicator shows how researchers are spread out across different sectors, broken down by sex.

To compute this indicator, perform these two calculations for *each* sector:

Distribution of female researchers across sectors = F_i / F_a

Distribution of male researchers across sectors = M_i / M_a

where:

i denotes the sector (HES, GOV, BES or PNP);

a denotes all sectors;

F_i denotes the number of female researchers in a particular sector;

M_i denotes the number of male researchers in a particular sector;

F_a denotes the number of female researchers in all sectors;

M_a denotes the number of male researchers in all sectors.

For each sex, the proportions for the sectors are shown alongside one another (with a sum total of 100%).

For example, suppose there are 1 000 female researchers. Of these, 350 are in the HES, 224 are in the GOV sector, 326 are in the BES, and 100 are in the PNP. The proportion of female researchers in each sector would be as follows:

HES: 350 / 1000 = 35 %

GOV: 224 / 1000 = 22.4 %

BES: 326 / 1000 = 32.6 %

PNP: 100 / 1000 = 10 %

Sum total of 100 %.

2.4.5.4. *Specifications*

The Frascati Manual (OECD, 2015) identifies and defines **five sectors of the economy**: HES, GOV, BES, PNP and Rest of the world. The definitions for sectors included in this indicator are:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market NPIs that are controlled by government units that are not part of the Higher education sector'.

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e., units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

PNP (§3.75): 'The Private non-profit sector comprises all non-profit institutions serving households (NPISH), as defined in the System of National Accounts (SNA) 2008, except those classified as part of the Higher education sector. For completeness of presentation it comprises also, households and private individuals engaged or not engaged in market activities, as explained in the section 'Criteria for the classification of institutional sectors for R&D statistics' earlier in this chapter'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

2.4.6. *Distribution of researchers in the higher education sector (HES) across fields of Research and Development, by sex*

2.4.6.1. *Definition of indicator*

This indicator focuses on the higher education sector (HES) and presents the distribution of female and male researchers across the six major fields of Research and Development: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; humanities.

2.4.6.2. *Rationale*

Although women are more likely than men to have a higher education degree, they remain over-represented in fields of study that are linked to traditional female roles such as care-related fields and are under-represented in science, mathematics, IT, engineering and related careers. As a result, inequality in occupations is taking new forms and, despite their investment in education, young women are still twice as likely as young men to be economically inactive (European Commission, 2016).

Consistent with the indicators on PhD graduations, this indicator sheds light on the extent of gender segregation across different fields of R&D in the higher education sector (HES). It is particularly important to consider this sector, given that it is one of the main sources of employment for researchers in the EU. According to recent data (2017), a great proportion of the researchers (43.25 %) are employed in the higher education sector.[15]

2.4.6.3. *Computation method*

Data needed

(F) Number of female researchers in the Higher Education Sector (HES), in all fields of Research and Development. **Unit: Head count**.

(F_i) Number of female researchers in the Higher Education Sector (HES), in each field of Research and Development. **Unit: Head count**.

(M) Number of male researchers in the Higher Education Sector (HES), in all fields of Research and Development. **Unit: Head count**.

(M_i) Number of male researchers in the Higher Education Sector (HES), in each field of Research and Development. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment, field of R&D and sex)*

Computation formula

This indicator shows how researchers are spread out across different fields of Research and Development, broken down by sex.

[15] See Eurostat, total R&D personnel by sectors of performance, occupation and sex [rd_p_persocc].

To compute this indicator, perform these two calculations for each field of Research and Development in turn:

Distribution of female researchers across fields of R&D = $\mathrm{F_i} / F$

Distribution of male researchers across fields of R&D = $\mathrm{M_i} / M$

where:

i denotes a particular field of Research and Development;

$\mathrm{F_i}$ denotes the number of female researchers in the HES, in a given field of R&D;

$\mathrm{M_i}$ denotes the number of male researchers in the HES, in a given field of R&D.

For each sex, the proportions for the fields of Research and Development are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 1 000 female researchers in the HES. Of these, 150 are in natural sciences, 170 in engineering and technology, 200 in medical sciences, 82 in agricultural and veterinary sciences, 250 in social sciences, and 148 in humanities. The proportion of female researchers in the HES in each field of Research and Development would be as follows:

Natural sciences: 150 / 1000 = 15 %

Engineering and technology: 170 / 1000 = 17 %

Medical sciences: 200 / 1000 = 20 %

Agricultural and veterinary sciences: 82 / 1000 = 8.2 %

Social sciences: 250 / 1000 = 25 %

Humanities and arts: 148 / 1000 = 14.8 %

Sum total of 100 %.

2.4.6.4. *Specifications*

The Frascati Manual (OECD, 2015) provides definitions for the six **main fields of R&D** (Table 2.2, p. 59) that are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H).

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.6.5. *Comments and critical issues*

The breakdown of researchers by field of R&D is performed according to the field in which they work and not according to the field of their qualification.

2.4.7. ***Compound annual growth rate (CAGR) of female researchers in the higher education sector (HES), by field of Research and Development***

2.4.7.1. *Definition of indicator*

This indicator presents the compound annual growth rate of female researchers in the Higher education sector (HES) in six major fields of Research and Development: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; humanities.

2.4.7.2. *Rationale*

Although women are more likely than men to have a higher education degree, they remain over-represented in fields of study that are linked to traditional female roles such as care-related fields and are under-represented in science, mathematics, IT, engineering and related careers. As a result, inequality in occupations is taking new forms and, despite their investment in education, young women are still twice as likely as young men to be economically inactive (European Commission, 2016).

Consistent with the indicators on PhD graduations, this indicator sheds light on the extent of gender segregation across different fields of R&D in the higher education sector (HES). It is particularly important to consider this sector, given that it is the main source of employment for researchers in the EU. According to recent data (2017), a great proportion of the researchers (43.25 %) are employed in the higher education sector.[16]

2.4.7.3. *Computation method*

Data needed

(F) Number of female researchers in each of the fields of Research and Development in the higher education sector, in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment, field of R&D and sex)*

Computation formula

The CAGR shows the average rate of growth per year**,** for a given period. In this case, it shows the average percentage growth of female researchers in each main field of Research and Development in the higher education sector (HES) in a given period.

[16] See Eurostat, total R&D personnel by sectors of performance, occupation and sex [*rd_p_persocc*].

For each field of Research and Development respectively, perform this calculation:

CAGR of female researchers in each field of R&D = $(F_e/F_s)^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

F_s the number of female researchers in the chosen field of R&D (HES) in the start year;

F_e the number of female researchers in the chosen field of R&D (HES) in the end year.

2.4.7.4. *Specifications*

The Frascati Manual (OECD, 2015) provides definitions for the six **main fields of R&D**, which are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H).

The breakdown of researchers by field of Research and Development is performed according to the field in which they work and not according to the field of their qualification.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.7.5. *Comments and critical issues*

In fields of Research and Development where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance in that field.

2.4.8. ***Proportion of women among researchers, by main field of Research and Development (FORD) and by sector (HES, GOV and BES)***

2.4.8.1. *Definition of indicator*

This indicator presents the proportion of female researchers in each of the six fields of Research and Development: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; and humanities. It does so for the higher education sector (HES), the government sector (GOV) and the business enterprise sector (BES) in turn.

2.4.8.2. *Rationale*

The EU's commitment to tackling 'gender segregation at all levels in education and employment' encompasses the research fields in which women and men work (Council of the European Union, 2014). In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the research and innovation (R&I) workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society (EIGE, 2016). This indicator sheds light on the extent of gender segregation across different fields of research in the HES, GOV and BES sectors.

2.4.8.3. *Computation method*

Data needed

(F) Number of female researchers, broken down by sector (HES, GOV, BES) and field of Research and Development (FORD). **Unit: Head count**.

(T) Total number of researchers, broken down by sector (HES, GOV, BES) and field of Research and Development (FORD). **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment, field of R&D and sex)*

Computation formula

For each field of Research and Development, perform this calculation:

Proportion of women among researchers in a FORD in the HES = F_{hi} / T_{hi}

Proportion of women among researchers in a FORD in the GOV = F_{gi} / T_{gi}

Proportion of women among researchers in a FORD in the BES = F_{bi} / T_{bi}

where:

i denotes a particular field of R&D (FORD);

h denotes the higher education sector;

g denotes the government sector;

b denotes the business enterprise sector;

F_{hi} denotes the number of female researchers working in the HES in a particular field of R&D;

T_{hi} denotes the total number of researchers working in the HES in the same field of R&D as that in F_{hi} ;

F_{gi} denotes the number of female researchers working in GOV in a particular field of R&D;

T_{gi} denotes the total number of researchers working in GOV in the same field of R&D as that in F_{gi} ;

F_{bi} denotes the number of female researchers working in the BES in a particular field of R&D;

T_{bi} denotes the total number of researchers working in the BES in the same field of R&D as that in F_{bi} .

For example, in a particular sector, suppose there are 1 200 people working as researchers. Of these, 150 work in natural sciences (68 of them women), 245 work in engineering and technology (80 of them are women), 300 work in medical sciences (178 of them are women), 95 work in agricultural and veterinary sciences (34 of them are women), 140 work in social sciences (75 are women), and finally, 270 work in humanities (125 are women).

The proportion of women among researchers in each field of R&D is as follows:

natural sciences: 68 / 150 = 45.3 %

engineering and technology: 80 / 245 = 32.7 %

medical sciences: 178 / 300 = 59.3 %

agricultural and veterinary sciences: 34 / 95 = 35.8 %

social sciences: 75 / 140 = 53.6 %

humanities and arts: 125 / 270 = 46.3 %.

2.4.8.4. *Specifications*

The Frascati Manual (OECD, 2015) provides definitions for the six **main fields of R&D** (p. 59), which are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H).

The breakdown of researchers by field of Research and Development is according to the field in which they work and not according to the field of their qualification.

The Frascati Manual (OECD, 2015) identifies and defines **five sectors of the economy**: HES, GOV, BES, PNP and Rest of the world. The definitions for the first three of these (included in this indicator) are:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market NPIs that are controlled by government units that are not part of the Higher education sector'.

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e., units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.8.5. *Comments and critical issues*

In the body of the She Figures, this indicator is presented for two reference years in order to show the evolution of the proportion of female researchers in different fields and sectors (i.e., the extent of change over time).

2.4.9. ***Distribution of researchers in the government sector (GOV) across fields of Research and Development, by sex***

2.4.9.1. *Definition of indicator*

This indicator focuses on the government sector (GOV) and presents the distribution of male and female researchers across the six fields of Research and Development: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; and humanities.

2.4.9.2. *Rationale*

The EU is committed to reducing 'gender segregation at all levels in education and employment', which includes the research fields in which women and men work. Indicators on horizontal segregation tend to focus on the higher education sector. However, in 2017, the government sector employed about 10 % of researchers in the

EU,[17] making it another sector of interest when considering researchers' career patterns and the extent of horizontal segregation.[18]

2.4.9.3. *Computation method*

Data needed

(F) Number of female researchers in the government sector (GOV), in all fields of Research and Development. **Unit: Head count**.

(F_i) Number of female researchers in the government sector (GOV), in each field of Research and Development. **Unit: Head count**.

(M) Number of male researchers in the government sector (GOV), in all fields of Research and Development. **Unit: Head count**.

(M_i) Number of male researchers in the government sector (GOV), in each field of Research and Development. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment, field of R&D and sex)*

Computation formula

This indicator shows how researchers are spread out across different fields of Research and Development (FORD).

For each field of Research and Development, the formula for this indicator is:

Distribution of female researchers in GOV sector across FORD = F_i / F

Distribution of male researchers in GOV sector across FORD = M_i / M

where:

i refers to a particular field of R&D;

F_i denotes the number of female researchers in the GOV sector, in a given field of R&D;

M_i denotes the number of male researchers in the GOV sector, in a given field of R&D.

For each sex, the proportions for the fields of R&D are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 1 000 female researchers in the GOV sector. Of these, 150 are in natural sciences, 170 in engineering and technology, 200 in medical sciences, 82 in agricultural and veterinary sciences, 250 in social sciences, and 148 in humanities.

[17] See Eurostat, total R&D personnel by sectors of performance, occupation and sex [rd_p_persocc].

[18] For a definition of horizontal segregation, please refer to Annex 2.

The proportion of female researchers in the GOV sector in each field of R&D would be as follows:

natural sciences: 150 / 1000 = 15 %

engineering and technology: 170 / 1000 = 17 %

medical sciences: 200 / 1000 = 20 %

agricultural and veterinary sciences: 82 / 1000 = 8.2 %

social sciences: 250 / 1000 = 25 %

humanities and arts: 148 / 1000 = 14.8 %

sum total of 100 %.

2.4.9.4. *Specifications*

The Frascati Manual (OECD, 2015) provides definitions for the six **main fields of R&D** (p. 59), which are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.9.5. *Comments and critical issues*

The breakdown of researchers by field of Research and Development is performed according to the field in which they work and not according to the field of their qualification.

2.4.10. ***Compound annual growth rates (CAGR) of female researchers in the government sector (GOV) by field of Research and Development***

2.4.10.1. *Definition of indicator*

This indicator presents the compound annual growth rate of female researchers in the government sector across the six fields of Research and Development: natural sciences;

engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; and humanities.

2.4.10.2. *Rationale*

The EU is committed to reducing 'gender segregation at all levels in education and employment', which includes the research fields in which women and men work. Indicators on horizontal segregation tend to focus on the higher education sector. However, in 2017, the government sector employed about 10 % of researchers in the EU,[19] making it another sector of interest when considering researchers' career patterns and the extent of horizontal segregation.[20]

2.4.10.3. *Computation method*

Data needed

(F) Number of female researchers in the government sector, working in each of the fields of Research and Development, in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment, field of R&D and sex)*

Computation formula

The CAGR shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of female researchers in each main field of Research and Development in the government (GOV) sector, in a given period.

For each field of Research and Development respectively, the following calculation is performed:

CAGR of female researchers in each field of R&D = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

Fs is the number of female researchers (GOV) in the chosen field of Research and Development in the start year;

$\mathrm{F}e$ is the number of female researchers (GOV) in the chosen field of Research and Development in the end year.

[19] See Eurostat, total R&D personnel by sectors of performance, occupation and sex [rd_p_persocc].

[20] For a definition of horizontal segregation, please refer to Annex 2.

2.4.10.4. *Specifications*

The Frascati Manual (OECD, 2015) provides definitions for the six **main fields of R&D** (p. 59), which are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.10.5. *Comments* and critical issues

In areas where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance.

The breakdown of researchers by field of Research and Development is performed according to the field in which they work and not according to the field of their qualification.

2.4.11. ***Distribution of researchers across economic activities (NACE Rev. 2) in the business enterprise sector (BES), by sex***

2.4.11.1. *Definition of indicator*

This indicator presents the distribution of male and female researchers across specific economic activities in the business enterprise sector: manufacturing; services of the business economy; and all other economic activities.

2.4.11.2. *Rationale*

Previous editions of the She Figures (2006, 2009, 2012, 2015, 2018) have shown that, out of the three main sectors of the economy (HES, GOV and BES), female researchers are less represented in the business enterprise sector, making up less than a fifth of such employees. Whilst other She Figures indicators give a picture of women's overall representation in this sector, this indicator provides an insight into the economic activities being pursued by female and male researchers within the sector. According to a report by the European Foundation for the Improvement of Living and Working Conditions, segregation by gender in the labour market is far-reaching (Eurofound, 2013). Given this situation, it is worthwhile to investigate whether this also holds for individual economic activities within the business enterprise sector.

2.4.11.3. *Computation method*

Data needed

(F) Number of female researchers in the business enterprise sector (BES), in all economic activities (**Unit: Head count**), as well as:

Number of female researchers in the BES, in the economic activity 'Manufacturing'. **Unit: Head count**.

Number of female researchers in the BES, in the economic activity 'Services of the business economy'. **Unit: Head count**.

Number of female researchers in the BES, in all NACE economic activities except 'Manufacturing' and 'Services of the business economy'. **Unit: Head count**.

(M) Number of male researchers in the business enterprise sector (BES) in all economic activities (**Unit: Head count**), as well as:

Number of male researchers in the business enterprise sector (BES), in the economic activity 'Manufacturing'. **Unit: Head count**.

Number of male researchers in the business enterprise sector (BES), in the economic activity 'Services of the business economy'. **Unit: Head count**.

Number of male researchers in the business enterprise sector (BES), in all NACE Rev. 2 economic activities except 'Manufacturing' and 'Services of the business economy'. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_bempoccr2)*

Note that this data code from Eurostat already combines codes G–N as 'Services of the business economy', as well as some of the 'Other NACE codes'.

Computation formula

This indicator covers three types of economic activities in the NACE Rev. 2 classifications:

- 'Manufacturing' – Code C;
- 'Services of the business economy – Codes G–N combined;
- 'Other NACE codes' – Codes A, B, D–F, O–U.

The formula for this indicator is:

Distribution of female researchers (BES) across economic activities = Fi/F

Distribution of male researchers (BES) across economic activities = Mi/M

where:

i denotes a particular economic activity (for this indicator, either 'Manufacturing', 'Services of the business economy' or 'Other NACE codes');

F denotes the Number of female researchers in the BES;

M denotes the Number of male researchers in the BES;

Fi denotes the number of female researchers in the BES, in a given economic activity;

Mi denotes the number of male researchers in the BES, in a given economic activity.

For each sex, the proportions for the three types of economic activity are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 1 000 female researchers in the BES. Of these, 240 work in manufacturing, 340 in 'services of the business economy', and 420 in the remaining economic activities ('Other NACE codes'). The proportion of female researchers in the BES in each type of economic activity would be as follows:

manufacturing: 240 / 1000 = 24 %

services of the business economy: 340 / 1000 = 34 %

other NACE codes: 420 / 1000 = 42 %

sum total of 100 %.

2.4.11.4. *Specifications*

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

Researchers in the business enterprise sector are categorised using the Statistical **Classification of Economic Activities in the European Community, Rev. 2 (NACE Rev. 2)**. This has 21 main sections:

A Agriculture, forestry and fishing

B Mining and quarrying

C Manufacturing

D Electricity, gas, steam and air conditioning supply

E Water supply, sewerage, waste management and remediation activities

F Construction

G Wholesale and retail trade; repair of motor vehicles and motorcycles

H Transportation and storage

I Accommodation and food service activities

J Information and communication

K Financial and insurance activities

L Real estate activities

M Professional, scientific and technical activities

N Administrative and support service activities

O Public administration and defence; compulsory social security

P Education

Q Human health and social work activities

R Arts, entertainment and recreation

S Other service activities

T Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use

U Activities of extraterritorial organisations and bodies.

For a full listing of the NACE Rev. 2 categories (including divisions and groups), please see:
http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=NACE_REV2

NACE Rev. 2 was adopted in December 2006, building on NACE Rev. 1.1. In general, the updated version of NACE is used in statistics on economic activities from 1 January 2008 onwards.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.12. ***Proportion of women among researchers in the business enterprise sector (BES), by economic activity (NACE)***

2.4.12.1. *Definition of indicator*

This indicator allows comparison of the proportion of female researchers across five different economic activities in the business enterprise sector: manufacturing; manufacturing of chemicals and chemical products; manufacturing of basic pharmaceutical products and pharmaceutical preparations; services of the business economy; other NACE codes.

2.4.12.2. *Rationale*

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the research and innovation (R&I) workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society (EIGE, 2016).

Previous editions of the She Figures (2006, 2009, 2012, 2015, 2018) have shown that, of the three main sectors of the economy (HES, GOV and BES), female researchers are worst represented in the business enterprise sector, making up less than a fifth of such

employees. By considering individual economic activities, this indicator enables one to assess if this picture also holds in key sections of the sector.

2.4.12.3. *Computation method*

Data needed

(F) Number of female researchers in the business enterprise sector (BES), in each of the following activities/divisions:

Economic activity, 'Manufacturing'. **Unit: Head count**.

Division 'Manufacturing of chemicals and chemical products' of the economic activity, 'Manufacturing'. **Unit: Head count**.

Division 'Manufacture of basic pharmaceutical products and pharmaceutical preparations' of the economic activity, 'Manufacturing'. **Unit: Head count**.

Economic activity, 'Services of the business economy'. **Unit: Head count**.

All NACE economic activities, except 'Manufacturing' and 'Services of the business economy'. **Unit: Head count**.

(T) Number of researchers in the business enterprise sector (BES), in each of the following activities/divisions.

Economic activity, 'Manufacturing'. **Unit: Head count**.

Division 'Manufacturing of chemicals and chemical products' of the economic activity, 'Manufacturing'. **Unit: Head count**.

Division 'Manufacture of basic pharmaceutical products and pharmaceutical preparations' of the economic activity, 'Manufacturing'. **Unit: Head count**.

Economic activity, 'Services of the business economy'. **Unit: Head count**.

All NACE economic activities, except 'Manufacturing' and 'Services of the business economy'. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_bempoccr2)*

Note that this data code from Eurostat (*rd_p_bempoccr2*) already combines codes G–N as 'Services of the business economy', as well as some of the 'Other NACE codes'.

Computation formula

This indicator covers five types of economic activities/divisions in the NACE Rev. 2 classifications:

- 'Manufacturing' – Code C
- 'Manufacturing of chemicals and chemical products'- Code C20
- 'Manufacture of basic pharmaceutical products and pharmaceutical preparations' – Code C21

- 'Services of the business economy – Codes G–N combined
- 'Other NACE codes' – Codes A, B, D–F, O–U.

Applied to each activity/division in turn, the formula for this indicator is:

Proportion of women among researchers in a given economic activity = Fi/Ti

where:

i denotes a particular economic activity (one of the five covered by this indicator);

Fi denotes the number of female researchers in the BES in a particular economic activity;

Ti denotes the total number of researchers in the BES, in a given economic activity.

Note: Ensure that the economic activity covered by Fi and Ti is the same.

2.4.12.4. *Specifications*

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

Researchers in the business enterprise sector are categorised using the Statistical **Classification of Economic Activities in the European Community, Rev. 2 (NACE Rev. 2)**. This has 21 main sections. For this indicator, the most relevant sections are:

C Manufacturing, which includes two divisions:

C20: Manufacture of chemicals and chemical products;

C21: Manufacture of basic pharmaceutical products and pharmaceutical preparations.

G Wholesale and retail trade; repair of motor vehicles and motorcycles

H Transportation and storage

I Accommodation and food service activities

J Information and communication

K Financial and insurance activities

L Real estate activities

M Professional, scientific and technical activities

N Administrative and support service activities.

The remaining sections (covered in 'Other NACE codes') are:

A Agriculture, forestry and fishing

B Mining and quarrying

D Electricity, gas, steam and air conditioning supply

E Water supply, sewerage, waste management and remediation activities

F Construction

O Public administration and defence; compulsory social security

P Education

Q Human health and social work activities

R Arts, entertainment and recreation

S Other service activities

T Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use

U Activities of extraterritorial organisations and bodies.

For a full listing of the NACE Rev. 2 categories (including divisions and groups), please see:
http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=NACE_REV2

NACE Rev. 2 was adopted in December 2006, building on NACE Rev. 1.1. In general, the updated version of NACE is used in statistics on economic activities from 1 January 2008 onwards.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.13. *Compound annual growth rate (CAGR) of researchers in the higher education sector (HES), by sex*

2.4.13.1. *Definition of indicator*

This indicator compares the average annual rate of growth in women and men's employment as researchers in the higher education sector over a particular period.

2.4.13.2. *Rationale*

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the research and innovation (R&I) workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society (EIGE, 2016).

This indicator enables one to gauge changes in the patterns of women and men's employment as researchers over time, in the higher education sector (HES). Through comparing these results with those of the equivalent indicators for the government (GOV)

sector and business enterprise sector (BES), it is also possible to consider whether increases/decreases in one sector are offset by those in another.

2.4.13.3. *Computation method*

Data needed

(F) Number of female researchers in the higher education sector in a start and an end year. **Unit: Head count**.

(M) Number of male researchers in the higher education sector in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment and sex)*

Computation formula

The CAGR shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of female researchers and male researchers in the higher education sector (HES) each year in a given period.

It is calculated in the following way:

CAGR of female researchers in the HES = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

CAGR of male researchers in the HES = $(\mathrm{M_e}/\mathrm{M_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

Fs denotes the number of female researchers in the HES in the start year;

Fe denotes the number of female researchers in the HES in the end year;

$\mathrm{M_s}$ denotes the number of male researchers in the HES in the start year;

$\mathrm{M_e}$ denotes the number of male researchers in the HES in the end year.

For example, if there were 100 female researchers in the HES in 2012, and 150 in 2016, the calculation would be:

CAGR of female researchers = $(150/100)^{1/4} - 1 = 10.7\,\%$.

2.4.13.4. *Specifications*

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

The Frascati Manual (OECD, 2015) identifies and defines **five sectors of the economy**: HES, GOV, BES, PNP and Rest of the world. The definitions for the HES sector included in this indicator is:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.13.5. *Comments and critical issues*

In areas where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance.

Ensure the same reference period when calculating the CAGR for women and men respectively.

2.4.14. ***Compound annual growth rate (CAGR) of researchers in the government sector (GOV), by sex***

2.4.14.1. *Definition of indicator*

This indicator compares the average annual rate of growth in women and men's employment as researchers in the government sector, over a particular period.

2.4.14.2. *Rationale*

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the research and innovation (R&I) workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society (EIGE, 2016).

This indicator enables one to gauge changes in the patterns of women and men's employment as researchers over time, in the government sector (GOV). Through comparing these results with those of the equivalent indicators for the higher education sector (HES) and business enterprise sector (BES), it is also possible to consider whether increases/decreases in one sector are offset by those in another.

2.4.14.3. *Computation method*

<u>Data needed</u>

(F) Number of female researchers in the government sector in a start and an end year. **Unit: Head count**.

(M) Number of male researchers in the government sector in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment and sex)*

Computation formula

The compound annual growth rate (CAGR) shows the average rate of growth per year**,** for a given period. In this case, it shows the average percentage growth of female researchers and male researchers in the government sector (GOV) each year in a given period.

It is calculated in the following way:

CAGR of female researchers in the GOV = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

CAGR of male researchers in the GOV = $(\mathrm{M_e}/\mathrm{M_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

Fs denotes the number of female researchers in the GOV in the start year;

Fe denotes the number of female researchers in the GOV in the end year;

Ms denotes the number of male researchers in the GOV in the start year;

Me denotes the number of male researchers in the GOV in the end year.

For example, if there were 100 female researchers in the GOV in 2012, and 150 in 2016, the calculation would be:

CAGR of female researchers = $(150/100)^{1/4} - 1 = 10.7\ \%$.

2.4.14.4. *Specifications*

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

The Frascati Manual (OECD, 2015) identifies and defines **five sectors of the economy**: HES, GOV, BES, PNP and Rest of the world. The definitions for the GOV sector included in this indicator is:

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market NPIs that are controlled by government units that are not part of the Higher education sector'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.14.5. *Comments and critical issues*

In areas where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance.

Ensure the same reference period when calculating the CAGR for women and men respectively.

2.4.15. ***Compound annual growth rate (CAGR) of researchers in the business enterprise sector (BES), by sex***

2.4.15.1. *Definition of indicator*

This indicator compares the average annual rate of growth in women and men's employment as researchers in the business enterprise sector, over a particular period.

2.4.15.2. *Rationale*

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the research and innovation (R&I) workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society (EIGE, 2016).

This indicator enables one to gauge changes in the patterns of women and men's employment as researchers over time, in the business enterprise sector (BES). Through comparing these results with those of the equivalent indicators for the higher education sector (HES) and government sector (GOV), it is also possible to consider whether increases/decreases in one sector are offset by those in another.

2.4.15.3. *Computation method*

<u>Data needed</u>

(F) Number of female researchers in the business enterprise sector in a start and an end year. **Unit: Head count**.

(M) Number of male researchers in the business enterprise sector in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment and sex)*

Computation formula

The CAGR shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of female researchers and male researchers in the business enterprise sector (BES) each year in a given period.

It is calculated in the following way:

CAGR of female researchers in the BES = $(F_e/F_s)^{1/N} - 1$

CAGR of male researchers in the BES = $(M_e/M_s)^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

F_s denotes the number of female researchers in the BES in the start year;

F_e denotes the number of female researchers in the BES in the end year;

M_s denotes the number of male researchers in the BES in the start year;

M_e denotes the number of male researchers in the BES in the end year.

For example, if there were 100 female researchers in the BES in 2012, and 150 in 2016, the calculation would be:

CAGR of female researchers = $(150/100)^{1/4} - 1 = 10.7\ \%$

2.4.15.4. *Specifications*

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

The Frascati Manual (OECD, 2015) identifies and defines **five sectors of the economy**: HES, GOV, BES, PNP and Rest of the world. The definitions for BES sector included in this indicator is:

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e., units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.15.5. *Comments and critical issues*

In areas where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance.

Ensure the same reference period when calculating the CAGR for women and men respectively.

2.4.16. ***Distribution of researchers in the higher education sector (HES) across age groups, by sex***

2.4.16.1. *Definition of indicator*

This indicator shows the distribution of both male and female researchers in the higher education sector (HES) across different age groups.

2.4.16.2. *Rationale*

This indicator focuses on the higher education sector (HES), and can be compared with the results of the equivalent indicator for the government sector (GOV).

Considering the age distribution of researchers, it may reveal differences in the career patterns of women and men. For example, according to Eurostat, a higher proportion of women are outside of the labour force due to caring responsibilities, including for children.[21] This may reduce their participation in the labour market during the key childbearing years of a particular country. On another level, by taking older age as a 'proxy' for seniority, this indicator can be used to gauge women and men's relative presence in the top research positions, against a backdrop of far-reaching under-representation of women in decision-making roles (EIGE's Gender Statistics Database).

2.4.16.3. *Computation method*

Data needed

(F) Number of female researchers in the higher education sector (HES) aged 25 and over. **Unit: Head count**.

$(\mathrm{F_i})$ Number of female researchers in the higher education sector (HES), in each of these age categories: 25–34; 35–44; 45–54; 55 and over. **Unit: Head count**.

(M) Number of male researchers in the higher education sector (HES) aged 25 and over. **Unit: Head count**.

$(\mathrm{M_i})$ Number of male researchers in the higher education sector (HES), in each of these age categories: 25–34; 35–44; 45–54; 55 and over. **Unit: Head count**.

[21] In 2019, in the EU, 37.3 % of women (aged 25 to 49) who were outside of the labour force were in the position due to looking after children or incapacitated adults. For men of the same age group outside of the labour force, the rate was 3.9 %. See Eurostat, 'Inactive Population – Main reason for not seeking employment – Distributions by sex and age (%)', data table *lfsa_igar*.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persage)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment, age and sex)*

Computation formula

The formula for this indicator is:

Distribution of female researchers across age groups = $\mathrm{F_i}/F$

Distribution of male researchers across age groups = $\mathrm{M_i}/M$

where:

i denotes a particular age group;

$\mathrm{F_i}$ denotes the number of female researchers in the HES, in a given age group;

$\mathrm{M_i}$ denotes the number of male researchers in the HES, in a given age group.

For each sex, the proportions for the age groups are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 100 male researchers (aged 25 and over) in the HES in one country. Of these, 12 are aged 25–34; 26 aged 35–44; 38 aged 45–54; and 24 aged 55 and over. The proportion of men in each age group would be as follows:

aged 25-34: 12 / 100 = 12 %

aged 35-44: 26 / 100 = 26 %

aged 45-54: 38 / 100 = 38 %

aged 55 and over: 24 / 100 = 24 %

sum total of 100 %.

2.4.16.4. *Specifications*

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

The Frascati Manual (OECD, 2015) identifies and defines **five sectors of the economy**: HES, GOV, BES, PNP and Rest of the world. The definitions for the HES sector included in this indicator is:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.17. *Distribution of researchers in the government sector (GOV) across age groups, by sex*

2.4.17.1. *Definition of indicator*

This indicator shows the distribution of both male and female researchers in the government sector (GOV) across different age groups.

2.4.17.2. *Rationale*

This indicator focuses on the government sector (GOV), and can be compared with the results of the equivalent indicator for the higher education sector (HES).

Considering the age distribution of researchers may reveal differences in the career patterns of women and men. For example, according to Eurostat, a higher proportion of women are outside of the labour force due to caring responsibilities, including for children.[22] This may reduce their participation in the labour market during the key childbearing years of a particular country. On another level, by taking older age as a 'proxy' for seniority, this indicator can be used to gauge women and men's relative presence in the top research positions, against a backdrop of far-reaching under-representation of women in decision-making roles (EIGE's Gender Statistics Database).

2.4.17.3. *Computation method*

Data needed

(F) Number of female researchers in the GOV sector aged 25 and over. **Unit: Head count**.

$(\mathrm{F_i})$ Number of female researchers in the GOV sector, in each of these age categories: 25–34; 35–44; 45–54; 55 and over. **Unit: Head count**.

(M) Number of male researchers in the GOV sector, aged 25 and over. **Unit: Head count**.

$(\mathrm{M_i})$ Number of male researchers in the GOV sector, in each of these age categories: 25–34; 35–44; 45–54; 55 and over. **Unit: Head count**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persage)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment,age and sex)*

Computation formula

The formula for this indicator is:

[22] Ibid.

Distribution of female researchers across age groups = $\mathrm{F_i}/F$

Distribution of male researchers across age groups = $\mathrm{M_i}/M$

where:

i denotes a particular age group;

$\mathrm{F_i}$ denotes the number of female researchers in the GOV sector, in a given age group;

$\mathrm{M_i}$ denotes the number of male researchers in the GOV sector, in a given age group.

For each sex, the proportions for the age groups are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 100 male researchers (aged 25 and over) in the GOV sector in one country. Of these, 12 are aged 25–34; 26 aged 35–44; 38 aged 45–54; and 24 aged 55 and over. The proportion of men in each age group would be as follows:

aged 25-34: 12 / 100 = 12 %

aged 35-44: 26 / 100 = 26 %

aged 45-54: 38 / 100 = 38 %

aged 55 and over: 24 / 100 = 24 %

sum total of 100 %.

2.4.17.4. *Specifications*

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

The Frascati Manual (OECD, 2015) identifies and defines **five sectors of the economy**: HES, GOV, BES, PNP and Rest of the world. The definitions for the GOV sector included in this indicator is:

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market NPIs that are controlled by government units that are not part of the Higher education sector'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.18. *Dissimilarity Index for researchers in the higher education sector (HES) and government sector (GOV)*

2.4.18.1. *Definition of indicator*

The Dissimilarity Index (DI) provides a theoretical measurement of the percentage of women and men in a field of R&D who would have to move to another field of R&D to ensure that the proportions of women were the same across all the possible fields of R&D. It can therefore be interpreted as the hypothetical distance from a balanced sex distribution across fields of R&D, based upon the overriding proportion of women (National Science Foundation, 2000).

2.4.18.2. *Rationale*

Although women are more likely than men to have a higher education degree, they remain over-represented in fields of study that are linked to traditional female roles such as care-related fields and are under-represented in science, mathematics, IT, engineering and related careers (European Commission, 2016). This indicator shows the proportion of one sex or all employees that would need to change field of R&D in order to achieve a gender balance across those fields.

2.4.18.3. *Computation method*

Data needed

(F) Number of female researchers across all fields of R&D. **Unit: Head count**.

(Fi) Number of female researchers in each field of R&D. **Unit: Head count**.

(M) Number of male researchers across all fields of R&D. **Unit: Head count**.

(Mi) Number of male researchers in each field of R&D. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

UIS data centre *(http://data.uis.unesco.org; Researchers by sector of employment, field of R&D and sex)*

Computation formula

This table presents the values of the Dissimilarity Index (DI) in the different countries for researchers in two sectors: higher education and government. Seven fields were considered in computing the DI: natural sciences; engineering and technology; medical and health sciences; agricultural and veterinary sciences; social sciences; humanities; and any other field of Research and Development. The full calculation method is explained under 'Specifications' below.

The formula for the Dissimilarity Index is:

$$DI = 1/2 \sum i \mid Fi / F - Mi / M \mid$$

where:

i denotes a particular R&D field.

For example, if we have three fields, A, B and C, with 17, 37 and 91 women, and 108, 74, 182 men respectively, the overall proportion of women is 28.5 %. We therefore need to calculate:

$$\frac{\left|\frac{17}{145}-\frac{108}{364}\right|+\left|\frac{37}{145}-\frac{74}{364}\right|+\left|\frac{91}{145}-\frac{182}{364}\right|}{2}=\frac{0.1795+0.0519+0.1276}{2}=0.1795$$

This means that 18 % of researchers will have to change field in order to maintain the background proportion of 28.5 % women in each field.

2.4.18.4. *Specifications*

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.18.5. *Comments and critical issues*

In order to interpret the DI correctly, it is important to know which sex is in the majority overall. The maximum value is 1, which indicates the presence of either women or men only in each of the occupations, depending on the majority gender. The minimum value of 0 indicates a distribution between women and men within each occupation which is equal to the overall average proportion of women. If the same occupational categories are used for different countries, the DI yields a comparable and descriptive statistic that reflects the extent to which the two sexes are differently distributed. The results also depend on the number of categories. If more categories are used, the indicator will reflect greater variability in the distribution, which in turn will yield results indicating a higher level of segregation.

The index shown in the She Figures is the Duncan and Duncan Index of Dissimilarity, first developed in the 1950s and now used extensively for international comparisons of inequality and dissimilarity (not solely between the sexes but also between other groups).

2.4.19. ***Distribution of R&D personnel across occupations, by sector of the economy and sex***

2.4.19.1. *Definition of indicator*

This indicator presents the distribution of research and development (R&D) personnel across three occupations (researchers, technicians, and other supporting staff), by sex in the three main sectors of the economy: higher education sector (HES), government (GOV) sector and business enterprise sector (BES).

2.4.19.2. *Rationale*

This indicator focuses on R&D personnel across all three sectors, namely the higher education sector, the government sector and the business enterprise sector. Since this indicator corrects for the total number of personnel for each sex, it allows for a comparison of the presence of each sex across the different occupations.

2.4.19.3. *Computation method*

Data needed

$(M_{s,i})$ Number of men in a given R&D occupation and sector. **Unit: Head count**.

$(F_{s,i})$ Number of women in a given R&D occupation and sector. **Unit: Head count**.

(M_s) Number of men in all R&D occupations in a given sector. **Unit: Head count**.

(F_s) Number of women in all R&D occupations in a given sector. **Unit: Head count**.

(i) Denotes a particular R&D occupation:

Researchers

Technicians

Other supporting staff

Technicians and other supporting staff (in cases where technicians and other supporting staff are not provided separately, see 2.4.19.5).

(s) Denotes a sector of activity:

higher education sector (HES)

government sector (GOV)

business enterprise sector (BES)

total of all sectors (HES, GOV and BES).

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

UIS data centre *(http://data.uis.unesco.org; R&D personnel by function and sector of employment)*

Computation formula

This indicator presents the relative proportion of personnel per occupation by sex.

Distribution of female personnel across occupations by sector = $\mathrm{F_{s,i}}/\mathrm{F_s}$

Distribution of female personnel across occupations by sector = $\mathrm{M_{s,i}}/\mathrm{M_s}$

For each sex, the proportions for the occupations are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 1 000 female R&D personnel in these three sectors. Of these, 390 work as researchers, 260 work as technicians and 350 work as other supporting staff. The proportion of women in each occupation would be as follows:

researchers: 390 / 1000 = 39 %

technicians: 260 / 1000 = 26 %

other supporting staff: 350 / 1000 = 35 %

sum total of 100 %.

2.4.19.4. *Specifications*

The Frascati Manual (OECD, 2015) provides an international definition for **R&D personnel**, §5.6: 'All persons employed directly on R&D should be counted, as well as those providing direct services such as R&D managers, administrators, and clerical staff.' R&D personnel comprise three categories of occupations:

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

Technicians (and equivalent staff) are persons whose main tasks require technical knowledge and experience in one or more fields of engineering, physical and life sciences or social sciences and humanities. They participate in R&D by performing scientific and technical tasks involving the application of concepts and operational methods, normally under the supervision of researchers. Equivalent staff performs the corresponding R&D tasks under the supervision of researchers in the social sciences and humanities (§5.40, Frascati Manual, OECD, 2015).

Other supporting staff includes skilled and unskilled craftsmen, secretarial and clerical staff participating in R&D projects or directly associated with such projects (§5.43, Frascati Manual, OECD, 2015).

The Frascati Manual (OECD, 2015) identifies and defines **five sectors of the economy**: HES, GOV, BES, PNP and Rest of the world. The definitions for the first three of these (included in this indicator) are:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market NPIs that are controlled by government units that are not part of the Higher education sector'.

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e., units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.4.19.5. *Comments and critical issues*

From the reference year 2012 onwards, it is not compulsory for countries to report technicians separately from other supporting staff when providing data for their R&D personnel (European Parliament, 2012). Therefore, distribution of R&D personnel across occupations is presented for the categories each country provides. In cases where data for technicians and other supporting staff are not provided, in any way, these are calculated as the subtraction of researchers from the total R&D personnel.

2.4.20. ***Total intramural R&D expenditure in purchasing power standards (PPS) per capita researcher in FTE, by sector of the economy***

2.4.20.1. *Definition of indicator*

This indicator breaks down R&D expenditure per capita researcher in FTE by sector of economy (business enterprise, government, higher education or private non-profit) for a given year. To account for differences in prices, currency and exchange rates, the data are expressed in purchasing power standards (PPS).

2.4.20.2. *Rationale*

Although it does not provide any gender-specific information, the indicator should be viewed in conjunction with the indicator that addresses the distribution of researchers across sectors by sex, in order to see if there is any correlation between R&D spending and female researchers' presence.

2.4.20.3. *Computational method*

Data needed

(T_i) The overall number of researchers, in full-time equivalent (FTE), by sector of the economy (HES, GOV, BES, PNP, Total). **Unit: Full-Time Equivalent**.

(E_i) R&D expenditure in millions of purchasing power standards (PPS), by sector of the economy (HES, GOV, BES, PNP, Total). **Unit: Million PPS**.

(i) Denotes a particular sector:

higher education sector (HES)

government sector (GOV)

business enterprise sector (BES)

private non-profit (PNP)

sum of the sectors (HES+GOV+BES+PNP).

Source of data

For T_i: Eurostat – Statistics on research and development *(online data: rd_p_persocc)*

For E_i: Eurostat – Statistics on research and development *(online data: rd_e_gerdtot)*

Computation formula

R&D expenditure in PPS per capita researcher (in a given sector) = $(E_i \times 1{,}000{,}000)/T_i$

2.4.20.4. *Specifications*

The definition of the **full-time equivalent (FTE)** unit of measurement of personnel employed in R&D, as proposed by the Frascati Manual corresponds to one year's work by one person on R&D.

The Frascati Manual defines intramural expenditures on **R&D as all expenditures for R&D performed** within a statistical unit or sector of the economy during a specific period, whatever the source of funds. It recommends using purchasing power parities (PPP) to express R&D statistics in monetary terms.

The **PPPs** are currency conversion rates that convert to a common currency and equalise the purchasing power of different currencies. They eliminate the differences in price levels between countries because economic indicators expressed in a national currency are converted into an artificial common currency, called the purchasing power standards (PPS).

2.5. MORE Survey

Content-based rationale

Directive 2006/54/EC of 5 July 2006 lays down the principle of equal treatment of men and women in the EU, in relation to their working conditions, access to promotion and access to occupational security schemes. Amongst other things, the Mobility and Career Paths of Researchers in Europe (MORE) Survey investigates gender differences in the working conditions of female and male researchers working in higher education institutions, including their contractual status and level of mobility. She Figures indicators from this data source include sex differences in mobility; part-time employment of researchers in the higher education sector (HES), by sex; and precarious working contracts of researchers in HES out of total researcher population, by sex.

Broad overview of the source

The MORE Surveys are part of the Mobility and Career Paths of Researchers in Europe (MORE) Project (European Commission, 'MORE', 'MORE2', 'MORE3' and 'MORE4'), funded by the European Commission. The project was set up to improve understanding of research careers in Europe. The survey is an important source of data on the research profession, including researchers' career progression and ability to move/work between different countries (European Commission, 2017b).

To date, there have been four editions of the MORE Survey: the MORE1 Survey (2009 data), the MORE2 Survey (2012 data; released in 2013), MORE3 Survey (2016 data; released in 2017) and MORE4 Survey (2019 data; data not released yet). As discussed under individual indicators, the results of these surveys are not fully comparable due to some differences in the questionnaire design.

Although the MORE Surveys produced questionnaires for researchers in multiple sectors, the She Figures uses only the survey of higher education institutions (HEIs). This is because the HEI survey contained the most relevant questions on the contractual status and mobility of researchers (European Commission, 2017c).

A complete list of indicators falling into this category can be found in Annex 3 and their detailed description follows below.

2.5.1. ***Sex differences in international mobility of researchers during their PhD***

2.5.1.1. *Definition of indicator*

The indicators show the difference in the percentage of female / male researchers who – during their PhD – moved for at least three months to a country other than that where they attained (or will attain) their PhD. It covers researchers in the early stages of their careers (R1 and R2).

2.5.1.2. *Rationale*

One of the main European Commission's plans for the new European Research Area is to strengthen the mobility of researchers and the flow of knowledge (European Commission, 2020c). However, there are some concerns that women may be less mobile than men at certain stages of their life. For example, there are signs that women continue to bear the

majority of childbearing responsibilities in the EU[23] and – as the European Parliament warns – mobility 'can be difficult to reconcile with family life'. According to the EU funded Gendered Innovations Expert Group report, 'gender roles that limit women's mobility interfere with careers in science and engineering' (European Commission, 2020d).

This indicator aims to identify if there are indeed such differences in the mobility of women and men, focusing on researchers' experiences of mobility during their PhDs.

2.5.1.3. *Computation method*

Data needed

$(\mathrm{F_i})$ Number of female researchers (R1 and R2 career stages) who – during their PhD – moved for at least three months to a country other than that where they attained (or will attain) their PhD. **Unit: Number**.

(F) Number of female researchers (R1 and R2 career stages). **Unit: Number**.

$(\mathrm{M_i})$ Number of male researchers (R1 and R2 career stages) who – during their PhD – moved for at least three months to a country other than that where they attained (or will attain) their PhD. **Unit: Number**.

(M) Number of male researchers (R1 and R2 career stages). **Unit: Number**.

(S) Sampling weights for individual survey results, by country and field of Research and Development.

Source of data

European Commission - MORE Survey on mobility patterns and career paths of researchers (https://www.more-4.eu/). Custom extraction of MORE4 Survey data.

Computation formula

This indicator presents the percentage point difference in the percentage of female / male researchers (R1 and R2) who were 'internationally mobile' during their PhDs (using the definition provided in this indicator). It is calculated by subtracting women's rate of mobility from men's rate.

Before calculating this indicator, one must weigh the survey results to increase their representativeness. Pre-calculated sampling weights (by country and field of Research and Development) were included in the MORE4 dataset.

Following the weighting phase, calculate the indicator as normal, using the weighted numbers.

Using the weighted values, perform these calculations:

Percentage of internationally mobile female researchers (R1 and R2) = $\mathrm{F_{iw}} / \mathrm{F_w}$

[23] For example, the gap between the EU employment rate of men and women widens with the arrival of dependent children. See Eurostat, 'Employment rate of adults by sex, age groups, highest level of education attained, number of children and age of youngest child (%)' [lfst_hheredch].

Percentage of internationally mobile male researchers (R1 and R2) = $\mathrm{M_{iw}} / \mathrm{M_{w}}$

Differences between genders in international mobility during PhD

$= (\mathrm{M_{iw}} / \mathrm{M_{w}}) - (\mathrm{F_{iw}} / \mathrm{F_{w}})$

where:

i denotes international mobility (using the definition provided for this indicator);

w denotes that the values are weighted.

For example, F_{iw} refers to the number of internationally mobile female researchers (R1 and R2) (weighted, using the definition described above), whilst M_w refers to the total number of male researchers (R1 and R2) (weighted).

2.5.1.4. *Specifications*

This indicator focuses on researchers in the higher education sector only. The MORE4 Survey applies the Frascati Manual definition of researchers.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

In this indicator, **researchers who are 'internationally mobile'** are defined as those who have moved abroad for at least three months during their PhD to a country other than the one where they completed (or will obtain) their PhD. It is based on a direct question in the MORE4 Survey of Higher Education Institutions.[24]

The survey asks researchers to classify their career stage, using the categories defined in the European Framework for Research Careers (DG Research and Innovation, 2011):

- R1: First Stage Researcher (up to the point of PhD)
- R2: Recognised Researcher (PhD holders or equivalent who are not yet fully independent)
- R3: Established Researcher (researchers who have developed a level of independence)
- R4: Leading Researcher (researchers leading their research area or field).

This indicator focuses on those who classified themselves as being in the R1 and R2.

[24] Changes to the MORE Surveys affect the comparability across different editions of the She Figures. In She Figures 2012, mobile researchers were defined as researchers who, in the last three years, had moved from the country where they had obtained their highest level of qualification to work as a researcher for at least three months in another country. The definition did not distinguish between career stages. In contrast to the 2012 edition, this indicator has been separated into two indicators, one focusing on mobility during their PhD – for researchers in the early career stages (R1 and R2 combined) – and another focusing on mobility in the last 10 years in the post-PhD phases (R2–R4).

2.5.1.5. *Comments and critical issues*

- There are large differences in the distribution of researchers across different career stages, depending on the country.

- The selected method for the survey results is direct weighting, so that the multiplication of each variable with the weighting coefficient returns the reference population. Here, sampling weights were calculated using the estimated number of researchers in each field of Research and Development in a given country (2019 data). Note that the sampling error can be higher at subpopulation level (including for gender).

2.5.2. ***Sex differences in international mobility in post-PhD career stages***

2.5.2.1. *Definition of indicator*

The indicators present the percentage point difference in the proportion of female/male researchers who – in the last 10 years – have worked abroad for at least three months in a country other than the country where they attained their highest educational degree. It focuses on researchers in the post-PhD phases of their careers (phases R2–R4).

2.5.2.2. *Rationale*

One of the main European Commission's plans for the new European Research Area is to strengthen the mobility of researchers and the flow of knowledge (European Commission, 2020c). However, there are some concerns that women may be less mobile than men at certain stages of their life. For example, there are signs that women continue to bear the majority of childbearing responsibilities in the EU[25] and – as the European Parliament warns – mobility 'can be difficult to reconcile with family life'.[26] According to the EU funded project Gendered Innovations, 'gender roles that limit women's mobility interfere with careers in science and engineering' (European Commission, 2020d).

This indicator aims to identify if there are indeed such differences in the mobility of women and men, focusing on researchers' experience of mobility in their later career stages (after gaining a PhD).

2.5.2.3. *Computation method*

Data needed

(F_i) Number of female researchers (R2, R3 and R4 career stages combined) who – in the last 10 years – moved for at least three months to a country other than that where they attained their highest educational qualification. **Unit: Number**.

(F) Number of female researchers (R2, R3 and R4 career stages combined). **Unit: Number**.

[25] For example, the gap between the EU employment rate of men and women widens with the arrival of dependent children. See Eurostat, 'Employment rate of adults by sex, age groups, highest level of education attained, number of children and age of youngest child (%)' [lfst_hheredch].

[26] European Parliament resolution of 21 May 2008 on women and science (2007/2206(INI)) https://www.europarl.europa.eu/sides/getDoc.do?type=TA&language=EN&reference=P6-TA-2008-0221

(M_i) Number of male researchers (R2, R3 and R4 career stages combined) who – in the last 10 years – moved for at least three months to a country other than that where they attained their highest educational qualification. **Unit: Number**.

(M) Number of male researchers (R2, R3 and R4 career stages combined). **Unit: Number**.

(S) Calibrated sampling weights for individual survey results, by country and field of Research and Development.

Source of data

European Commission - MORE Survey on mobility patterns and career paths of researchers (https://www.more-4.eu/). Custom extraction of MORE4 Survey data.

Computation formula

This indicator presents the percentage points difference in the percentage of female/male researchers (R2, R3 and R4 combined) who were 'internationally mobile' in the last 10 years (using the definition provided in this indicator). It is calculated by subtracting women's rate of mobility from that of men.

Before calculating this indicator, one must weigh the survey results to increase their representativeness. Pre-calculated sampling weights (by country and field of Research and Development) were included in the MORE4 dataset.

Following the weighting phase, calculate the indicator as normal. Using the weighted values, perform these calculations:

Percentage of internationally mobile female researchers (R2, R3 and R4) = F_{iw} / F_w

Percentage of internationally mobile male researchers (R2, R3 and R4) = M_{iw} / M_w

Differences between genders in international mobility in post-PhD career stages

$= (M_{iw} / M_w) - (F_{iw} / F_w)$

where:

i denotes international mobility (using the definition provided for this indicator);

w denotes that the values are weighted.

For example, F_{iw} refers to the number of internationally mobile female researchers (R2–R4) (weighted, using definition described above), whilst M_w refers to the total number of male researchers (R2–R4) (weighted).

2.5.2.4. *Specifications*

This indicator focuses on researchers in the higher education sector only. The MORE4 Survey applies the Frascati Manual definition of researchers.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

In this indicator, **researchers who are 'internationally mobile'** are defined as those who have worked abroad for more than three months at least once in the last 10 years,

since obtaining their highest educational qualification (PhD or other). It is based on a direct question in the MORE4 Survey of Higher Education Institutions[27].

The survey asks researchers to classify their career stage, using the categories defined in the European Framework for Research Careers (DG Research and Innovation, 2011):

- R1: First Stage Researcher (up to the point of PhD)
- R2: Recognised Researcher (PhD holders or equivalent who are not yet fully independent)
- R3: Established Researcher (researchers who have developed a level of independence)
- R4: Leading Researcher (researchers leading their research area or field).

This indicator focuses on those who classified themselves as being in the R2, R3 and R4 phases (combined).

2.5.2.5. *Comments and critical issues*

- There are large differences in the distribution of researchers across different career stages, depending on the country.
- The selected method for the survey results is direct weighting, so that the multiplication of each variable with the weighting coefficient returns the reference population. Here, sampling weights were calculated using the estimated number of researchers in each field of Research and Development in a given country (2009 data). Note that the sampling error can be higher at subpopulation level (including for gender).

2.5.3. ***Proportion of researchers employed part-time among researchers in the higher education sector (HES), by sex***

2.5.3.1. *Definition of indicator*

This indicator compares the proportion of persons employed part-time among female and among male researchers. It covers the higher education sector (HES) only. The researcher's country is his/her country of current employment. Organisations based in countries outside of those covered in She Figures are excluded.

2.5.3.2. *Rationale*

Part-time work is an important feature of working conditions with noteworthy gender aspects. The predominance of women in part-time work is on the one hand often explained by gender stereotypes related to family responsibilities but is also linked to gender segregation in employment. On the other hand, part-time work might be seen as

- [27] Changes to the MORE Surveys affect the comparability across different editions of the She Figures. In She Figures 2012, mobile researchers were defined as researchers who, in the last three years, had moved from the country where they had obtained their highest level of qualification to work as a researcher for at least three months in another country. The definition did not distinguish between career stages. In contrast to the 2012 edition, this indicator has been separated into two indicators, one focusing on mobility during their PhD – for researchers in the early career stages (R1 and R2 combined) – and another focusing on mobility in the last 10 years in the post-PhD phases (R2–R4).

an instrument to increase the labour market participation – and thus, to a certain extent at least, the economic independence of women (Burri S. and Aune H., 2013). Different types of work flexibility may have fewer negative, gender-specific consequences, as a recent critical analysis of part-time work in the Netherlands shows (Vinkenburg C.J., van Engen M., Peters C.P., 2015).

As a first step towards understanding this situation better, this indicator aims to consider the relative propensity of female and male researchers to be employed part-time.

2.5.3.3. *Computation method*

Data needed

(F_p) Number of female researchers in the HES who indicated that they worked part-time, combining all of these three categories: part-time (more than 50 %), part-time (50 %) and part-time (less than 50 %). **Unit: Number**.

(M_p) Number of male researchers in the HES who indicated that they work part-time, combining all of these three categories: part-time (more than 50 %), part-time (50 %) and part-time (less than 50 %). **Unit: Number**.

(F) Number of female researchers in the HES who indicated their employment status. **Unit: Number**.

(M) Number of male researchers in the HES who indicated their employment status. **Unit: Number**.

(S) Sampling weights for individual survey results, by country and field of Research and Development.

Source of data

European Commission - MORE Survey on mobility patterns and career paths of researchers (https://www.more-4.eu/). Custom extraction of MORE4 Survey data.

Computation formula

This indicator compares the proportion of the female researcher population that work part-time with the proportion of the male researcher population that work part-time.

Before calculating this indicator, one must weigh the survey results to increase their representativeness. Pre-calculated sampling weights (by country and field of Research and Development) were included in the MORE4 dataset.

Following the weighting phase, calculate the indicator as normal, but using the weighted numbers. Perform these calculations:

Proportion of persons employed part time among female researchers in the HES

$= \mathrm{F_{pw}} / \mathrm{F_{w}}$

Proportion of persons employed part time among male researchers in the HES

$= \mathrm{M_{pw}} / \mathrm{M_{w}}$

where:

p denotes part-time employment;

w denotes that the values are weighted.

For example, F_{pw} indicates the number of part-time female researchers in the higher education sector (weighted), whilst M_w indicates the number of male researchers in the higher education sector (weighted).

2.5.3.4. *Specifications*

This indicator focuses on researchers in the higher education sector only. It covers researchers at all career stages. The MORE4 Survey applies the Frascati Manual definition of researchers.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

2.5.3.5. *Comments and critical issues*

- The results can give an indication of the relative working conditions of male and female researchers, but it is worth bearing in mind that this indicator does not explore the reasons behind potential differences, nor does it provide a value judgement as to relative merits of working part-time or full-time. Using this indicator alone, it is not possible to judge the extent to which part-time employment is a free choice or a constraint.

- The selected method for the survey results is direct weighting, so that the multiplication of each variable with the weighting coefficient returns the reference population.

2.5.4. ***Proportion of researchers in the higher education sector (HES) working under precarious contracts, by sex***

2.5.4.1. *Definition of indicator*

This indicator compares the proportion of persons with precarious working contracts among female and among male researchers in the higher education sector (HES). Each researcher's country is his/her country of current employment. Organisations based in countries outside of those covered in She Figures were excluded.

2.5.4.2. *Rationale*

The existence and increase of precarious employment is subject to debate throughout the EU (DG for Internal Policies, 2016). Researchers with 'precarious working contracts' are those with no contracts, with fixed term contracts of up to one year, or with other contracts. Most affected are junior academic positions or other positions relying on third-party funding. The provision of research jobs that are associated with precariousness is in sharp conflict with the EU wide goal to provide attractive and secure positions in academia to fully exploit Europe's talent pool for HES within the new European Research Area (ERA) and Horizon Europe initiatives (DG for Research and Innovation, 2020).

This indicator aims to measure the relative propensity of female and male researchers to be employed on such contracts.

2.5.4.3. *Computation method*

Data needed

(F_i) Number of female researchers in the HES who indicated that they worked on a 'precarious' working contract. **Unit: Number**.

(M_i) Number of male researchers in the HES who indicated that they worked on a 'precarious' working contract. **Unit: Number**.

(F) Number of female researchers who indicated their contractual status. **Unit: Number**.

(M) Number of male researchers who indicated their contractual status. **Unit: Number**.

(S) Sampling weights for individual survey results, by country and field of Research and Development ('weihc'), available through the MORE dataset.

Source of data

European Commission - MORE Survey on mobility patterns and career paths of researchers (https://www.more-4.eu/). Custom extraction of MORE4 Survey data.

Computation formula

Consistent with the approach followed in the MORE2 and MORE3 Surveys, this indicator considers the following researchers to have precarious working contracts:

- researchers who indicated they have a fixed-term contract of one year or less;
- researchers who indicated they have no contract;
- researchers who indicated they have an 'other' type of contract (often associated with student status).

This indicator compares the proportion of female researchers with precarious contracts with the proportion of the male researchers in the same position.

Before calculating this indicator, one must weigh the survey results to increase their representativeness. Pre-calculated sampling weights (by country and field of Research and Development) were included in the MORE4 dataset.

Following the weighting phase, calculate the indicator as normal, but using the weighted numbers. Perform these calculations:

Proportion of persons with precarious working contracts among female researchers in the HES = F_{iw} / F_w

Proportion of persons with precarious working contracts among male researchers in the HES = M_{iw} / M_w

where:

i denotes employment on a 'precarious' working contract;

w denotes that the values are weighted.

For example, F_{iw} indicates the number of female researchers on precarious contracts in the higher education sector (weighted), whilst M_w indicates the number of male researchers in the higher education sector (weighted).

2.5.4.4. *Specifications*

This indicator focuses on researchers in the higher education sector only. It covers researchers at all career stages. The MORE4 Survey applies the Frascati Manual definition of researchers.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

2.5.4.5. *Comments and critical issues*

- The results can give an indication of the relative working conditions for male and female researchers, but it is worth bearing in mind that this indicator does not explore the reasons behind potential differences, nor does it provide a value judgement as to relative merits of working on different contracts. Using this indicator alone, it is not possible to judge the extent to which the use of different contracts is a free choice or a constraint.

- The selected method for the survey results is direct weighting, so that the multiplication of each variable with the weighting coefficient returns the reference population.

2.5.5. ***Proportion of researchers in the higher education sector (HES) working under precarious contracts, by sex and career stage***

2.5.5.1. *Definition of indicator*

This indicator compares the proportion of persons with precarious working contracts in the subcategories created by sex and career stage in the higher education sector (HES). Each researcher's country is his/her country of current employment. Organisations based in countries outside of those covered in She Figures were excluded.

2.5.5.2. *Rationale*

This indicator aims to provide information on whether sex intersects with career stage in precarious forms of employment. Junior academic staff are more likely to work under precarious working contracts than senior staff. The provision of research jobs that are associated with precariousness is in sharp conflict with the EU wide goal to provide attractive and secure positions in academia to fully exploit Europe's talent pool for HES within the new European Research Area (ERA) and Horizon Europe initiatives (DG for Research and Innovation, 2020).

2.5.5.3. *Computation method*

<u>Data needed</u>

($F_{i,c}$) Number of female researchers in the HES who indicated that they worked on a 'precarious' working contract, by career stage. **Unit: Number**.

($M_{i,c}$) Number of male researchers in the HES who indicated that they worked on a 'precarious' working contract, by career stage. **Unit: Number**.

(F_c) Number of female researchers who indicated their contractual status, by career stage. **Unit: Number**.

(M_c) Number of male researchers who indicated their contractual status, by career stage. **Unit: Number**.

(S) Sampling weights for individual survey results, by country and sex, available through the MORE dataset.

Source of data

European Commission - MORE Survey on mobility patterns and career paths of researchers (https://www.more-4.eu/). Custom extraction of MORE4 Survey data.

Computation formula

Consistent with the approach followed in the MORE2 and MORE3 Surveys, this indicator considers the following researchers to have precarious working contracts:

- researchers who indicated they have a fixed-term contract of one year or less;
- researchers who indicated they have no contract;
- researchers who indicated they have an 'other' type of contract (often associated with student status).

Before calculating this indicator, one must weigh the survey results to increase their representativeness. Pre-calculated sampling weights (by country and sex) were included in the MORE4 dataset.

Following the weighting phase, calculate the indicator as normal, but using the weighted numbers. Perform these calculations:

Proportion of women with precarious working contracts among female researchers in career stage 'c' in the HES = $F_{i,c[w]} / F_{c[w]}$

Proportion of persons with precarious working contracts among male researchers in the HES = $M_{i,c[w]} / M_{c[w]}$

where:
i denotes employment on a 'precarious' working contract;
c denotes career stage;
w denotes that the values are weighted.

2.5.5.4. *Specifications*

This indicator focuses on researchers in the higher education sector only. It covers researchers at all career stages. The MORE4 Survey applies the Frascati Manual definition of researchers.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

2.5.5.5. *Comments and critical issues*

- This indicator breaks down for two variables, one of which have four categories. Given this, in most sub-categories, data are limited or missing.

- The selected method for the survey results is direct weighting, so that the multiplication of each variable with the weighting coefficient returns the reference population. This indicator is weighted by country and field of R&D.

2.5.6. *Proportion of researchers in the higher education sector (HES) working under precarious contracts, by sex and family status*

2.5.6.1. *Definition of indicator*

This indicator compares the proportion of persons with precarious working contracts in the subcategories created by sex and family status in the higher education sector (HES). Each researcher's country is his/her country of current employment. Organisations based in countries outside of those covered in She Figures were excluded.

2.5.6.2. *Rationale*

This indicator aims to provide information on whether sex intersects with family status in precarious forms of employment. Women tend to have unequal caring responsibilities when they are in couple with children, which may result in differences in the working conditions. The provision of research jobs that are associated with precariousness is in sharp conflict with the EU wide goal to provide attractive and secure positions in academia to fully exploit Europe's talent pool for HES within the new European Research Area (ERA) and Horizon Europe initiatives (DG for Research and Innovation, 2020).

2.5.6.3. *Computation method*

Data needed

$(F_{i,d})$ Number of female researchers in the HES who indicated that they worked on a 'precarious' working contract, by family status. **Unit: Number**.

$(M_{i,d})$ Number of male researchers in the HES who indicated that they worked on a 'precarious' working contract, by family status. **Unit: Number**.

(F_d) Number of female researchers who indicated their contractual status, by family status. **Unit: Number**.

(M_d) Number of male researchers who indicated their contractual status, by family status. **Unit: Number**.

(S) Sampling weights for individual survey results, by country and sex, available through the MORE dataset.

Source of data

European Commission - MORE Survey on mobility patterns and career paths of researchers (https://www.more-4.eu/). Custom extraction of MORE4 Survey data.

Computation formula

Consistent with the approach followed in the MORE2 and MORE3 Surveys, this indicator considers the following researchers to have precarious working contracts:

- researchers who indicated they have a fixed-term contract of one year or less;
- researchers who indicated they have no contract;

- researchers who indicated they have an 'other' type of contract (often associated with student status).

Before calculating this indicator, one must weigh the survey results to increase their representativeness. Pre-calculated sampling weights (by country and sex) were included in the MORE4 dataset.

Following the weighting phase, calculate the indicator as normal, but using the weighted numbers. Perform these calculations:

Proportion of women with precarious working contracts among female researchers in family status 'd' in the HES = $F_{i,d[w]} / F_{d[w]}$

Proportion of persons with precarious working contracts among male researchers in the HES = $M_{i,d[w]} / M_{d[w]}$

where:
i denotes employment on a 'precarious' working contract;
d denotes family status;
w denotes that the values are weighted.

2.5.6.4. *Specifications*

This indicator focuses on researchers in the higher education sector only. It covers researchers at all career stages. The MORE4 Survey applies the Frascati Manual definition of researchers.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods (§5.35, Frascati Manual, OECD, 2015).

2.5.6.5. *Comments and critical issues*

- This indicator breaks down for two variables, one of which have four categories. Given this, in most sub-categories, data are limited or missing.

- The selected method for the survey results is direct weighting, so that the multiplication of each variable with the weighting coefficient returns the reference population. This indicator is weighted by country and sex in contrast to the other MORE4 indicators (which are weighted by country and field of R&D) as it breaks down for family status.

- It must be noted that researcher couples are not uncommon. Due to this fact, when one compares the values of the indicator between women and men whose family status is "in couple" one should keep in mind that several of these women's and men's couple partner is also in the same proportion.

2.6. Institutional change indicator through web-scraping techniques

Content-based rationale

This indicator, that was calculated with the method of web-scraping, presents the proportion of RPOs that report in their websites measures and actions taken towards gender equality by type of research performing organisation (higher education institutions (HEIs); public research organisations (PROs)).

The web-scraping was performed using the SerpApi, a Google search application programming interface (API) though Python scripts. The detailed descriptions of the indicators based on web-scraping techniques follow below.

2.6.1. ***Proportion of RPOs that have taken measures and actions to promote Gender Equality, by type of organisation***

2.6.1.1. *Definition of indicator*

The proportion of research performing organisations that report in their websites to have taken actions and measures to promote gender equality, by type of organisation (higher education institutions and public research organisations).

2.6.1.2. *Rationale*

Institutional change is embedded in the EC's Gender Equality Strategy 2020-2025 (European Commission, 2020b) and Strategic Plan for Horizon Europe 2021-2024 (DG Research and Innovation, 2020) that sees it as a necessary tool to remove barriers to gender equality through the implementation of Gender Equality Plans (GEPs). Institutional change is also a key priority of the new European Research Area (ERA), as the European Commission 'will propose as of 2021 […] the development of inclusive gender equality plans with the European Union Member States (EU MS) in order to promote gender equality in R&I' (European Commission, 2020b).

A key instrument for institutional change is the development and implementation of targeted actions and measures towards gender equality in RPOs. This indicator relies on web-scraping techniques to capture the proportion of RPOs that report in their website that they have taken actions and measures towards gender equality. This indicator builds on a similar indicator that was published in previous editions of 'She Figures', 'Proportion of RPOs that have adopted GEPs'. The strict definition of Gender Equality Plans[28] could not be used because of the technique. Instead, a broader term of 'measures and actions towards gender equality' was used.

2.6.1.3. *Computation method*

<u>Data needed</u>

(X) The number of RPOs in whose website was detected at least one specific search term indicating the presence of measures and actions taken by it to promote Gender Equality.

[28] A GEP is a set of measures and actions aiming to create institutional and cultural change towards gender equality. It is put in place following a self-assessment of the organisation. GEPs sets concrete objectives and targets to be achieved, the strategies and practices adopted for their achievement, and the creation of effective monitoring and evaluation systems using indicators.

(T) Total number of RPOs whose website's content was scraped and searched for the specific terms.

Source of data

a) the higher education institutions (HEIs) listed in the European Tertiary Education Register (ETER);

b) the public bodies and research organisations that participated in projects under Framework Programme 7 (FP7) and Horizon 2020 (H2020) framework programme for research and innovation

c) Input from statistical correspondents potentially revising list (b)

Computation formula

Proportion of ROs that have taken actions and measures towards Gender Equality = X/T

2.6.1.4. *Specifications*

The organisations' websites were scraped using a specific list of terms and phrases, translated in each country's official language(s). The initial list of search phrases was developed in English though manual search of a random sample of websites (their English versions) from all countries and identification of common terms and phrases in the websites indicating actions and measures taken by the organisation to promote gender equality. Afterwards, exploratory web-scrapes were used to identify which of these terms and phrases capture more accurately actions and measures towards gender equality. The final list of search phrases (in English) is:

- Gender Equality
- Gender Equality Plan
- Equal opportunities officer
- Equal participation officer
- Eliminate/Prevent sex discrimination
- Eliminate/Prevent harassment
- Harassment policy
- Gender diversity committee
- Gender diversity office
- Gender diversity task force

The final list of search phrases was translated in all official languages of the participating countries and was sent to the Statistical Correspondents to review and correct if needed. For countries that have more than one official language (e.g., Belgium), the Statistical Correspondents indicated which organisations should be scraped in each language.

The list of public research organisations initially comprised the public bodies and research organisations that participated in projects under Framework Programme 7 (FP7) and Horizon 2020 (H2020) framework programme for research and innovation. The initial lists were downloaded from CORDIS (https://cordis.europa.eu/en) on 11/05/2020 and were cleaned of duplicate records before being sent to the Statistical Correspondents for review. The Statistical Correspondents replied either with a revised list from which any remaining higher education institutes and business enterprises had been removed or with a list of public organisations from their national R&D survey.

2.6.1.5. *Comments and critical issues*

It must be noted that the results of this indicator are estimates. The accuracy of this indicator was calculated at 86 % during the exploratory web-scraping phase. This means

that the indicator correctly assigned organisations as having or not having taken actions and measures towards gender equality in 86% of the cases.

Moreover, not all organisations funded by FP7 or H2020 are research organisations (e.g., municipal authorities). The review of the lists extracted from CORDIS by the statistical correspondents contributed greatly to the removal of such organisations from the lists.

Since not all national lists of public organisations were prepared in the same way (some are curated lists of FP7/H2020 participants, others are the lists of public organisations surveyed for the production of R&D statistics) no meaningful way of weighing the national results was identified and the indicator has not been computed at EU-level.

Also, some of the search phrases may be more or less common across countries (and under-/over-estimate the indicator). In order to have a comparable research in all counties, country-specific terms were excluded even if they may have had increased accuracy at country level.

It must be noted that for this indicator there is a break in time series. Neither the method of measurement not the definition of indicator are the same as the indicators that were presented in past editions of She Figures.

Finally, during the exploratory web-scraping analysis, it was noted that when scraping PDF files attached to the websites the false positive rate (i.e., marking an organisation as having taken actions and measures towards gender equality when it really has not) increased and overall accuracy decreased. Therefore, the indicator was calculated based only on the content of the websites and not of any attached PDFs.

2.7. Women in Science (WiS) questionnaire

Content-based rationale

The indicators that stem from this source investigate the under-representation of women at the higher levels of the academic career path and in positions of power (known as the 'glass ceiling' phenomenon – whereby the representation of women decreases as the seniority of the role increases). They cover a wide range of sectors, particularly in science and technology, as well as the differences in success in obtaining research funding, by sex. Indicators computed include the proportion of women academic staff by grade and in total; the proportion of women grade A staff by main field of Research and Development; the distribution of grade A staff across fields of Research and Development by gender; and the Glass Ceiling Index.

Broad overview of the source

The Women in Science (WiS) questionnaires were sent to appointed Statistical Correspondents and provide data in support of the sets of indicators investigating the under-representation of women at the higher levels of the academic career path, as detailed in the rationale above.

The detailed descriptions of the indicators based on Women in Science questionnaire follow below.

2.7.1. ***Proportion of women among academic staff, by grade***

2.7.1.1. *Definition of indicator*

This indicator presents the proportion of women among the persons occupying positions at different grades of an academic career for a given year.

2.7.1.2. *Rationale*

By looking at the proportion of women present at each grade, one can track their progress in advancing through the stages of the academic career and identify the levels at which women are lost. Indeed, in 2016, despite accounting for almost 60 % of all bachelor's and master's graduates in the EU-28, women were still severely under-represented at the higher levels of academic career path as only 24 % of full professors, 22 % of heads of higher education institutions and 27 % of board members in research decision-making are women (DG Research and Innovation, 2019b). As such, it is interesting to monitor the number of women present at each level of academia in order to observe whether there is progress towards reducing vertical segregation, which can be measured by looking at the relative under- or over-representation of women versus men at consecutive grades, hierarchical levels or career stages (Fox et al 2017).

2.7.1.3. *Computation method*

Data needed

(F_{GY}) Number of female academic staff at a given grade G (G = A, B, C, D, or T [Total]) for a given year Y. **Unit: Head count**.

(M_{GY}) Number of male academic staff at a given grade G (G = A, B, C, D, or T [Total]) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among academic staff, by grade = $\frac{F_{GY}}{F_{GY}+M_{GY}}$

2.7.1.4. *Specifications*

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e., below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.1.5. *Comments and critical issues*

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

It is important to note that these data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

2.7.2. ***Proportion of women among grade A positions***

2.7.2.1. *Definition of indicator*

This indicator presents the proportion of women among persons occupying the single highest-level academic positions in a given year.

2.7.2.2. *Rationale*

By comparing different years, this indicator allows one to track the progress made with regard to women's presence at the highest level of the academic career path. Indeed, in

2016, despite accounting for almost 60 % of all bachelor's and master's graduates in the EU-28, women were still severely under-represented at the higher levels of academic career path as only 24 % of full professors, 22 % of heads of higher education institutions and 27 % of board members in research decision-making are women (DG Research and Innovation, 2019b). As such, it is interesting to monitor the number of women present at each level of academia in order to observe whether there is progress towards reducing vertical segregation, which can be measured by looking at the relative under- or over-representation of women versus men at consecutive grades, hierarchical levels or career stages (Fox et al 2017).

2.7.2.3. *Computation method*

Data needed

(F_{AY}) Number of women in grade A academic positions for a given year Y. **Unit: Head count**.

(M_{AY}) Number of men in grade A academic positions for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among grade A positions = $\frac{F_{AY}}{F_{AY}+M_{AY}}$

2.7.2.4. *Specifications*

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.2.5. *Comments and critical issues*

The classification of academic positions into grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

It is important to note that these data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

2.7.3. ***Proportion of grade A among academic staff, by sex***

2.7.3.1. *Definition of indicator*

This indicator allows for a comparison of the number of male and female staff at the highest-level academic positions compared to the number of staff of the same sex across all academic positions, for a given year.

2.7.3.2. *Rationale*

In 2016, despite accounting for almost 60 % of all university graduates in the EU-28, women were still severely under-represented at the higher levels of academic career path as only 24 % of full professors, 22 % of heads of higher education institutions and 27 % of board members in research decision-making are women (DG Research and Innovation, 2019b).

In this indicator, the low number of women grade A staff is compared to the overall number of female staff in academia, thereby correcting for the relative presence of women in academic positions overall. The advantage of such a calculation is that it moves beyond the absolute numbers of men and women in academic positions, which enhances comparability of the measure across different settings.

2.7.3.3. *Computation method*

Data needed

(F_{GY}) Number of female academic staff at a given grade G (G = A or T [Total]) for a given year Y. **Unit: Head count**.

(M_{GY}) Number of male academic staff at a given grade G (G = A or T [Total]) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of grade A among female academic staff = $\frac{\mathrm{F_{AY}}}{\mathrm{F_{TY}}}$

Proportion of grade A among male academic staff = $\frac{\mathrm{M_{AY}}}{\mathrm{M_{TY}}}$

2.7.3.4. *Specifications*

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e., below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.3.5. *Comments and critical issues*

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

It is important to note that these data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

2.7.4. ***Proportion of women among academic staff, by main field of Research and Development and grade***

2.7.4.1. *Definition of indicator*

This indicator looks at the presence of women in top academic positions across different fields of R&D, allowing for the identification of the fields in which women are more or less present for a given year.

2.7.4.2. *Rationale*

In 2016, despite accounting for almost 60 % of all university graduates in the EU-28, women were still severely under-represented at the higher levels of academic career path as only 24 % of full professors, 22 % of heads of higher education institutions and 27 % of board members in research decision-making are women (DG Research and Innovation, 2019b).

However, there may be some differences in the employment of women in top positions across different fields of Research and Development. As such, looking at the proportion of women in different positions of seniority reveals which fields of Research and Development have seen a more successful integration of female staff in top positions over time.

2.7.4.3. *Computation method*

Data needed

(F_{GSY}) Number of women at a given seniority grade G (G = A, B, C, D, or T [Total]), in main field of R&D S for reference year Y. **Unit: Head count**.

(M_{GSY}) Number of men at a given seniority grade G (G = A, B, C, D, or T [Total]), in main field of R&D S for reference year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among academic staff at grade G in main field S = $\frac{F_{GSY}}{F_{GSY}+M_{GSY}}$

Proportion of women among academic staff at grade G in main field S = $\frac{M_{GSY}}{F_{GSY}+M_{GSY}}$

2.7.4.4. *Specifications*

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e., below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.

The Frascati Manual (OECD, 2015) provides definitions for the six main fields of Research and Development (p.95). The following abbreviations are used:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities (H)
- unknown (U).
  - Unknown is not a field in Frascati; it has been added in the WiS questionnaire so that data can also be provided for academic staff whose field is unknown.

The breakdown of researchers by field of Research and Development is according to the field in which they work and not according to the field of their qualification.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.4.5. *Comments and critical issues*

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

It is important to note that these data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

2.7.5. ***Distribution of grade A staff across fields of Research and Development, by sex***

2.7.5.1. *Definition of indicator*

This indicator reveals differences in the distribution of male and female grade A staff across the different fields of Research and Development for a given year, by presenting the relative proportion of grade A staff of a given sex by field.

2.7.5.2. *Rationale*

In 2016, despite accounting for almost 60 % of all university graduates in the EU-28, women were still severely under-represented at the higher levels of academic career path as only 24 % of full professors, 22 % of heads of higher education institutions and 27 % of board members in research decision-making are women (DG Research and Innovation, 2019b).

Since this indicator corrects for the total number of grade A staff for each sex, it allows for a comparison of the fields of R&D in which each sex is more or less present in the top levels.

2.7.5.3. *Computation method*

Data needed

(F_{ASY}) Number of grade A women in main field of R&D S for year Y. **Unit: Head count**.

(M_{ASY}) Number of grade A men in main field of R&D S for year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of S field of R&D among grade A women = $\frac{\mathrm{F_{ASY}}}{\sum_{\mathrm{S}} \mathrm{F_{ASY}}}$

Proportion of S field of R&D among grade A men = $\frac{\mathrm{M_{ASY}}}{\sum_{\mathrm{S}} \mathrm{M_{ASY}}}$

2.7.5.4. *Specifications*

The grades presented in She Figures are based upon national mappings according to the following definitions:

(A) The single highest grade / post at which research is normally conducted within the institutional or corporate system

The Frascati Manual (OECD, 2015) provides definitions for the six main fields of Research and Development (p.95). The following abbreviations are used:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities (H)
- unknown (U).
  - Unknown is not a field in Frascati; it has been added in the WiS questionnaire so that data can also be provided for academic staff whose field is unknown.

The breakdown of researchers by field of Research and Development is according to the field in which they work and not according to the field of their qualification.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.5.5. *Comments and critical issues*

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

It is important to note that these data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

2.7.6. ***Glass Ceiling Index***

2.7.6.1. *Definition of indicator*

The Glass Ceiling Index (GCI) is a relative index comparing the proportion of women in academia (grades A, B, and C) to the proportion of women in top academic positions (grade A positions; equivalent to full professorships in most countries), for a given year. The GCI is a positive function that ranges above 0 to infinity. A GCI of 1 indicates that there is no difference between women and men in the chance of being promoted. A score

of less than 1 means that women are over-represented at grade A level and a GCI score of more than 1 points towards a glass ceiling effect, meaning that women are under-represented in grade A positions. In other words, the interpretation of the GCI is that the higher the value, the stronger the glass ceiling effect and the more difficult it is for women to move into a higher position.

2.7.6.2. *Rationale*

Both the Gender Statistics Database on women and men in decision-making and the Gender Equality Index of EIGE demonstrate the under-representation of women in positions of power, across a wide range of sectors in the EU. Moreover the European Parliament 'notes that despite positive changes in recent years, gender equality in science and academia has still not been achieved […] and points out the strikingly low presence of women in the highest academic and decision-making positions in scientific institutions and universities, which indicates the existence of a glass ceiling, that is, invisible barriers based on prejudices which stand in the way of women accessing positions of responsibility'.[29]

The version of the index presented here measures the relative chance for women (as compared with men) of reaching a top academic position, correcting for the relative presence of women (as compared with men) in academic positions overall. As such, it indicates the opportunity, or lack of it, for women to move up the hierarchical ladder in their academic profession. The advantage of the GCI being a relative index is that it moves beyond the absolute numbers of men and women in possible academic positions, which enhances comparability of the measure across different settings.

2.7.6.3. *Computation method*

Data needed

(F_{GY}) Number of grade A, B and C (G subscript) women for a given year Y. **Unit: Head count**.

(M_{GY}) Number of grade A, B and C (G subscript) men for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

$$Glass\ Ceiling\ Index = \left(\frac{\frac{F_{AY} + F_{BY} + F_{CY}}{F_{AY} + F_{BY} + F_{CY} + M_{AY} + M_{BY} + M_{CY}}}{\frac{F_{AY}}{F_{AY} + M_{AY}}} \right)$$

[29] European Parliament, Texts adopted: European Parliament resolution of 9 September 2015 on women's careers in science and universities, and glass ceilings encountered (2014/2251(INI)) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.C_.2017.316.01.0173.01.ENG

2.7.6.4. *Specifications*

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e., below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.6.5. *Comments and critical issues*

The classification of academic positions into A, B, C grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

It is important to note that these data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

2.7.7. ***Proportion of women among grade A staff, by age group***

2.7.7.1. *Definition of indicator*

This indicator presents the proportion of women among grade A staff across different age groups (less than 35 years, 35–44 years, 45–54 years, and 55 years or more) for a given year.

2.7.7.2. *Rationale*

In 2016, despite accounting for almost 60 % of all university graduates in the EU-28, women were still severely under-represented at the higher levels of academic career path as only 24 % of full professors, 22 % of heads of higher education institutions and 27 % of board members in research decision-making are women (DG Research and Innovation, 2019b).

This indicator sheds light on the representation of women in grade A research positions in different age groups. There are various reasons why this may be of interest. For example, according to Eurostat, a higher proportion of women are outside of the labour

force due to caring responsibilities, including for children.[30] This may reduce their participation in the labour market during the key childbearing years of a particular country.

2.7.7.3. *Computation method*

Data needed

(F_{AOY}) Number of grade A women in age group O (<35 years, 35–44 years, 45–54 years, 55+ years) for a given year Y. **Unit: Head count**.

(M_{AOY}) Number of grade A men in age group O (<35 years, 35–44 years, 45–54 years, 55+ years) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among grade A staff in age group O = $\frac{\mathrm{F_{AOY}}}{\mathrm{F_{AOY}+M_{AOY}}}$

2.7.7.4. *Specifications*

The grades presented in She Figures are based upon national mappings according to the following definition:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.7.5. *Comments and critical issues*

Given that, in some countries, the proportion of academic staff at grade A level is very small in the youngest age group (those aged under 35), it is best not to comment on this group for these countries. The existence of a generational effect could be exemplified by the fact that the proportion of women is larger in the younger age groups. In addition, the classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

It is important to note that these data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education

[30] In 2017, in the EU, 37.8 % of women (aged 25 to 49) who were outside of the labour force were in the position due to looking after children or incapacitated adults. For men of the same age group outside of the labour force, the rate was 3.9 %. See Eurostat, 'Inactive population not seeking employment by sex, age and main reason', data table lfsa_igar.

institutions (excluding staff involved in teaching or administration only and not at all in research).

2.7.8. *Distribution of grade A staff across age groups, by sex*

2.7.8.1. *Definition of indicator*

This indicator presents the distribution of male and female grade A staff across age groups for a given year.

2.7.8.2. *Rationale*

In 2016, despite accounting for almost 60 % of all university graduates in the EU-28, women were still severely under-represented at the higher levels of academic career path as only 24 % of full professors, 22 % of heads of higher education institutions and 27 % of board members in research decision-making are women (DG Research and Innovation, 2019b).

This indicator corrects for the total number of grade A staff for each sex and therefore it allows for a comparison of the presence of each sex across the different age groups.

2.7.8.3. *Computation method*

Data needed

(F_{AOY}) Number of grade A women in age group O (<35 years, 35–44 years, 45–54 years, 55+ years) for a given year Y. **Unit: Head count**.

(M_{AOY}) Number of grade A men in age group O (<35 years, 35–44 years, 45–54 years, 55+ years) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of age group O among women grade A staff for year Y = $\frac{\mathrm{F_{AOY}}}{\sum_{\mathrm{O}} \mathrm{F_{AOY}}}$

Proportion of age group O among men grade A staff for year Y = $\frac{\mathrm{M_{AOY}}}{\sum_{\mathrm{O}} \mathrm{M_{AOY}}}$

2.7.8.4. *Specifications*

The grades presented in She Figures are based upon national mappings according to the following definition:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.8.5. *Comments and critical issues*

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics. It is important to note that these data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

2.7.9. ***Proportion of women among heads of institutions in the higher education sector (HES)***

2.7.9.1. *Definition of indicator*

This indicator looks at the proportion of women among the heads of institutions in the higher education sector (HES) for a given year.

2.7.9.2. *Rationale*

The under-representation of women in leadership positions has broad implications for scientific advancement and for industries with a strong need for a technologically educated workforce. An increasing number of science institutions have been adopting in recent years a variety of measures to make improvements (Gvozdanović and Maes, 2018), such as leadership training, implicit bias training, Gender Equality Plans and the Human Resources Strategy for Researchers (Cameron *et al*, 2015). To confront this under-representation of women, part of the EC's strategy is to lead by example and to reach gender balance of 50 % at all levels of its management by the end of 2024 (European Commission, 2020b).

This indicator shows the proportion of women in decision-making positions as heads of institutions in the HES.

2.7.9.3. *Computation method*

Data needed

(F_Y) Number of women heads of institutions (in the higher education sector) for a given year Y. **Unit: Head count**.

(M_Y) Number of men heads of institutions (in the higher education sector) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among heads of institution in the HES = $\frac{F_Y}{F_Y+M_Y}$

2.7.9.4. *Specifications*

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.10. ***Proportion of women among heads of universities or assimilated institutions based on capacity to deliver PhDs***

2.7.10.1. *Definition of indicator*

This indicator looks at the proportion of women among only the heads of universities or assimilated institutions which can deliver PhDs (as opposed to the proportion of women among the heads of institutions in the higher education sector (HES) indicator, which considered all HES institutions), for a given year.

2.7.10.2. *Rationale*

The under-representation of women in leadership positions has broad implications for scientific advancement and for industries with a strong need for a technologically educated workforce. An increasing number of science institutions have been adopting in recent years a variety of measures to make improvements (Gvozdanović and Maes, 2018), such as leadership training, implicit bias training, Gender Equality Plans and the Human Resources Strategy for Researchers (Cameron *et al*, 2015).

This under-representation of women in decision-making positions is very well documented with this indicator that measures the proportion of women heads of institutions in the HES. Here, the scope is limited to universities or assimilated institutions based on capacity to deliver PhDs. These differ from general 'institutions in the higher education sector' as the HES sector 'comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions' (§3.67, OECD, 2015), many of which may not offer PhD programmes.

2.7.10.3. *Computation method*

Data needed

(F_Y) Number of women heads of universities or assimilated institutions which can deliver PhDs for a given year Y. **Unit: Head count**.

(M_Y) Number of men heads of universities or assimilated institutions which can deliver PhDs for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among heads of universities or assimilated institutions = $\frac{F_Y}{F_Y+M_Y}$

2.7.10.4. *Specifications*

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.11. ***Proportion of women on boards***

2.7.11.1. *Definition of indicator*

This indicator presents the proportion of women members of boards, top decision-making committees that have a crucial impact on the orientation of research in a given year.

2.7.11.2. *Rationale*

Since research funding applications are reviewed by scientific boards, the success of women in this process depends on the boards' members that make such decisions, who are often men. It is important to include women in this 'gate-keeping' procedure in order to ensure equal access to funding (DG Research, 2008; Bagihole, 2005).

Furthermore, the boards of research organisations have the potential to exercise extensive influence on scientific policy, either through directing core aspects of the agenda or supporting research through an advisory and coordinating role. Given that both advisory and executive boards have considerable decision-making power, the indicator assesses the proportion of women sitting on such boards in order to further investigate decision-making by women in academic careers.

2.7.11.3. *Computation method*

Data needed

(F_Y) Number of women on boards for a given year Y. **Unit: Head count**.

(M_Y) Number of men on boards for a given year Y. **Unit: Head count**.

The list of boards taken into account is given in the methodological Appendix of the main She Figures publication.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women on boards $= \frac{F_Y}{F_Y + M_Y}$

2.7.11.4. *Specifications*

In order to enhance cross-country comparability, Statistical Correspondents were asked to provide data only on the boards of umbrella, national-level research performing organisations (RPOs) and research funding organisations (RFOs). This does not include the boards / councils of individual higher education institutes. Instead, the aim is to capture the highest-level board[s] operating in the country. Umbrella, national

organisations which fund industrial research are taken into account only if they also perform / fund public research.

Scientific board of research organisation: A publicly or privately managed and financed group of elected or appointed experts that exists to implement scientific policy by, among other things, directing the research agenda, resource allocation and management within scientific research.

Administrative / advisory board of research organisation: A publicly or privately managed and financed group of elected or appointed experts that exists to support the research agenda in a nonexecutive function by, among other things, administering research activities, consulting and coordinating different actors and taking a general advisory role.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, Frascati Manual, OECD, 2015).

2.7.11.5. *Comments and critical issues*

No common definition of boards exists and the number of boards varies significantly between countries. It was requested that the metadata submitted should distinguish between boards of organisations performing research and the boards of organisations that are funding research, although both are included in the final computations.

2.7.12. ***Funding success rate difference between women and men***

2.7.12.1. *Definition of indicator*

This indicator presents research funding success-rate differences between women and men. A positive difference means that men have a higher success rate whereas a negative difference means that women have a higher success rate.

2.7.12.2. *Rationale*

The European Parliament has recognised that 'whereas despite all ongoing efforts to promote gender equality and equal opportunities, women still experience unequal access to research positions, funding, publishing and academic awards, and are also affected by rigid criteria for promotion and recognition and lack of funding or suitable policies to support them'.[31]

As such, this indicator looks at the differences in the success rate of men and women when applying for research funding. The calculation of a success rate rather than the use of raw numbers allows one to normalise for the total number of applications.

[31] European Parliament, Texts adopted: European Parliament resolution of 9 September 2015 on women's careers in science and universities, and glass ceilings encountered (2014/2251(INI)) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.C_.2017.316.01.0173.01.ENG

2.7.12.3. *Computation method*

Data needed

(F_{AY}) Number of female applicants for research funding for a given year Y. **Unit: Head count**.

(F_{BY}) Number of female beneficiaries of research funding for a given year Y. **Unit: Head count**.

(M_{AY}) Number of male applicants for research funding for a given year Y. **Unit: Head count**.

(M_{BY}) Number of male beneficiaries of research funding for a given year Y. **Unit: Head count**.

The list of national research funds taken into account is given in the methodological Appendix of the main She Figures publication.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Success rate difference between men and women = $\frac{\mathrm{M_{BY}}}{\mathrm{M_{AY}}} - \frac{\mathrm{F_{BY}}}{\mathrm{F_{AY}}}$

2.7.12.4. *Specifications*

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.7.12.5. *Comments and critical issues*

No common definition of funds exists and the total number of funds varies significantly between the countries and over the time period being considered. However, in an attempt to harmonise the data on funds provided by Statistical Correspondents of different countries, it was requested that data should cover all publicly managed research funds (funds granted by institutions in the public sector, excluding private sector funding). Furthermore, Statistical Correspondents were asked to exclude from reporting any funds which allocate funding exclusively on a first-come, first-served basis, i.e., without other selection criteria.

2.7.13. ***Funding success rate difference between women and men, by field of Research and Development***

2.7.13.1. *Definition of indicator*

This indicator presents research funding success-rate differences between women and men across different fields of Research and Development. A positive difference means that men have a higher success rate whereas a negative difference means that women have a higher success rate.

2.7.13.2. *Rationale*

The European Parliament has recognised that 'whereas despite all ongoing efforts to promote gender equality and equal opportunities, women still experience unequal access to research positions, funding, publishing and academic awards, and are also affected by rigid criteria for promotion and recognition and lack of funding or suitable policies to support them'.[32]

As such, this indicator looks at the differences in the success rate of men and women when applying for research funding. The calculation of a success rate rather than the use of raw numbers allows one to normalise for the total number of applications.

2.7.13.3. *Computation method*

Data needed

(F_{ASY}) Number of female applicants for research funding for a given year Y in a given field of Research and Development S. **Unit: Head count**.

(F_{BSY}) Number of female beneficiaries of research funding for a given year Y in a given field of Research and Development S. **Unit: Head count**.

(M_{ASY}) Number of male applicants for research funding for a given year Y in a given field of Research and Development S. **Unit: Head count**.

(M_{BSY}) Number of male beneficiaries of research funding for a given year Y in a given field of Research and Development S. **Unit: Head count**.

The list of national research funds taken into account is given in the methodological Appendix of the main She Figures publication.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Success rate difference between men and women for field S = $\frac{\mathrm{M_{BSY}}}{\mathrm{M_{ASY}}} - \frac{\mathrm{F_{BSY}}}{\mathrm{F_{ASY}}}$

2.7.13.4. *Specifications*

The Frascati Manual (OECD, 2015) provides definitions for the six main fields of Research and Development (p.95). The following abbreviations are used:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)

[32] Ibid

- social sciences (SS)
- humanities (H)
- multi-disciplinary (MU)
- unknown (U).
  - Unknown and multi-disciplinary are not fields in Frascati; they have been added in the WiS questionnaire so that data can also be provided for applications whose field is unknown or cover more than one specific field respectively.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) (§5.58, OECD, 2015).

2.8. Scopus™

Content-based rationale

Bibliometric indicators derived from Scopus™ are integrated into the She Figures 2021 publication. These indicators are used to identify differences between women and men regarding several indicators by country, year and field of research and development (FORD). The indicators in this section focus on revealing gender disparities as they relate to author pool (the group of individuals authoring publications), author impact (the publications and citations accrued by authors) and author team composition (the gender composition of research teams contributing as authors on publications).

- Additionally, the percentage of a country's research output integrating a gender dimension in its research and innovation content (GDRIC) will be assessed using this data source.

Broad overview of the source

The indicators presented in this section were computed by Elsevier using raw bibliographic data derived from the Scopus ™ database.

Scopus™ is Elsevier's abstract and citation database of peer-reviewed literature, covering 78.9 million documents. The data included in She Figures 2021 are based on documents published in over 40,000 journals, book series and conference proceedings by a broad number of publishers (over 5,000 publishers).

Its coverage is multi-lingual and global: approximately 46% of its titles are published in languages other than English (or published in both English and another language).

Its coverage is also inclusive across all major research fields, with 13,300 titles in the Physical Sciences, 14,500 in the Health Sciences, 7,300 in the Life Sciences, and 12,500 in the Social Sciences (the latter including some 4,000 Arts & Humanities related titles). Titles which are covered are predominantly serial publications (journals, trade journals, book series and conference material), but considerable numbers of conference papers are also covered from stand-alone proceedings volumes (a major dissemination mechanism, particularly in the Computer Sciences).

For this report, a static version of the database covering the period 2010-2019 inclusive was aggregated by country, region, and field.

Publications included in the computation of indicators are limited to peer-reviewed publications, that is, articles, reviews, and conference papers.

In addition to its content coverage, Scopus identifies individual authors and institutions, allowing for a robust analysis at the desired level of granularity whether it is country, state, institution, or author.

Country and regional attribution of each publication is based on the authorship by-line of each article. A publication is attributed to a country if an author indicates affiliation with an institute in that country.

A body of literature is available on the limitations and caveats in the use of 'bibliometric' data, such as the accumulation of citations over time, the skewed distribution of citations across articles and differences in publication and citation practices between fields of research, different languages, and applicability to social sciences and humanities research (Bar-Ilan, 2008). In social sciences and humanities, the bibliometric indicators presented in She Figures must be interpreted with caution because a reasonable proportion of research outputs in such fields take the form of books, monographs, and non-textual

media. As such, analyses of journal articles, their usage, and citation, provide a less comprehensive view for social sciences and humanities than for other fields, where journal articles comprise the vast majority of research outputs.

It is important to note that bibliometric approaches can only be used as a proxy to assess productivity and impact of publishing authors. All indicators based on bibliometric approaches can only give insights into the output of authors publishing in Scopus-indexed publications. For this reason, throughout the She Figures handbook, the term 'author' is used to indicate in fact publishing authors. However, a researcher in the sense of the OECD definition is not limited to these authors, and the metrics provided in this section do not assess all researchers, including, for example, researchers in the corporate sector who may not be publishing authors.

Calculation Period

Scopus indexed full names when available for all participating authors and for all publication years dating back to 1996. Author-level indicators are calculated based on authors during the period 2015-2019. Publication-level analyses cover publications over the period 2010 to 2019 and is calculated per annum when possible as well as in aggregate for the full ten-year period.

Included authors

The authors included in the indicator calculation are defined per indicator.

Possible Issues with hypercollaboration

The terms 'Hypercollaborative co-authorship' and 'hypercollaboration' have been coined to classify the growing phenomenon of articles that have hundreds or even thousands of co-authors. The rise of so-called 'Big Science' – a term used to describe research that requires major capital investment and is often, but not always, international in nature – may be one of the causes of this phenomenon. The frequency of such articles is still relatively small: just 827 articles published between 2010 and 2015 had more than 1,000 authors. Most of these came from CERN's Large Hadron Collider in Switzerland, and include, in May 2015, the most multi-authored research paper published to date, with 5,154 authors (Aad et al., 2015).

While hyper-collaborated publications may represent extreme outliers in co-authorship data and remain proportionally few, such hypercollaborative articles are included throughout the analyses. Like other collaborative articles, they are counted as single internationally co-authored articles for each country represented in them and for each country pairing.

Journal Classification

Journals may be assigned to several major and minor subject areas. Major subject areas are defined according to 27 All Science Journal Classification (ASJC) categories. These are 27 subject categories that all Scopus indexed journals are classified as. Each of the 27 subject categories is further subdivided into a total of 334 minor sub-categories. Because some journals can be classified as multi-category (i.e., more than one subject), each publication may fall into more than one subject classification. For the analyses in She Figures 2021, the ASJC classifications were mapped to the Fields of Research and Development (OECD, 2015) classifications. A full table of the mapping of Fields of Research and Development (FORD) classifications with the ASJC sub-categories can be found in Annex 4.

For the calculation of indicators based on FORD classification, it is important to note that publications are not classified in mutually exclusive fields, rather, some will be classified

in more than one field. For example, publication P may belong in both Medical Sciences (FORD 3) and Social Sciences (FORD 5). Although publication P will contribute once to the publication count of Medical Sciences and once to the publication count of Social Sciences, this publication will not be counted twice in the aggregated count of 'All' publications.

For citation-based measures fractional calculation has been used - distributing publication and citation counts equally across multiple journal categories; publication P would be counted as 0.5 publications for each of Medical Sciences and Social Sciences, and its citations would be shared equally between these subject areas.

Gender Classification

A binary gender was inferred for author IDs using the NamSor API (January 2019 release). The API provides a Gender Probability Score and gender classification based on three data points: country of origin, first name and last name. These three data points for authors were generated based on information related to each author ID.

- *Determination of author country of origin:*

Each author's country of origin was determined based on the country of affiliation listed on the publications from their first year of publication in Scopus (i.e., articles, reviews and conference papers). In some cases, authors had published in more than one country in their first year of publication. In these cases, the country with the largest number of publications was designated as the author's country of origin. Authors with equal numbers of publications in two or more countries were excluded from the gender disambiguation analysis. The process used to determine the author country of origin is summarized here.

For each author ID:

1. Identify year of first publication in Scopus.
2. Identify all publications from the first year that the author published.
3. Identify the country affiliation indicated by the author in the first-year publications.
4. Tabulate the countries of affiliation.
5. Tabulate the number of times each country was listed as the country of affiliation in the first year of publication.
6. Assign the author ID to the country most often indicated as the country of affiliation in the first year of publication as the country of origin.
7. If two countries appeared as the country of affiliation for an author an equal number of times in the first year of publication, then that author was excluded from the analysis.

- *First name determination:*

First and last name are required as input data for NamSor. Therefore, only author IDs with a first and last name were passed through the NamSor API to retrieve a Gender Probability Score. All author IDs for whom no first name data was available were not included in the analysis.

Different variants of an author's name are commonly observed across their publications. To identify the best first name to pass through NamSor for each author, all the name variants associated with each author ID were assessed. For each author ID in the Scopus

snapshot, all publications on which the author ID appears in the author field were examined and a list of all distinct first names associated with the author ID was generated. Based on this list, a table was generated with a revised first name for each author ID. The process used to determine the best first name to pass through NamSor is described here.

In cases where only a single first name was associated with an author ID:

1. When the name was of zero length, the best first name was assigned as null

e.g., author first names: []

best first name: null

These author IDs were excluded from the analysis.

2. When the name was not of zero length, the following nonsensical characters were removed if they were leading or trailing: "-!#&"

e.g., author first names: ["Tom&"]

best first name: Tom

In cases where multiple first names were associated with an author ID, the longest available name following removal of nonsensical characters was selected, provided this name was not composed of a string of initials, according to the following steps for each author ID:

1. Author first names were collected into a list.

2. Author first names were initialized as the empty string ' '.

3. The list of names was looped through. Each name in the list was stripped of nonsensical leading and trailing characters (-!#&).

4. The next name in the list was considered and subjected to the same treatment. If the length of the string was longer and not composed of a string of initials, this next name was then assigned as the revised first name. A string of initials among author first names was identified by comparing the number of periods that appear in the string to the number of characters (excluding whitespace, periods and nonsensical characters). When the number of periods was equal to the number of characters, this string was identified as a string of initials. This was done using Regex expressions.

5. This process was repeated until the end of the list of author first names is reached.

e.g., author first names: ["Samantha", "#Sam", "Sam", "S. E."]

best first name: Samantha

e.g., author first names: ["Samuel", "#Sam", "Sam", "S. E."]

best first name: Samuel

- *Gender Probability Score:*

The NamSor Gender Probability Score was used to predict the gender of each author. The Gender Probability Score is the natural log of the ratio of probabilities, as determined by a Naïve-Bayes model, of the name receiving the classification of either male or female. The Gender Probability Score is based on the best first name, last name and country of origin.

To predict the gender of each author, a table containing unique name-country combinations (represented in the Scopus snapshot) from the following three fields was created:

- best first name (based on the process described in the section "First name determination")

- last name (based on author ID)

- country of origin (based on the process described in the section "Determination of author country of origin")

Next the resulting name-country combinations were taken and passed through the NamSor API, which generated a classification (either male or female) with the associated Probability Calibrated Score for each combination.

The Probability Calibrated score reflects the confidence in the gender assignment returned by NamSor. The score ranges from 0.5 to 1.0. A score of 0.5 indicates high uncertainty that the inferred gender is correct and a score of 1.0 indicates the highest level of certainty that the inferred gender is correct. Previous work with the returned data from the NamSor API indicates that a cutoff Probability Calibrated score of 0.85 would yield a fairly high F1 score for male names (0.95) and female names (0.93).

Name-country combinations that fell short of this threshold were classified as unknown. The gender inferred for each name-country combination was then matched to author IDs based on the best first name, last name and country of origin.

The detailed descriptions of the indicators based on Scopus follow below.

2.8.1. ***Ratio of women to men amongst active authors***

2.8.1.1. *Definition of indicator*

This indicator compares the number of women who can be classified as "active authors" to the number of men who can be classified as "active authors" within a seniority level, country or region and field of research. A value above 1 indicates that among active authors, women outnumber men; a value below 1 indicates that among active authors, men outnumber women.

2.8.1.2. *Rationale*

Representation of women and men in the authorship team is partly dependent of the "available workforce" of authors. Assessing the ratio of women to men who can be categorized as "active authors", thereby shedding light on the level of gender balance among the authors within a selected category of country or region, field of research, and seniority level can shed light onto metrics related to inclusion on authorship teams. This metric is based on "active authors" to provide insight into the author group that forms the basis for the indicators 2.9.3 and 2.9.4. The insights gained from this analysis will

inform whether policies or interventions should focus on increasing the participation of women or men as active authors in particular countries or regions, fields, and career stages.

2.8.1.3. *Computation method*

This calculation is based on **active authors** only. Active authors are defined as those that produced 10 or more papers in the last 20 years (2000-2019) and at least 1 paper in the last 5 years (2015-2019) OR those who produced 4 or more papers in last 5 years (2015-2019).

To estimate career stage, the following definitions are applied:

Seniority level is estimated via the time elapsed since an author's first publication in a journal indexed in Scopus and has three categories:

- **<5**: authors whose first paper in Scopus was published in the last 5 years (2015-2019).
- **5 to 10**: authors whose first paper in Scopus was published more than 5 years ago and up to 10 years ago (2010-2014).
- **>10**: authors whose first paper in Scopus is published more than 10 years ago (2009 or prior).

Country designation is attributed to authors based on their publication output. An author counts towards a country's metrics if at least 30% of his/her publications during the period 2015-2019 list the country in the affiliation details.

Subject designation is attributed to authors based on their publication output. An author counts towards a FORD subject's metrics if at least 30% of his/her publications during the period 2015-2019 are in that given subject.

Data needed

$(\sum FAA_{xci})$ Number of women who are considered active authors (FAA) in a given seniority category (x), country (c), and field (s) during the period 2015-2019. **Unit: Number.**

$(\sum MAA_{xci})$ Number of men who are considered active authors (MAA) in a given seniority category (x), country (c), and field (s) during the period 2015-2019. **Unit: Number.**

Computation formula

Among all authors of peer-reviewed publications considered "active" based on the criteria above, the following indicator is calculated for each seniority level (x), country (c) and FORD (i):

Ratio of women to men amongst active authors: $\frac{(\sum FAA_{xci})}{(\sum MAA_{xci})}$

Source of data

Computed using Scopus™ data and NamSor™

2.8.2. ***Ratio of women to men amongst all authors***

2.8.2.1. *Definition of indicator*

This indicator compares the number of women authors to the number of men authors within a seniority level, country or region and field of research. A value above 1 indicates that among authors, women outnumber men; a value below 1 indicates that among authors, men outnumber women.

2.8.2.2. *Rationale*

Comparing the women and men who have authored publications may reveal the level of gender balance among the authors within a selected category of country or region, field of research, and seniority level. This metric therefor provides a point of comparison with metrics related to researcher gender ratio and inclusion of women on author teams. The insights gained from this analysis will inform whether policies or interventions should focus on increasing the participation of women or men as authors in particular countries or regions, fields, and career stages.

2.8.2.3. *Computation method*

This calculation is based on **all authors**. Authors are defined as those that produced at least one publication in the last 5 years (2015-2019).

To estimate career stage, the following definitions are applied:

Seniority level is estimated via the time elapsed since an author's first publication in a journal indexed in Scopus and has three categories:

- **<5**: authors whose first paper in Scopus was published in the last 5 years (2015-2019).
- **5 to 10**: authors whose first paper in Scopus was published more than 5 years ago and up to 10 years ago (2010-2014).
- **>10**: authors whose first paper in Scopus is published more than 10 years ago (2009 or prior).

Country designation is attributed to authors based on their publication output. An author counts towards a country's metrics if at least 30% of his/her publications during the period 2015-2019 list the country in the affiliation details.

Subject designation is attributed to authors based on their publication output. An author counts towards a FORD subject's metrics if at least 30% of his/her publications during the period 2015-2019 are in that given subject.

Data needed

$(\sum FA_{xci})$ Number of women who have authored a publication (FA) in a given seniority category (x), country (c), and field (i) during the period 2015-2019. **Unit: Number.**

$(\sum MA_{xci})$ Number of men who have authored a publication (MA) in a given seniority category (x), country (c), and field (i) during the period 2015-2019. **Unit: Number.**

Computation formula

Among all authors of peer-reviewed publications based on the criteria above, the following indicators is calculated for each seniority level (x), country (c) and FORD (i):

Ratio of women to men amongst active authors: $\frac{(\sum FA_{xci})}{(\sum MA_{xci})}$

Source of data

Computed using Scopus™ data and NamSor™

2.8.3. *Ratio of average number of publications by women to those by men*

2.8.3.1. *Definition of indicator*

This indicator compares the average number of publications by women who can be classified as "active authors" to the average number of publications by men who can be classified as "active authors" within a seniority level, country or region and field of research. A value above 1 indicates that among active authors, women publish more than men on average; a value below 1 indicates that among active authors, men publish more than women on average.

2.8.3.2. *Rationale*

An author's publication output is sometimes used to assess the productivity of a researcher. Comparing the average publication count of women and men in particular countries or regions, fields, and career stages may reveal differences in overall output within those categories. The insights gained from this indicator will inform whether policies based on publication output (such as promotional policies) should be benchmarked against norms within a gender category.

2.8.3.3. *Computation method*

This calculation is based on the publication output of **active authors** only. Active authors are defined as those that produced 10 or more papers in the last 20 years (2000-2019) and at least 1 paper in the last 5 years (2015-2019) OR those who produced 4 or more papers in last 5 years (2015-2019).

To estimate career stage, the following definitions are applied:

Seniority level is estimated via the time elapsed since an author's first publication in a journal indexed in Scopus and has three categories:

- **<5**: authors whose first paper in Scopus was published in the last 5 years (2015-2019).
- **5 to 10**: authors whose first paper in Scopus was published more than 5 years ago and up to 10 years ago (2010-2014).
- **>10**: authors whose first paper in Scopus is published more than 10 years ago (2009 or prior).

Country designation is attributed to authors based on their publication output. An author counts towards a country's metrics if at least 30% of his/her publications during the period 2015-2019 list the country in the affiliation details.

Subject designation is attributed to authors based on their publication output. An author counts towards a FORD subject's metrics if at least 30% of his/her publications during the period 2015-2019 are in that given subject.

Data needed

$(\sum FAA_{xci})$ Number of women who are considered active authors (FAA) in a given seniority category (x), country (c), and field (i) during the period 2015-2019. **Unit: Number.**

$(\sum MAA_{xci})$ Number of men who are considered active authors (MAA) in a given seniority category (x), country (c), and field (i) during the period 2015-2019. **Unit: Number.**

P Number of publications by an author. **Unit: Number.**

Computation formula

Among all authors of peer-reviewed publications from a country, the following metrics are computed for each seniority level:

Average number of publications per woman: $\frac{1}{(\sum FAA_{xci})}\sum_{j=1}^{FAA_{xci}}(P)_j$

Average number of publications per woman: $\frac{1}{(\sum MAA_{xci})}\sum_{j=1}^{MAA_{xci}}(P)_j$

Ratio of average number of publications by women to those by men : $\frac{\frac{1}{(\sum FAA_{xci})}\sum_{j=1}^{FAA_{xci}}(P)_j}{\frac{1}{(\sum MAA_{xci})}\sum_{j=1}^{MAA_{xci}}(P)_j}$

Source of data

Computed using Scopus™ data and NamSor™

2.8.4. ***Ratio of average FWCI of publications by women to that of men***

2.8.4.1. *Definition of indicator*

This indicator compares the average citation impact (based on field-weighted citation impact, FWCI) of publications by women who can be classified as "active authors" to the citation impact of publications by men who can be classified as "active authors" within a seniority level, country or region and field of research. A value above 1 indicates that among active authors, women's publications have a citation impact that is on average, higher than that of men's; a value below 1 indicates that among active authors, men's publications have a citation impact that is on average, higher than that of women's.

2.8.4.2. *Rationale*

Comparing the citation impact of men and women may reveal differences in citation behaviours. The insights gained from this analysis inform whether policies based on publication impact (such as promotional policies) should be benchmarked against norms within a gender category.

2.8.4.3. *Computation method*

This calculation is based on the publication output of **active authors** only. Active authors are defined as those that produced 10 or more papers in the last 20 years (2000-2019) and at least 1 paper in the last 5 years (2015-2019) OR those who produced 4 or more papers in last 5 years (2015-2019).

To estimate career stage, the following definitions are applied:

Seniority level is estimated via the time elapsed since an author's first publication in a journal indexed in Scopus and has three categories:

- **<5**: authors whose first paper in Scopus was published in the last 5 years (2015-2019).
- **5 to 10**: authors whose first paper in Scopus was published more than 5 years ago and up to 10 years ago (2010-2014).
- **>10**: authors whose first paper in Scopus is published more than 10 years ago (2009 or prior).

Country designation are attributed to authors based on their publication output. An author counts towards a country's metrics if at least 30% of his/her publications during the period 2015-2019 list the country in the affiliation details.

Subject designation is attributed to authors based on their publication output. An author counts towards a FORD subject's metrics if at least 30% of his/her publications during the period 2015-2019 are in that given subject.

Data needed

$(\sum FAA_{xci})$ Number of women who are considered active authors (FAA) in a given seniority category (x), country (c), and field (i) during the period 2015-2019. **Unit: Number.**

$(\sum FAA_{xci})$ Number of men who are considered active authors (MAA) in a given seniority category (x), country (c), and field (i) during the period 2015-2019. **Unit: Number.**

$\overline{FWCI}$ Mean FWCI of publications by an author. **Unit: Number.**

Field-weighted citation impact (FWCI): Field-weighted citation impact (FWCI) is an indicator of mean citation impact and compares the actual number of citations received by an article with the expected number of citations for articles of the same document type (article, review or conference proceeding paper), publication year and subject field. When an article is classified in two or more subject fields, the harmonic mean of the actual and expected citation rates is used. The indicator is therefore always defined with reference to a global baseline of 1 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document types (reviews typically attract more citations than research articles, for example) as well as subject-specific differences in citation frequencies. In general, the mean Field-Weighted Citation Impact (FWCI) for a set of N publications is defined as:

$$\overline{FWCI} = \frac{1}{N}\sum_{j=1}^{N}\frac{C_j}{E_j}$$

with

C_j = citations received by publication i

E_j = expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than one discipline, the harmonic mean is used to calculate E_j.

Computation formula

Among all authors of peer-reviewed publications from a country, the following metrics is computed for each seniority level:

Average FWCI for publications by women: $\frac{1}{\sum FAA_{xci}} \sum_{j=1}^{FAA_{xci}} (\overline{FWCI})_j$

Average FWCI for publications by men: $\frac{1}{\sum MAA_{xci}} \sum_{j=1}^{MAA_{ci}} (\overline{FWCI})_j$

Ratio of average FWCI of publications by women to that of men : $\frac{\frac{1}{\sum FAA_{xci}} \sum_{j=1}^{FAA_{xci}} (\overline{FWCI})_j}{\frac{1}{\sum MAA_{xci}} \sum_{j=1}^{MAA_{ci}} (\overline{FWCI})_j}$

Source of data

Computed using Scopus data ™ and NamSor ™

2.8.5. ***Average proportion of women among authors on publications***

2.8.5.1. *Definition of indicator*

This indicator is the average proportion of women among authors on publications from a given country or region and field of research. A value near 0.5, indicates that on average, women and men are represented at equal proportions on teams; a value above 0.5, indicates that on average, women are more highly represented than men on teams; a value below 0.5, indicates that on average, men are more highly represented than women on teams.

2.8.5.2. *Rationale*

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in research teams is valuable for ensuring that research outcomes reflect the experience of both men and women. This indicator looks at the contribution of women and men to research teams across countries or regions and fields of research and development (FORD).

2.8.5.3. *Computation method*

Country attributions of each publication is based on the authorship by-line of each article. A publication is attributed to a country if an author indicates affiliation with an institute in that country.

Subject designation of each publication is based on the source title of each article. A publication is attributed to a subject if the source title has been mapped to a subject in that field.

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum P_{cyi})$ The number of publications in a given country (c), year (y) and field (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and field (i), the formula for average proportion of women among authors on publications (APW) is:

$$(APW)_{CYi} = \frac{1}{\sum P_{cyi}} \sum_{j=1}^{P_{CYi}} \frac{\sum F}{\sum F + \sum M}$$

Source of data

Computed using Scopus ™ data and NamSor ™

2.8.6. ***Compound annual growth rate (CAGR) of average proportion of women among authors on publications***

2.8.6.1. *Definition of indicator*

Compound annual growth rate (CAGR) is defined as the year-over-year constant growth rate over a specified period of time. Starting with the first value in any series and applying this rate for each of the time intervals yields the amount in the final value of the series. Throughout the term CAGR is also referred to as '(yearly) growth rate.'

The indicator is calculated for the average proportion of women among authors on publications.

2.8.6.2. *Rationale*

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in research teams is valuable for ensuring that research outcomes reflect the experience of both men and women. This indicator looks at the growth rates of the scientific contribution of women on authorship teams across different countries and FORD based on authorship.

2.8.6.3. *Computation method*

Data needed

$(APW)_{cyi}$ Average proportion of women among authors on publications in a given country (c), year (y) and field (i). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Computation formula

CAGR for average proportion of women among authors = $(\mathrm{APW_{Cei}}/\mathrm{APW_{Csi}})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{APW_{Csi}}$ denotes the average proportion of women among authors in the start year;

$\mathrm{APW_{Cei}}$ denotes the average proportion of women among authors in the end year.

Source of data

Computed using Scopus data ™ and NamSor ™

2.8.7. ***Average proportion of women among authors on publications resulting from international collaborations***

2.8.7.1. *Definition of indicator*

This indicator is the average proportion of women among authors on publications resulting from international collaboration. A value is near 0.5, indicates that on average, women and men are represented at equal proportions on international authorship teams; a value above 0.5, indicates that on average, women are more highly represented than men on international authorship teams; a value below 0.5, indicates that on average, men are more highly represented than women on international authorship teams.

2.8.7.2. *Rationale*

Women are less likely than men to collaborate internationally on research papers (Elsevier, 2017) and this may have effects on the impact of their publications (as measured by citations). Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men.

This indicator looks at the average proportion of women among authorship teams resulting from internationally co-authored publications across different countries and fields of research and development (FORD).

2.8.7.3. *Computation method*

International collaboration is defined as multi-authored research outputs, where at least one author is from an institution inside the country of interest and at least one author is from an institution outside the country of interest (or EU for EU-27 and EU-28 calculations).

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum PI_{cyi})$ The number of publications resulting from international collaboration in a given country (c), year (y) and field (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and field (i), the formula for average proportion of women among authors on internationally collaborated publications (AIPW) is:

$$(AIPW)_{cyi} = \frac{1}{\sum PI_{cyi}} \sum_{i=1}^{PI_{cyi}} \frac{\sum F}{\sum F + \sum M}$$

The indicator is calculated for each country (c), year (y) and FORD (i).

Source of data

Computed using Scopus data ™ and NamSor ™

2.8.8. ***Compound annual growth rate (CAGR) of average proportion of women among authors on publications resulting from international collaborations***

2.8.8.1. *Definition of indicator*

Compound annual growth rate (CAGR) is defined as the year-over-year constant growth rate over a specified period of time. Starting with the first value in any series and applying this rate for each of the time intervals yields the amount in the final value of the series. Throughout the term CAGR is also referred to as '(yearly) growth rate.'

The indicator is calculated for the average proportion of women among authors on publications resulting from international collaboration across different countries and fields of research and development (FORD).

2.8.8.2. *Rationale*

Women publish fewer research papers on average than men and women are less likely than men to collaborate internationally on research papers (Elsevier, 2017) and this may have effects on the impact of their publications (as measured by citations). Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men. This indicator looks at the change in the average proportion of women among authors on publications resulting from international collaboration.

2.8.8.3. *Computation method*

Data needed

$(AIPW)_{cyi}$ Average proportion of women among authors on publications resulting from international collaboration in a given country (c), year (y) and field (i). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Computed using Scopus™ data and NamSor™

Computation formula

CAGR for average proportion of women among authors = $(\mathrm{AIPW_{Cei}}/\mathrm{AIPW_{Csi}})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{AIPW_{Csi}}$ denotes the average proportion of women authors on publications resulting from international collaboration in a given country (c), and field (i) in the start year;

$\mathrm{AIPW_{CeS}}$ denotes the average proportion of women authors on publications resulting from international collaboration in a given country (c), and field (i) in the end year.

2.8.9. ***Average proportion of women among authors on publications resulting from national collaboration***

2.8.9.1. *Definition of indicator*

This indicator is the average proportion of women among authors on publications resulting from national collaboration. A value is near 0.5, indicates that on average, women and men are represented at equal proportions on national authorship teams; a value above 0.5, indicates that on average, women are more highly represented than men on national authorship teams; a value below 0.5, indicates that on average, men are more highly represented than women on national authorship teams.

2.8.9.2. *Rationale*

Women are less likely than men to collaborate internationally on research papers (Elsevier, 2017) and this may have effects on the impact of their publications (as measured by citations). Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men. Statistics related to national collaboration can provide insight into whether there are barriers to authors collaborating outside of their own institution, which may be a first step in a researcher's path towards international collaboration.

This indicator looks at the average proportion of women among authorship teams resulting from nationally co-authored publications across different countries and fields of research and development (FORD).

2.8.9.3. *Computation method*

National collaboration is defined as multi-authored research outputs, where authors are affiliated with more than one institution within the same country.

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum PN_{cyi})$ The number of publications resulting from national collaboration in a given country (c), year (y) and field (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and field (i), the formula for average proportion of women among authors on internationally collaborated publications (ANPW) is:

$$(ANPW)_{cyi} = \frac{1}{\sum PN_{cyi}} \sum_{j=1}^{PN_{cyi}} \frac{\sum F}{\sum F + \sum M}$$

The indicator is calculated for each country (c), year (y) and FORD (i).

Source of data

Computed using Scopus ™ data and NamSor ™

2.8.10. ***Average proportion of women among authors on publications resulting from intra-EU27+ collaborations***

2.8.10.1. *Definition of indicator*

This indicator is the average proportion of women among authors on publications resulting from intra-EU27+ collaboration. A value near 0.5, indicates that on average, women and men are represented at equal proportions on EU27+ authorship teams; a value above 0.5, indicates that on average, women are more highly represented than men on EU27+ authorship teams; a value below 0.5, indicates that on average, men are more highly represented than women on EU27+ authorship teams.

2.8.10.2. *Rationale*

Women are less likely than men to collaborate internationally on research papers (Elsevier, 2017) and this may have effects on the impact of their publications (as measured by citations). Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men. Statistics related to intra-EU27+ collaboration can provide insight into whether there are barriers to authors collaborating outside of their own institution.

This indicator looks at the average proportion of women among authorship teams resulting from intra-EU27+ co-authored publications across different countries and fields of research and development (FORD).

2.8.10.3. *Computation method*

Intra-EU27+ collaboration is defined as multi-authored research outputs, where authors are affiliated with institutions in more than one of the 44 countries (EU-27 and Associated Countries), but all authors are based within the 44 countries.

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum PIE_{cyi})$ The number of publications resulting from intra-EU27+ collaboration in a given country (c), year (y) and field (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and field (i), the formula for average proportion of women among authors on internationally collaborated publications (AIEPW) is:

$$(AIEPW)_{cyi} = \frac{1}{\sum PIE_{cyi}} \sum_{j=1}^{PIE_{cyi}} \frac{\sum F}{\sum F + \sum M}$$

The indicator is calculated for each country (c), year (y) and FORD (i).

Source of data

Computed using Scopus ™ data and NamSor ™

2.8.11. ***Ratio of FWCI for women to men based on fractional authorship***

2.8.11.1. *Definition of indicator*

This indicator compares the fractional citation impact (based on field-weighted citation impact, FWCI) of publications by women to that of men on a publication by publication basis. Fractional calculation refers to distributing publication and citation counts equally across multiple authors; publication P with 2 authors would be counted as 0.5 publications for each author, and its citations would be shared equally between these authors. A value above 1 indicates that women are more represented on highly cited publications; a value below 1 indicates that men are more represented on highly cited publications. The ratio is based for the fractional FWCI for all women and men on a publication.

2.8.11.2. *Rationale*

Women publish fewer research papers on average than men (Elsevier, 2017) and women in a leading role of authorship receive fewer citations than in cases when a man was in one of these roles (Larivière et al., 2013) and this may have effects on the impact of their publications. Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research

evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men.

This indicator looks at the contribution of women and men authors to the mean FWCI in a given country and subject.

2.8.11.3. *Computation method*

FWCI is an indicator of citation impact of a publication based on the actual number of citations received by an article compared to the expected number of citations for articles of the same document type (article, review or conference proceeding paper), publication year and subject field. When an article is classified in two or more subject fields, the harmonic mean of the actual and expected citation rates is used. The indicator is therefore always defined with reference to a global baseline of 1 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document types (reviews typically attract more citations than research articles, for example) as well as subject-specific differences in citation frequencies overall and over time and document types.

In general, the Field-Weighted Citation Impact (FWCI) for a publication is defined as:

$$FWCI = \frac{C_j}{E_j}$$

where

C_j: citations received by publication i

E_j: expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than one discipline, the harmonic mean is used to calculate E_i.

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M)$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum A)$ Number of gendered authors in authorship byline. **Unit: Number.**

$(\sum P_{cyi})$ The number of publications in a given country (c), year (y) and FORD (i). **Unit: Number.**

$(FWCI)$ FWCI for given publication

Computation formula

For a given country (c), field (i) and year (y), the formula for ratio of FWCI for women to men based on fractional authorship (RFWtMFA) is:

$$fractional\ FWCI\ for\ women\ \ authors = \frac{\sum_{j=1}^{P_{cyi}}\left((FWCI) * (\sum F / \sum A)\right)}{\sum_{j=1}^{P_{cyi}}(\sum F / \sum A)}$$

$$fractional\ FWCI\ for\ men\ \ authors = \frac{\sum_{j=1}^{P_{cyi}}\left((FWCI) * (\sum M / \sum A)\right)}{\sum_{j=1}^{P_{cyi}}(\sum M / \sum A)}$$

$$(RFWtMFA) = \frac{\frac{\sum_{j=1}^{P_{cyi}}\left((FWCI) * (\sum F / \sum A)\right)}{\sum_{j=1}^{P_{cyi}}(\sum F / \sum A)}}{\frac{\sum_{j=1}^{P_{cyi}}\left((FWCI) * (\sum M / \sum A)\right)}{\sum_{j=1}^{P_{cyi}}(\sum M / \sum A)}}$$

Source of data

Computed using Scopus ™ data and NamSor ™

2.8.12. ***Compound annual growth rate (CAGR) of ratio of FWCI for women to men based on fractional authorship.***

2.8.12.1. *Definition of indicator*

This indicator presents the compound annual growth rate of the ratio of FWCI for women to men based on fractional authorship, meaning the average yearly percentage increase/decrease, year on year.

2.8.12.2. *Rationale*

Women publish fewer research papers on average than men (Elsevier, 2017) and women in a leading role of authorship receive fewer citations than in cases when a man was in one of these roles (Larivière et al., 2013) and this may have effects on the impact of their publications. Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men. This indicator assesses whether there has been any change in.

2.8.12.3. *Computation method*

Data needed

$(RFWtMFA)_{cyi}$ Ratio of FWCI for women to men based on fractional authorship in a given country (c), year (y) and FORD (i). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Computation formula

CAGR for ratio of FWCI for women to men based on fractional authorship = $(\mathrm{i})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$RFWtMFA_{\mathrm{Csi}}$ denotes the ratio of FWCI for women to men based on fractional authorship in a given country (c), and FORD (i) in the start year;

$RFWtMFA_{\mathrm{Cei}}$ denotes the ratio of FWCI for women to men based on fractional authorship in a given country (c), and field (i) in the end year.

Source of data

Computed using Scopus™ data and NamSor™

2.8.13. ***Ratio of corresponding authorships of women to men***

2.8.13.1. *Definition of indicator*

This indicator is the ratio of publications in which a woman is corresponding author to those in which a man is corresponding author. It is based on peer-reviewed scientific publications (articles, reviews, conference papers). A score above 1 indicates that women in a given country contribute more to the research output as corresponding author than men whereas a score below 1 means the opposite.

2.8.13.2. *Rationale*

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women.

This indicator looks at the contribution of women and men as corresponding author, to research across countries and fields of research and development (FORD).

2.8.13.3. *Computation method*

Data needed

$(\sum PCF_{cyi})$ The number of publications with a woman as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

$(\sum PCM_{cyi})$ The number of publications with a man as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and field (i), the formula for the indicator, Ratio of Corresponding authorship for Women to Men (RC) is:

$$(RC)_{cyi} = \frac{\sum PCF_{cyi}}{\sum PCM_{cyi}}$$

Source of data

Computed using Scopus™ data and NamSor™

2.8.14. *Compound annual growth rate (CAGR) of ratio of corresponding authorships of women to men*

2.8.14.1. *Definition of indicator*

Compound annual growth rate (CAGR) is defined as the year-over-year constant growth rate over a specified period of time. Starting with the first value in any series and applying this rate for each of the time intervals yields the amount in the final value of the series. Throughout the term CAGR is also referred to as '(yearly) growth rate.'

The indicator is calculated for the ratio of corresponding authorships of women to men.

2.8.14.2. *Rationale*

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women. This indicator looks at the growth rates of the scientific contribution of women as corresponding author compared to men as corresponding author across different countries and fields of research and development (FORD).

2.8.14.3. *Computation method*

Data needed

$(RC)_{ci}$ Ratio of corresponding authorships of women to men in a start and an end year in a given country (c) and FORD (i). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Computation formula

CAGR for ratio of corresponding authorships of women to men = $(\mathrm{RC_{Cei}}/\mathrm{RC_{Csi}})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{RC_{Csi}}$ denotes the ratio of Corresponding authorship for women to men in a given country (c), and FORD (i) in the start year;

$\mathrm{RC_{Cei}}$ denotes the ratio of Corresponding authorship for women to men in a given country (c), and FORD (i) in the end year.

Source of data

Computed using Scopus™ data and NamSor™

2.8.15. *Ratio of corresponding authorships of women to men on publications resulting from international collaborations*

2.8.15.1. *Definition of indicator*

This indicator is the ratio of publications resulting from international collaboration in which a woman is corresponding author to those in which a man is corresponding author. A score above 1 indicates that women in a given country contribute more to international collaborations as corresponding author than men whereas a score below 1 means the opposite.

2.8.15.2. *Rationale*

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women. Women are less likely than men to lead internationally collaborated research (Elsevier, 2017) and this may have effects on the impact of their publications (as measured by citations). Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men.

This indicator looks at the contribution of women and men as corresponding author on publications resulting from international collaboration, to research across countries and fields of research and development (FORD).

2.8.15.3. *Computation method*

International collaboration is defined as multi-authored research outputs, where at least one author is from an institution inside the country of interest and at least one author is from an institution outside the country of interest (or EU (for EU-27 and EU-28 calculations)).

Data needed

$(\sum PICF_{cyi})$ The number of publications resulting from international collaboration with a woman as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

$(\sum PICM_{cyi})$ The number of publications resulting from international collaboration with a man as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and FORD (i), the formula for the indicator, Ratio of Corresponding authorship for Women to Men on publications resulting from international collaborations (RC) is:

$$(RCI)_{cyi} = \frac{\sum PICF_{cyi}}{\sum PICM_{cyi}}$$

Source of data

Computed using Scopus™ data and NamSor™

2.8.16. *Ratio of corresponding authorships of women to men on publications resulting from national collaborations*

2.8.16.1. *Definition of indicator*

This indicator is the ratio of publications resulting from national collaboration in which a woman is corresponding author to those in which a man is corresponding author. A score above 1 indicates that women in a given country contribute more to national collaborations as corresponding author than men whereas a score below 1 means the opposite.

2.8.16.2. *Rationale*

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women. Women are less likely than men to lead internationally collaborated research (Elsevier, 2017) and this may have effects on the impact of their publications (as measured by citations). Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men. Statistics related to national collaboration can provide insight into whether there are barriers to authors collaborating outside of their own institution.

This indicator looks at the contribution of women and men as corresponding author on publications resulting from national collaboration, to research across countries and fields of research and development (FORD).

2.8.16.3. *Computation method*

National collaboration is defined as multi-authored research outputs, where authors are affiliated with more than one institution within the same country.

Data needed

$(\sum PNCF_{cyi})$ The number of publications resulting from national collaboration with a woman as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

$(\sum PNCM_{cyi})$ The number of publications resulting from national collaboration with a man as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and FORD (i), the formula for the indicator, Ratio of Corresponding authorship for Women to Men on publications resulting from national collaborations (RCN) is: $(RCN)_{cyi} = \frac{\sum PNCF_{cyi}}{\sum PNCM_{cyi}}$

Source of data

Computed using Scopus™ data and NamSor™

2.8.17. *Ratio of corresponding authorships of women to men on publications resulting from intra-EU27+ collaborations*

2.8.17.1. *Definition of indicator*

This indicator is the ratio of publications resulting from intra-EU27+ collaborations in which a woman is corresponding author to those in which a man is corresponding author. A score above 1 indicates that women in a given country contribute more to EU27+ collaborations as corresponding author than men whereas a score below 1 means the opposite.

2.8.17.2. *Rationale*

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women. Women are less likely than men to lead internationally collaborated research (Elsevier, 2017) and this may have effects on the impact of their publications (as measured by citations). Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men. Statistics related to intra-EU27+ collaboration can provide insight into whether there are barriers to authors collaborating outside of their own country.

This indicator looks at the contribution of women and men as corresponding author on publications resulting from intra-EU27+ collaboration, to research across countries and fields of science.

2.8.17.3. *Computation method*

Intra-EU27+ collaboration is defined as multi-authored research outputs, where authors are affiliated with institutions in more than one of the 44 countries (EU (for EU-27 and EU-28 calculations) and Associated Countries), but all authors are based within the 44 countries.

Data needed

$(\sum PIECF_{cyi})$ The number of publications resulting from Intra-EU27+ collaboration with a woman as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

$(\sum PIECM_{cyi})$ The number of publications resulting from Intra-EU27+ collaboration with a man as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and FORD (i), the formula for the indicator, Ratio of Corresponding authorship for Women to Men on publications resulting from Intra-EU27+ (RCIE) is: $(RCIE)_{cyi} = \frac{\sum PIECF_{cyi}}{\sum PIEM_{cyi}}$

Source of data

Computed using Scopus™ data and NamSor™

2.8.18. ***Ratio of FWCI for women to men based on corresponding authorship***

2.8.18.1. *Definition of indicator*

This indicator compares the average citation impact (based on field-weighted citation impact, FWCI) of publications with women as corresponding author to that of men. A value above 1 indicates that publications with women as corresponding author are more highly cited; a value below 1 indicates that publications with men as corresponding author are highly cited.

2.8.18.2. *Rationale*

Women publish fewer research papers on average than men (Elsevier, 2017) and women in a leading role of authorship receive fewer citations than in cases when a man was in one of these roles (Larivière et al., 2013) and this may have effects on the impact of their publications. Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men.

This indicator looks at the citations received by publications with women and men corresponding authors in a given country and subject.

2.8.18.3. *Computation method*

FWCI is an indicator of citation impact of a publication based on the actual number of citations received by an article compared to the expected number of citations for articles of the same document type (article, review or conference proceeding paper), publication year and subject field. When an article is classified in two or more subject fields, the harmonic mean of the actual and expected citation rates is used. The indicator is therefore always defined with reference to a global baseline of 1 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document types (reviews typically attract more citations than research articles, for example) as well as subject-specific differences in citation frequencies overall and over time and document types.

In general, the Field-Weighted Citation Impact (FWCI) for a publication is defined as:

$$FWCI = \frac{C_j}{E_j}$$

Where

C_j: citations received by publication i

E_j: expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than one discipline, the harmonic mean is used to calculate E_i.

Data needed

$(\sum PCF_{cyi})$ Number of publications with a woman as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

$(\sum PCM_{cyi})$ Number of publications with a man as corresponding author in a given country (c), year (y) and FORD (i). **Unit: Number.**

$(FWCI)$ FWCI for given publication

Computation formula

For a given country (c), year (y) and FORD (i), the formula for ratio of FWCI for women to men based on corresponding authorship (RFWtMCA) is:

$$FWCI\ for\ publications\ with\ women\ as\ corresponding\ authors = \frac{1}{\sum PCF_{cyi}} \sum_{j=1}^{PCF_{cyi}} (FWCI)$$

$$FWCI\ for\ publications\ with\ women\ as\ corresponding\ authors = \frac{1}{\sum PCM_{cyi}} \sum_{j=1}^{PCM_{cyi}} (FWCI)$$

$$(RFWtMCA) = \frac{\frac{1}{\sum PCF_{cyi}} \sum_{j=1}^{PCF_{cyi}} (FWCI)}{\frac{1}{\sum PCM_{cyi}} \sum_{j=1}^{PCM_{cyi}} (FWCI)}$$

Source of data

Computed using Scopus ™ data and NamSor ™

2.8.19. ***Compound annual growth rate (CAGR) of ratio of FWCI for women to men based on corresponding authorship***

2.8.19.1. *Definition of indicator*

This indicator presents the compound annual growth rate of the ratio of FWCI for women to men based on corresponding authorship, meaning the average yearly percentage increase/decrease, year on year.

2.8.19.2. *Rationale*

Women publish fewer research papers on average than men (Elsevier, 2017) and women in a leading role of authorship receive fewer citations than in cases when a man was in one of these roles (Larivière et al., 2013) and this may have effects on the impact of their publications. Funding agencies emphasize the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation (Jappelli et al., 2017), disadvantaging women in grant competitions with their

male counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men. This indicator assesses whether there has been any change in

2.8.19.3. *Computation method*

Data needed

$(RFWtMFA)_{cyi}$ Ratio of FWCI for women to men based on fractional authorship in a given country (c), year (y) and FORD (i). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Computation formula

CAGR for ratio of FWCI for women to men based on fractional authorship = $(RFWtMCA_{\mathrm{Cei}}/RFWtMCA_{\mathrm{Csi}})^{1/\mathrm{N}} - 1$

where

s refers to the start year

e refers to the end year

$RFWtMCA_{\mathrm{Csi}}$ denotes the ratio of FWCI for women to men based on corresponding authorship in a given country (c), and FORD (i) in the start year

$RFWtMCA_{\mathrm{Cei}}$ denotes the ratio of FWCI for women to men based on corresponding authorship in a given country (c), and FORD (i) in the end year.

Source of data

Computed using Scopus™ data and NamSor™.

2.8.20. ***Average proportion of women among authors on publications that list among the author affiliations, both a corporate entity and any other entity***

2.8.20.1. *Definition of indicator*

This indicator is the average proportion of women among authors on publications resulting from collaboration with a corporate entity. It is based on peer-reviewed scientific publications (articles, reviews, conference papers).

2.8.20.2. *Rationale*

- Mirrors indicator 3.2.2 "Public-private co-publications" in Innovation Union Scoreboard
- Directly relevant to the EU's Innovation Union Scoreboard and Open Innovation objectives
- Relevance for Digital Economy and Society policy objectives & Open Innovation

This indicator looks at the average proportion of women among authorship teams resulting from collaboration between a corporate entity and any other entity within countries and fields of research and development (FORD).

2.8.20.3. *Computation method*

Corporate collaboration is defined as publications that list among the author affiliations, both a corporate entity and any other entity (academic, government, or medical)

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum PC_{cys})$ The number of publications resulting from corporate collaboration in a given country (c), year (y) and FORD (i). **Unit: Number.**

Computation formula

For a given country (c), year (y) and FORD (i), the formula for average proportion of women among authors on corporate collaborated publications (ACPW) is:

$$(ACPW)_{cyi} = \frac{1}{\sum PC_{cyi}} \sum_{j=1}^{PI_{cyi}} \frac{\sum F}{\sum F + \sum M}$$

The indicator is calculated for each country (c), year (y) and FORD (i).

Source of data

Computed using Scopus ™ data and NamSor ™

2.8.21. ***Percent of a country's research output integrating a gender dimension in its research and innovation content (GDRIC)***

2.8.21.1. *Definition of indicator*

The indicator shows the proportion of peer-reviewed publications that integrate gender or sex-sensitive analysis and the impact of these publications, broken down by field and country.

2.8.21.2. *Rationale*

The European Commission seeks to promote the integration of the methods of gender analysis into research design and process as a way of preventing bias in research, promoting better quality of outcomes in S&T and achieving cross-cutting benefits. To assess whether the introduction of new policies has resulted in changes in the research landscape, this indicator will reveal changes over time in how much research:

- addresses gender issues;
- addresses male and female issues;
- addresses men and women issues;

- reflects in some way, consideration of both sexes as a proxy for gender dimension.

The above categories are by no means meant to be mutually exclusive. This assessment includes studies on non-human species, which can be models to study human conditions.

Bibliometric analyses are used to identify the body of research that explores 'gender dimension'. The European Commission created the Gendered Innovations/Innovation through Gender Expert Group, defining that research that explores 'gender dimension' integrates sex and gender analysis into research. 'Sex' refers to basic biological characteristics of females and males and 'gender' refers to cultural attitudes and behaviours that shape 'feminine' and 'masculine' behaviours, products, technologies, environments, and knowledge (European Commission, 2020d and European Commission, 2013b). The strengthening of the integration of the gender dimension into R&I content is one of the gender equality priorities set for Horizon Europe, the EU Framework Programme for Research and Innovation 2021-2027.

2.8.21.3. *Computation method*

Data needed

$\mathbf{P}_{cyi}$ Number of publications in a given country (c), year (y) and FORD (i). **Unit: Number**

$\mathbf{P}GDRIC_{cyi}$ Number of publications integrating GDRIC in a given country (c), year (y) and FORD (i). **Unit: Number**

Computation formula

For a given country (c), year (y) and FORD (i), the formula for this indicator is:

Percent of a country's publications integrating $GDRIC_{cyi}$ = $\frac{\mathbf{P}GDRIC_{cyi}}{\mathbf{P}_{cyi}}$

The aggregation over fields or countries is implicitly carried out by extending the range of fields or countries respectively over which the sums in the numerator and denominator extend. Even if a paper is assigned to more than one fields covered in the indicator, the algorithm counts it only once.

Source of data

Computed using Scopus™ data

2.8.21.4. *Specifications*

The bibliometric analysis is based on the following strategy:

1. Target research that mentions at least two sexes or genders (researchers often use 'gender' when in fact they mean 'sex') in the context of the abstract or title. This approach ensures identification of research in which the researchers have made some effort to compare characteristics/behaviour of females and males by including more than single sex or gender in the research rationale or output and therefore, the findings can be influenced by sex or gender-based variables.
2. Include research on non-human species in the search. Whilst discussions of the methods of analysis of gender dimension in research have been so far focused on

human context, it has been recognised that sex, in particular, can play an important role in controlling the 'lifepath' of non-human species (for example plants, animals or cells) that make up the natural ecosystems in which humans coexist and whose wellbeing they influence. Therefore, and especially in the context of societal and environmental challenges, such as those linked to the effects of climate change, it is important to also promote methods of sex/gender analysis in research more widely.

3. Exclude research which applies to only a single-sex or gender as a recognition that while some research may apply to only single-sex out of necessity (for example – pregnancy or prostate research), a large body of research in animals contains persistent historical male bias, thereby excluding females and biasing results (Berry and Zucker, 2011). This approach avoids including this latter body of animal research in which sex and gender dimension have not been considered in the research design.

For the categories listed above, that is, to assess research that accounts for both sexes in study design as a proxy for GDRIC, publications were identified that mention at least two sexes or genders in the context of the abstract or title. This approach has been chosen as a means to identify research in which the researchers have made some effort to compare characteristics/behaviour of women/females and men/males by including more than a single-sex or gender in the research rationale or output. Research on non-human species is included in the search.

Creation of datasets for gender dimension in research and innovation content

The publications included in the assessment were retrieved using any of the following eight queries:

i	(({men} OR {man} OR boy* OR {male} or mascul*) AND ({women} OR {woman} OR girl* OR {female} OR femin*))
ii	{sexual dimorphism}
iii	(sexual* AND dimorph*)
iv	"gender dimension"
v	"gender difference"
vi	{sex ratio}
vii	"sex difference"
viii	"sexual reproduction"

The assessment of the validity of each query used to create the publication set was done with input from gender experts selected by Portia Inc. These experts validated a set of 100 randomly selected publications retrieved by each query (eight queries in total) to generate a false positive rate. The false positive rate for each of the eight queries was less than 1%. Therefore, all eight queries were used in union to generate publication set.

2.9. EU Open Data Portal

Content-based rationale

The measurement of the extent to which research incorporates a gender dimension is not limited to research results published in peer-reviewed publications. A new indicator measures the proportion of projects stemming from Horizon 2020, European Union's

Framework Programme for Research and Innovation (2014-2020), which are integrating gender as part of their content.

In addition, a second indicator measures the proportion of Horizon 2020 projects, which, on top of integrating gender, are also integrating intersectional aspects in their content, i.e., they are seeking to improve understanding of how gender intersects with other characteristics (such as age, disability status and ethnicity) to affect experiences of disadvantage and discrimination.

Broad overview of the source

The indicators presented in this section were computed by Elsevier using Horizon 2020 data, retrieved from the EU Open Data Portal (https://data.europa.eu/data/datasets/cordish2020projects?locale=en).

The EU Open Data Portal contains all project deliverables and other publications linked to Horizon 2020 projects. This dataset contains projects and organisations funded by the European Union under the Horizon 2020 Framework Programme for Research and Innovation from 2014 to 2020. Various files available from the EU Open Data Portal have been used, linked by the project ID. However, only files containing report summaries, projects, publications and project deliverables were included in this process. It is important to note that since a limited number of files were checked, the calculated values may underestimate the actual integration of the gender dimension in R&I content in Horizon 2020 projects.

2.9.1. ***Gender dimension in research and innovation content in Horizon 2020 projects***

2.9.1.1. *Definition of indicator*

This indicator calculates the proportion of Horizon 2020 projects that integrate a gender dimension, broken down by country.

2.9.1.2. *Rationale*

The European Commission seeks to promote the integration of the methods of gender analysis into research design and process as a way of preventing bias in research, promoting better quality of outcomes in S&T and achieving cross-cutting benefits. To assess whether the introduction of new policies has resulted in changes in the research landscape, this indicator will reveal changes over time in how much projects stemming from Horizon 2020, European Union's Framework Programme for Research and Innovation (2014-2020), are integrating gender as part of the project content.

To ensure consistency with the indicator on the percent of a country's publication integrating a gender dimension in research and innovation content, the same bibliometric approach was used. Bibliometric analyses are used to identify Horizon 2020 projects that explore 'gender dimension.' The European Commission created the Gendered Innovations/Innovation through Gender Expert Group, defining that research that explores 'gender dimension' integrates sex and/or gender analysis into research. 'Sex' refers to basic biological characteristics of females and males and 'gender' refers to cultural attitudes and behaviours that shape 'feminine' and 'masculine' behaviours, products, technologies, environments, and knowledge (European Commission, 2020d and European Commission, 2013b). The strengthening of the integration of the gender

dimension into R&I content is one of the gender equality priorities set for Horizon Europe, the EU Framework Programme for Research and Innovation 2021-2027.

2.9.1.3. *Computation method*

Data needed

$\mathbf{HP}GDP_{cy}$ Number of Horizon 2020 projects, tagged as human or animal as the research subject, integrating a gender dimension in a given country (c) and year (y). **Unit: Number**

$\mathbf{HP}_{CY}$ Number of Horizon 2020 projects, tagged as human or animal as the research subject, in a given country (c) and year (y). **Unit: Number**

Computation formula

For a given country (c) and year (y), the formula for this indicator is:

Percent of a country's projects integrating GD_{cy} = $\frac{\mathbf{HP}GD_{cy}}{\mathrm{HP}_{cy}}$

Source of data

Computed using Horizon 2020 data, retrieved from the EU Open Data Portal (https://data.europa.eu/data/datasets/cordish2020projects?locale=en)

2.9.1.4. *Specifications*

The bibliometric analysis is based on the following strategy:

For each project, tag the project based on available text fields (like abstracts, titles, objectives, results etc.):

1. Research subject: classify as human, animal, or neither.
2. Include research on non-human species in the search. Whilst discussions of the methods of analysis of gender dimension in research have been so far focused on human context, it has been recognised that sex, in particular, can play an important role in controlling the 'lifepath' of non-human species (for example plants, animals or cells) that make up the natural ecosystems in which humans coexist and whose wellbeing they influence. Therefore, and especially in the context of societal and environmental challenges, such as those linked to the effects of climate change, it is important to also promote methods of sex/gender analysis in research more widely.
3. Exclude research which applies to only a single-sex or gender as a recognition that while some research may apply to only single-sex out of necessity (for example – pregnancy or prostate research), a large body of research in animals contains persistent historical male bias, thereby excluding females and biasing results (Berry and Zucker, 2011). This approach avoids including this latter body of animal research in which sex and gender dimension have not been considered in the research design.

For the categories listed above, that is, to assess research that accounts for both sexes in study design as a proxy for projects integrating a gender dimension were identified that mention at least two sexes or genders in the context of the abstract or title. This approach has been chosen as a means to identify research in which researchers have

made some effort to compare characteristics/behaviour of women/females and men/males by including more than a single-sex or gender in the research rationale or output. Research on non-human species is included in the search.

Creation of datasets for gender dimension in research and innovation content

The publications included in the assessment were retrieved using any of the following eight queries:

i	(({men} OR {man} OR boy* OR {male} or mascul*) AND ({women} OR {woman} OR girl* OR {female} OR femin*))
ii	{sexual dimorphism}
iii	(sexual* AND dimorph*)
iv	"gender dimension"
v	"gender difference"
vi	{sex ratio}
vii	"sex difference"
viii	"sexual reproduction"

The same query was used for indicator "Percent of a country's research output integrating a gender dimension in its research and innovation content (GDRIC)".

2.9.2. ***Percentage of country's Horizon 2020 projects integrating intersectional aspects, 2014-2020***

2.9.2.1. *Definition of indicator*

This exploratory indicator analyses the text fields used for the indicator on gender dimension in research and innovation content in Horizon 2020 projects and combines the results retrieved from this indicator with search queries on intersectional aspects, broken down by country.

2.9.2.2. *Rationale*

This indicator is relevant for EU's wider objective to improve understanding of how gender intersects with other characteristics (such as age, disability status and ethnicity) to affect experiences of disadvantage and discrimination (Gender Equality Strategy 2020-2025). To assess whether the introduction of new policies has resulted in changes in the research landscape, this indicator will reveal changes over time in how much projects stemming from Horizon 2020, European Union's Framework Programme for Research and Innovation (2014-2020), are integrating intersectional aspects as part of the project content.

This indicator builds on the approach for the gender dimension in research and innovation content in Horizon 2020 projects. Bibliometric analyses are used to identify Horizon 2020 projects that explore 'gender dimension.' The European Commission created the Gendered Innovations/Innovation through Gender Expert Group, defining that research that explores 'gender dimension' integrates sex and gender analysis into research. 'Sex' refers to basic biological characteristics of females and males and 'gender' refers to cultural attitudes and behaviours that shape 'feminine' and 'masculine' behaviours, products, technologies, environments, and knowledge. Based on the query to identify

gender aspects, additional queries for intersectional aspects were applied. The Gendered Innovations 2 report developed and highlighted a methodology for intersectionality in research, which was applied for this indicator.

2.9.2.3. *Computation method*

Data needed

$\mathbf{HP}GDP_{cy}$ Number of Horizon 2020 projects, tagged as human or animal as the research subject, integrating an intersectional aspect in a given country (c). **Unit: Number**

$\mathbf{HP}_{CY}$ Number of Horizon 2020 projects, tagged as human or animal as the research subject, in a given country (c). **Unit: Number**

Computation formula

For a given country (c), the formula for this indicator is:

Percent of a country's projects integrating intersectional aspects GDint_c = $\frac{\mathbf{HP}int_c}{\mathrm{HP}_c}$

Source of data

Computed using Horizon 2020 data, retrieved from the EU Open Data Portal (https://data.europa.eu/data/datasets/cordish2020projects?locale=en)

2.9.2.4. *Specifications*

The bibliometric analysis is based on the strategy following the gender dimension in research and innovation content for Horizon 2020 projects:

For each project, tag the project based on available text fields (like abstracts, titles, objectives, results etc.):

1. Horizon projects integrating a gender dimension were identified using the approach outlined in the indicator Gender dimension of research content in Horizon 2020 projects.
2. The resulting Horizon 2020 projects were again queried using a shortlist of keywords from the Gendered innovations 2 report (European Commission, 2020d). The keywords used for the queries were:
    i. "intersectional*"
    ii. "disabilit*"
    iii. "ethnic"
    iv. "LGBT*"
    v. "race" OR "racis*"
    vi. "socio-economic"
    vii. "religion"
    viii. "belief"
    ix. "class"
    x. "social origin"

xi. "sexual orientation"

xii. "vulnerable group" OR "vulnerable population"

3. It is acknowledged that this shortlist is not extensive and requires additional work, but it serves as a first entry point into addressing intersectional aspects in Horizon 2020 projects.

2.10. EPO Worldwide Patent Statistical Database (PATSTAT)

Content-based rationale

Women have been shown to lag behind men in terms of the size (as measured by the number of peer-reviewed scientific publications) and impact (as measured by citations to their publications) of their scientific production, as well as their propensity to partner on an international scale (as measured by the proportion of papers co-authored by researchers located in at least two countries) (DG Research and Innovation, 2016 and DG Research and Innovation, 2019b). Because of the emphasis placed by funding agencies on the above dimensions in the evaluation of research proposals, women could be disadvantaged in grant competitions relative to their men counterparts. In grant competitions focusing more heavily on applied research, the number of patent applications in which a researcher is listed as an inventor might also prove to be a decisive factor in the funding decision. Thus, techno metric indicators derived from PATSTAT are integrated in the She Figures publication to monitor gaps in the contribution of women and men to the production of inventions by country, year and technological fields.

Broad overview of the source

The indicators presented in this section were computed using raw bibliographic data derived from the European Patent Office (EPO) Worldwide Patent Statistical Database (PATSTAT). PATSTAT covers patent data from over 150 offices worldwide, including EPO, the United States Patent and Trademark Office (USPTO) and the Japan Patent Office (JPO). The USPTO covers the United States, the EPO covers Europe, the JPO covers Japan, and so forth. For the She Figures publication, the statistics are based on the EPO within PATSTAT, as the European market is one of the largest in the world and certainly the most relevant in the context of the She Figures publication, since it covers all countries associated with the European Research Area.

Note that statistics on inventorships can be produced by measuring issued patents or patent applications when working with EPO data. On a conceptual level, if the goal is to get a sense of the inventive/innovative capacity of a given entity (e.g., women in a given country) rather than of 'marketable/innovative outputs', as in this study, then applications are more appropriate. Furthermore, in cases where trends in the inventiveness of entities are to be investigated, also as in this study, the capacity to produce timely data are important. In this regard, issued patents have the disadvantage of running behind and becoming visible only years after the innovative activity has taken place. Thus, from a methodological standpoint, applications are still preferable. Consequently, EPO patent applications (kind codes: A1 and A2) were retained in computing this set of indicators.

Definitions and preliminary data treatment

A European patent application is characterized by one or more applicants, inventors and classified according to the different areas of technology to which they pertain.

The applicants are the natural and/or legal person and/or entity that filed the application. The inventor is the real creator of the invention. In EU this term is defined by each Member State's legislation. In general, in order for one to be considered an inventor, it is acknowledged that a certain level of contribution to the development of the creative

elements of an invention (technical creativity) must be met. Inventors are always private individuals and are always entitled to be designated on the patent, regardless of who files the application. Joint inventors or co-inventors exist when a patentable invention is the result of the inventive work of more than one inventor, even if they did not contribute in equal parts.

Concerning the technological classification system, this is based on the International Patent Classification (IPC) system used in over 100 countries to identify the content of patents in a uniform manner. It was created under the Strasbourg Agreement (1971).[33] The classification is updated on a regular basis (the 1st of January of each year) by a Committee of Experts, consisting of representatives of the Contracting States of that Agreement with observers from other organisations, such as the European Patent Office. The first level of the IPC hierarchy permits to identify eight sections; in particular:

A. Human Necessities
B. Performing Operations, Transporting
C. Chemistry, Metallurgy
D. Textiles, Paper
E. Fixed Constructions
F. Mechanical Engineering, Lighting, Heating, Weapons and Blasting
G. Physics
H. Electricity.

PATSTAT data are organised in a relational database, consisting in different tables that can be merged by means of specific keys. In this case, all the applications were considered with a first filing year from 2005 to 2018, attributing to each of these a binary variable (0=No/1=Yes) that specifies the IPC section to which it belongs. In a second step the application was attributed to the country of the first applicant (despite such information was not always filled).

As a last step, all the inventors referred to each application were extracted, to which a specific strategy had already attributed the sex.

The algorithm is based on a standard methodology offered by GendRE API, a package developed by NamSor™ (a European designer of name recognition software committed to promoting diversity and equal opportunity); in fact, this API permits to extract the sex from personal names, surnames and countries.

2.10.1. ***Ratio of women to men inventorships***

2.10.1.1. *Definition of indicator*

This indicator is the ratio of women to men inventorships, or equivalently, the ratio of the proportion of women inventorships (in total inventorships) compared to the equivalent proportion for men. The absolute number of inventorships used in computing this indicator is based on fractionalised counts of patent applications across their

[33] Summary of the Strasbourg Agreement concerning the International Patent Classification (1971) : https://www.wipo.int/treaties/en/classification/strasbourg/summary_strasbourg.html

corresponding inventors: for example, if a patent application involves 10 inventors, each inventor is attributed an equal fraction of the inventorships (i.e., 1/10 of the invention). A score above 1 indicates that women in a given country produced a larger proportion of the country's inventions than men, whereas a score below 1 means the opposite.

2.10.1.2. *Rationale*

Women still lag behind men in terms of their measurable scholarly output (Larivière et al., 2013, DG Research and Innovation, 2016 and DG Research and Innovation, 2019b). Given the increasing reliance on bibliometric statistics (i.e., statistical analyses of written publications such as books or articles) for research evaluation purposes in research assessment exercises and grant competitions, the lower scientific output of women could lead to reduced chances of being funded (or the receipt of lower funding amounts), which could in turn decrease their scientific output, thereby creating a vicious circle. In grant competitions focusing more heavily on applied research, the number of patent applications on which a researcher is listed as an inventor might also prove to be a decisive factor in the funding decision. This indicator looks at the size of the technological output of women compared to men across different countries and fields of technology.

2.10.1.3. *Computation method*

Once having identified the sex of the inventors of each application (following the approach described in the general part of this section), it is possible to obtain a dataset that constitutes the base of the computation of this indicator that will be exploited for each year, country and by crossing this information with the IPC classifications.

Data needed

(WI_{cyi}) Sum of fractionalised inventorships for women in a given country (c), year (y) and section (i, based on the International Patent Classification [IPC]). **Unit: Total of fractionalized counts**.

(MI_{cyi}) Sum of fractionalised inventorships for men in a given country (c), year (y) and IPC section (i). **Unit: Total of fractionalized counts**.

(TI_{cyi}) Sum of fractionalized inventorships across women and men in a given country (c), year (y) and IPC section (i). **Unit: Total of fractionalized counts**.

(N_{cyi}) Total number of fractionalized inventorships in a given country (c), year (y) and IPC section (i). **Unit: Total of fractionalized counts**.

Source of data

Computed using PATSTAT data

Computation formula

Ratio of inventorships for Women to Men, for a given country (C), year (Y) and IPC section

$$(I) = \frac{WI_{cyi}}{TI_{cyi}} \Big/ \frac{MI_{cyi}}{TI_{cyi}} = \frac{WI_{cyi}}{MI_{cyi}}$$

2.10.1.4. *Specifications*

All EPO patent applications are classified based on the **International Patent Classification (IPC) Version 2017.01** of the World Intellectual Property Organization (WIPO) in PATSTAT. This hierarchical classification is divided into eight sections (Level 1), which are further divided into classes (Level 2), subclasses (Level 3), main groups (Level 4) and subgroups (lower level). This classification is not mutually exclusive (i.e., each patent application is classified into one or more sections, classes, subclasses, main groups and subgroups). Thus, a given patent application can contribute to the scores of more than one of the eight IPC sections for which this indicator has been computed, in addition to the total for all EPO patent applications (subscript I in the above formula):

A. Human Necessities
B. Performing Operations, Transporting
C. Chemistry, Metallurgy
D. Textiles, Paper
E. Fixed Constructions
F. Mechanical Engineering, Lighting, Heating, Weapons and Blasting
G. Physics
H. Electricity

(T) Total across all sections including unclassified patent applications (unique/distinct count of patent applications across sections).

2.10.1.5. *Comments and critical issues*

Although the data covers all the patent applications registered in the considered years, it is necessary to observe that for some of these it was not possible to correctly derive the sex of all inventors. This can then be associated to a **confidence interval (CI),** for a given country, year and IPC classification. The CI takes into account the total number of inventorships (including those with an unclassified sex of the inventors), N_{cyi} and considers those with classified sex of all inventors as a random sample.

The 90% CI is in the interval:

$$RWMI \pm \frac{\frac{WI_{CYI}}{TI_{CYI}} - \frac{1.645\sqrt{\frac{N_{CYI}-TI_{CYI}}{N_{CYI}-1}}\sqrt{\frac{\frac{WI_{CYI}}{TI_{CYI}}\left(1-\frac{WI_{CYI}}{TI_{CYI}}\right)}{TI_{CYI}}}}{\frac{WI_{CYI}}{TI_{CYI}}}}{\frac{MI_{CYI}}{TI_{CYI}} - \frac{1.645\sqrt{\frac{N_{CYI}-TI_{CYI}}{N_{CYI}-1}}\sqrt{\frac{\frac{MI_{CYI}}{TI_{CYI}}\left(1-\frac{MI_{CYI}}{TI_{CYI}}\right)}{TI_{CYI}}}}{\frac{MI_{CYI}}{TI_{CYI}}}}$$

Because the confidence intervals of some of the smaller countries were sometimes relatively large on a yearly basis, due to the size of the available samples by IPC section, the ratios were computed using **a four-year period** (2015–2018). This way, the samples used were larger, providing estimates that are more robust.

Potential errors in coverage mainly refer to:

- Bias in the number of documents over time: there are applications for which the name of the inventor is missing. Despite this phenomenon, the available and correct sample for She Figures 2018 (DG Research and Innovation, 2019b) was sufficiently large (more than 50 million of patents) to produce accurate statistics over the entire time frame covered in the publication. A similar picture emerges from the PATSTAT, Autumn 2020 data.
- Bias in favour of some countries: more applications could be attributed to some countries and less to others. However, since the indicators are ratios of variables referring to the same given country it is expected that such bias do not affect the cross-country comparability of them.
- Bias in favour of disciplines: it was observed that the proportion of EPO patent applications for which the affiliation country and full given name is available for all inventors on them was high and relatively similar across IPC classes. It was therefore concluded that none of the IPC classes contributed significantly more or less than the others to the women-to-men ratios computed at higher aggregation levels (e.g., for IPC Sections or for all EPO patent applications). A similar picture emerges from the PATSTAT, autumn 2020 data.

2.10.2. ***Compound annual growth rate (CAGR) of the proportion of women inventorships***

2.10.2.1. *Definition of indicator*

This indicator presents the compound annual growth rate of the proportion of women inventorships meaning the average yearly percentage increase/decrease in the proportion, moving from one period to the next (using four-year moving periods, e.g., 2006–2009, 2007–2010 and so on, until 2015-2018), year on year.

2.10.2.2. *Rationale*

She Figures 2018 has shown that women still lag behind men in terms of measurable scholarly output (DG Research and Innovation, 2019b). Given the increasing reliance on bibliometric statistics (i.e., statistical analyses of written publications such as books or articles) for research evaluation purposes in research assessment exercises and grant competitions, the lower scientific output of women could lead to reduced chances of being funded or the receipt of lower funding amounts, which could in turn decrease their scientific output, thereby creating a vicious circle. This indicator looks at the size of the scientific output of women compared to men across different countries and IPC sections.

2.10.2.3. *Computation method*

<u>Data needed</u>

(F_s) Estimated proportion in the start period. **Unit: Unitless.**

(F_e) Estimated proportion in the end period. **Unit: Unitless.**

(N) Number of years in the reference period (i.e., last year of end period – last year of start period). **Unit: Year.**

Source of data

Computed using PATSTAT data

Computation formula

CAGR for the proportion of women inventorships = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

2.10.3. ***Patent applications, by type of sex composition of inventors' team***

2.10.3.1. *Definition of indicators*

The indicators analyse the sex composition of the inventors' team for each patent application (e.g., by teams consisting of females only) and are the proportions of each type of composition and ratios of such proportions over each other.

2.10.3.2. *Rationale*

A patent is a legal title granting its holder the right, in specific countries and for a certain period, to prevent third parties from exploiting an invention for commercial purposes without authorisation. Any legal entity (one or more individuals and/or firms) could register a patent; when doing such action, it is mandatory to specify also the related inventors (which could be different from the entity that applies for the patent).

Thus, each patent application can have one named inventor (a lone/individual inventor) or multiple inventors (working collaboratively as part of a team). The determination of the sex of each named inventor permits to identify mutually exclusive sets of applications, i.e., those referred to a working-alone female (or male), those developed by teams of the same sex and those referred to mixed-sex teams.

The indicators shed light on the propensity of the two sexes to work alone or in same-sex teams versus working in mixed-sex teams as well as on how such collaboration patterns vary between countries and evolve over time.

2.10.3.3. *Computation method*

An application is characterised by a reference year (*y*) and a list of applicants. The latter may be legal entities and may involve different persons than the inventors.

In case of multiple applicants, the order in which they were specified in the application is maintained also in the collected data. To associate an application to a specific country (*c*), it is considered the country of residence of the first (main) applicant.

Once the sexes of its inventors have been identified (see the general introduction to this section), it is possible to compute the following variables for each *y* and *c*:

- number of applications from female/male inventors working alone: WI_{cy} or MI_{cy}
- number of applications from teams with inventors of the same sex: $allWI_{cy}$ or $allMI_{cy}$
- number of applications from teams which consist predominantly of women: pWI_{cy}. These are teams with more as 60% women.

- number of applications from teams which consist predominantly of men: pMI_{cy}. These are teams with more as 60% men.
- number of applications from sex-balanced teams: bTI_{cy}. These are teams with exactly or between 40% and 60% women.

All the above situations are mutually exclusive, so that their sum corresponds to the total number of applications for each year (*y*) and country (*c*):

$$WI_{cy} + MI_{cy} + allWI_{cy} + allMI_{cy} + pWI_{cy} + pMI_{cy} + bTI_{cy} = TOT_{cy}$$

The indicators able to synthesize the phenomenon consider comparisons between 'similar' variables or with the total number of applications. It has to be noted that the IPC section was ignored when estimating this indicator, in order to avoid issues due to small samples for country/year combinations.

All the indicators were referred to TOT_{cy}; in particular:

- proportion of applications from female/male inventors working alone: WI_{cy}/TOT_{cy} or MI_{cy}/TOT_{cy}
- proportion of applications from teams with inventors of the same sex: $allWI_{cy}/TOT_{cy}$ or $allMI_{cy}/TOT_{cy}$
- proportion of applications from teams which consist predominantly of women: pWI_{cy}/TOT_{cy}. These are teams with more than 60% women.
- proportion of applications from teams consisting predominantly of men: pMI_{cy}/TOT_{cy}. These are teams with more than 60% men.
- proportion of applications from sex-balanced teams: bTI_{cy}/TOT_{cy}. These are teams with exactly or between between 40% and 60% women.

Data needed

(WI_{cy}) Number of applications attributed to the country (c) and year (y) from female inventors working alone. **Unit: Count**.

(MI_{cy}) Number of applications attributed to the country (c) and year (y) from male inventors working alone. **Unit: Count**.

($allWI_{cy}$) Number of applications attributed to the country (c) and year (y) from teams of female inventors. **Unit: Count**.

($allMI_{cy}$) Number of applications attributed to the country (c) and year (y) from teams of male inventors. **Unit: Count**.

(pWI_{cy}) Number of applications attributed to the country (c) and year (y) from predominantly female teams. **Unit: Count**.

(pMI_{cy}) Number of applications attributed to the country (c) and year (y) from predominantly male teams. **Unit: Count**.

(pTI_{cy}) Number of applications attributed to the country (c) and year (y) from sex-balanced teams. **Unit: Count**.

The ratios were computed using **a four-year period** (2015–2018). This way, the samples used were larger, providing more robust estimates.

Source of data

Computed using PATSTAT database.

2.10.3.4. *Comments and critical issues*

Potential errors in coverage mainly refer to:

- Bias in the number of documents over time: there are applications for which the name of the inventor is missing. Despite this phenomenon, the available and correct sample for She Figures 2018 (DG Research and Innovation, 2019b) was sufficiently large (more than 50 million of patents) to produce accurate statistics over the entire time frame covered in the publication. A similar picture emerges from the PATSTAT, autumn 2020 data.
- Bias in favour of some countries: more applications could be attributed to some countries and less to others. However, since the indicators are ratios of variables referring to the same given country it is expected that such bias do not affect the cross-country comparability of them.
- Bias in favour of disciplines: it was observed that the proportion of EPO patent applications for which the affiliation country and full given name is available for all inventors on them was high and relatively similar across IPC classes. It was therefore concluded that none of the IPC classes contributed significantly more or less than the others to the women-to-men ratios computed at higher aggregation levels (e.g., for IPC Sections or for all EPO patent applications). A similar picture emerges from the PATSTAT, autumn 2020 data.

2.10.4. ***Compound annual growth rate (CAGR) of types of inventors' team***

2.10.4.1. *Definition of indicators*

This indicator presents the compound annual growth rate of the types of working teams, representing the average yearly percentage of increase/decrease in the related proportions. To obtain more robust estimation, this indicator is built considering a four-year moving period, e.g., 2014–2017, 2015–2018.

2.10.4.2. *Rationale*

A patent is a legal title granting its holder the right, in specific countries and for a certain period, to prevent third parties from exploiting an invention for commercial purposes without authorisation. Any legal entity (one or more individuals and/or firms) could register a patent; when doing such action, it is mandatory to specify also the related inventors (which could be different from the entity that applies for the patent).

Thus, each patent application can have one named inventor (a lone/individual inventor) or multiple inventors (working collaboratively as part of a team). The determination of the sex of each named inventor permits to identify mutually exclusive sets of applications, i.e., those referred to a working-alone female (or male), those developed by teams of the same sex and those referred to mixed-sex teams.

The indicators shed light on the propensity of the two sexes to work alone or in same-sex teams versus working in mixed-sex teams as well as on how such collaboration patterns vary between countries and evolve over time.

2.10.4.3. *Computation method*

Data needed

(F) Estimated proportions of each type of working team in a start and an end year. **Unit: Unitless.**

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number.**

Source of data

Computed using PATSTAT data

Computation formula

CAGR for each type of working team = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{F_s}$ denotes the estimated proportion of each type of working team in the start year;

$\mathrm{F_e}$ denotes the estimated proportion of each type of working team in the end year.

The ratios were computed using **four-year moving periods** (e.g., 2005–2008, 2006-2009, …, 2013–2016). This way, the samples used were larger, providing estimates that are more robust.

3. QUALITY PLAN: VERIFICATION AND VALIDATION OF DATA

In preparing the present study, data quality was viewed as a multi-faceted concept. The quality framework suggested covered three different dimensions to be considered in selecting indicators: relevance, accuracy and availability (Table 3). Each indicator was to be evaluated by grading it for each dimension and by an overall assessment.

The relevance of an indicator was determined by a qualitative assessment of the value contributed by that indicator in terms of its policy relevance. An indicator had to be policy relevant by addressing key policy issues related to gender inequalities in the EU Member States and Associated Countries.

The accuracy of an indicator is the degree to which the indicator correctly estimates or describes the quantities or characteristics it is designed to measure. Accuracy has two dimensions: the data collection method and the degree of cross-country standardisation. The data collection method was considered sound if the data correctly estimated or described the quantities or characteristics that it was designed to measure. Thus, accuracy based on the data collection method refers to the closeness between the values provided and the (unknown) true value.

The evaluation of the accuracy of data collection methods was significant in the present study, given that the data used had to be collected not only from high-quality databases of national statistical offices and international organisations but also from other databases held by the European Commission and some of its agencies, as well as Statistical Correspondents. The latter may not have undergone formal quality reviews by statistical authorities. The accuracy of data collection methods in the present study can be evaluated as being very good, good and acceptable.

The other dimension of data accuracy was cross-country comparability: whether an indicator was comparable across countries required consideration as to the methods of data collection in the countries concerned. For example, an indicator was comparable if the same question was asked in all countries in the same way and by the same means. It was desirable to have the highest degree of comparability across countries. For data collected through the Statistical Correspondents, guidelines were prepared to maximise cross-country comparability. Metadata were also collected with the questionnaire from every participating country to allow an assessment of comparability. Additionally, much attention has been paid to ensuring data quality by regular consultation with Statistical Correspondents throughout the process of data gathering and input.

The concept of availability related to the accessibility of a given indicator in various countries and for a given time frame. It was desirable to have data from as many countries as possible, including the EU Member States and Associated Countries. In addition, an indicator that was available beyond the initial benchmark year was considered better than one that was available for only one year.

At the same time, ensuring maximum quality and reliability of the resulting data warehouse requires *a posteriori* verification and validation of the data received. The diagnosis of the accuracy and reliability of databases evolves over time, along with their content. In She Figures 2021, this diagnosis is primarily based on two approaches, which are further explained in this section.

Table 3 Dimensions of the data quality framework

	Depends on	Addressed by
RELEVANCE	• Relevance of selected indicators in the current models and measuring systems of research and innovation: are they up-to-date from a content/policy perspective?	• Steering group discussions • Mapping of state of the art with respect to R&I indicators • Identification of new indicators to introduce in She Figures 2021
ACCURACY OF DATA COLLECTION METHOD/COMPARABILITY	• Alignment between countries in reporting system, classifications used, etc., by data source	• Rely whenever possible on existing official classifications and manuals for data collection (e.g., Frascati Manual, etc.), international standards, etc. • Guidelines and aiming to have Statistical Correspondents adhere as much as possible to quality standards of data collection • Metadata sheets (to systematically register potential deviations from the defined classifications and standards) • Validity/coherence checks after data gathering and computation of confidence intervals (for certain indicators)
AVAILABILITY	• Capacity and resources of governments to collect the required information • Availability of secondary source databases	• Steering group discussions • Lessons from previous rounds • Flagging system (to systematically register missing data)

Coherence checks

For data broken down in categories, totals are also available. Categories are defined regarding:

- Sex
- Age groups (see later in this section 'Additional data considerations')
- Institutional sectors (see main grouping as defined by the Frascati Manual)
- NACE activities (see NACE Rev. 2.0 categories under Section 2.4.11)
- Fields of Research and Development (see 'Fields of Research and Development' in Section 2.4.6 as well as in pages 95 and 96 of Frascati Manual, OECD, 2015)
- Education levels and fields (see categories in Section 2.1.4) of education see UNESCO (2012) and UNESCO (2014)
- Grades (see definition of grades in Annex 2)
- ISCO-08 categories (see definition of Major Groups in International Labour Organization, 2012)
- R&D personnel categories
- Countries
- Years

Having available both the broken-down data and the totals allows for coherence checks by comparing provided totals with the sum of provided data by categories. For example, a check which can be performed on most of the tables is that the sum of the values for women (w) and men (m) should correspond to the reported totals (t), hence the verification is done by applying the definition $t = w + m$. A similar data verification procedure is followed to assess whether reported totals correspond to the sums of breakdowns at the level of the above-mentioned categorisations. Table 4 below shows the details of these coherence checks of WiS data.

Table 4 List of coherence checks on the WiS data

Data	Verification Description	Verification Formula
Tables T1 (HC) and T2 (FTE): RESEARCHERS and ACADEMIC STAFF BY SEX, GRADE AND MAIN FIELD OF RESEARCH AND DEVELOPMENT AND AGE GROUP (FOR LATEST YEAR ONLY)	Check Total sexes	{Men} + {Women} = Total
	Check Total age groups	{<35} + {35–44} + {45–54} + {55+} = Total
	Check Total fields of Research and Development	{NS} + {ET} + {MS} + {AS} + {SS} + {H} + {Unknown} = Total per grade
	Check Total grades	{A} + {B} + {C} + {D} = ALL grades
Tables T3 (Team leaders) and T4 (Team members): APPLICANTS AND BENEFICIARIES OF PUBLICLY MANAGED RESEARCH FUNDS BY SEX, MAIN FIELD OF RESEARCH AND DEVELOPMENT, AMOUNT APPLIED FOR AND AMOUNT RECEIVED	Check Total sexes	{Men} + {Women} = Total
	Check Total fields of Research and Development	{NS} + {ET} + {MS} + {AS} + {SS} + {H} + {MU} + {Unknown} = Total per fund
	Check Comparison applicants and beneficiaries	# applicants > # beneficiaries
Tables T5 (scientific boards) and T6 (administrative / advisory boards): PRESIDENTS / LEADERS AND MEMBERS OF BOARDS BY SEX, FIELD OF RESEARCH AND DEVELOPMENT AND POSITION	Check Total sexes	{Men} + {Women} = Total
	Check Total fields of Research and Development	{NS} + {ET} + {MS} + {AS} + {SS} + {H} + {Unknown} = Total per board
	Check coherence of president and member counts with the respective data in EIGE's 'Research funding organisations: presidents and members of the highest decision-making body'[34]	
Tables T7 (institutions) and (universities): HEADS OF INSTITUTIONS IN THE HES BY SEX & HEADS OF UNIVERSITIES OR ASSIMILATED INSTITUTIONS (BASED ON CAPACITY TO DELIVER PhDS) BY SEX	Check Total sexes	{Men} + {Women} = Total
	Check Number of institutes	# institutes T7 ≥ # institutes T8

[34] http://eige.europa.eu/gender-statistics/dgs/indicator/wmidm_educ__wmid_resfund

Note that a final visual scan of the formatted tables/charts appearing in the main publication is performed in the end to detect any inconsistencies that would have been overlooked in previous validation steps.

Additional data considerations

Age groups

Data referring to the labour force refer to all persons aged 15+ living in private households and include the employed and the unemployed. Data referring research personnel and to human resources in science and technology (HRST) refer to the age group 25–64.

Small numbers

For some countries with small populations, raw data relating to small numbers of people have been reported here. The percentages and indicators have not always been included (mostly growth rates) and this is identified in the footnotes to the indicators. The reader is therefore asked to bear this in mind when interpreting the most disaggregated data.

EU estimates

EU totals estimated by DG Research and Innovation (as noted in the footnotes) are based upon existing data for the reference year in combination with the next available year if the reference year is unavailable, in the following sequence (n-1, n+1, n-2, n+2, etc.).

The aggregates were estimated by DG Research and Innovation only when at least 60 % of the EU population on a given indicator was available. *These estimates are intended only as an indication for the reader.*

Rounding error

In some cases, the row or column totals do not match the sum of the data. This may be due to rounding error.

Decimal places

All the data in the figures have been calculated at the precision levels of one or two decimals. However, the values have been rounded in the figures to make them fit.

Cut-off date

The cut-off date for data downloaded from Eurostat's dissemination database (Eurostat) was the 18th of June 2020. Due to the large variety of data sources and variability in data availability, some other cut-off dates were used in order to gather all the required data. Details on the cut-off dates are released in the She Figures 2021.

ANNEXES

Annex 1: Changes to international classification standards

International Standard Classification of Education (ISCED)

The International Standard Classification of Education (ISCED) is the UN framework for classifying educational programmes at different levels.

The data shown until the 2015 edition of the She Figures publication were compiled following the ISCED 1997 version. The classification has been revised and the data shown from the 2018 edition of the She Figures publication onwards follow the ISCED 2011 classification of education levels and the ISCED-F classification of fields of education and training.

For the levels of education of interest for She Figures 2021, the changes are shown in the following table.

Table 5 Correspondence between ISCED 2011 and ISCED 1997 levels of interest for She Figures 2021

ISCED 2011	ISCED 1997
Level 5 – Short-cycle tertiary education	Level 5 – First stage of tertiary education (not leading directly to an advanced research qualification) (5A, 5B)
Level 6 – Bachelor's or equivalent level	
Level 7 – Master's or equivalent level	
Level 8 – Doctoral or equivalent level	Level 6 –Second stage of tertiary education (leading to an advanced research qualification)

Source: Reproduced from http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf

There are also some changes in broad and narrow fields of study between ISCED 1997 and ISCED-F:

- The 1997 broad field 3 'Social sciences, business and law' has been split into ISCED-F broad fields 03 'Social sciences, journalism and information' and 04 'Business, administration and law'.
- The 1997 broad field 4 'Science' has been split into ISCED-F broad fields 05 'Natural sciences, mathematics and statistics' and 06 'Information and Communication Technologies (ICTs)'.
- The 1997 broad field 6 'Agriculture' lost some of its field 62 'Agriculture, forestry and fishery' to the ISCED-F broad field 05 'Natural sciences, mathematics and statistics'.
- The 1997 broad field 8 'Services' lost some of its field 85 'Environmental protection' to the ISCED-F broad field 07 'Engineering, manufacturing and construction'.

Annex 2: Definitions of key terms

Gender refers to 'social attributes and opportunities associated with being female and male and to the relationships between women and men and girls and boys, as well as to the relations between women and those between men' (https://eige.europa.eu/thesaurus/terms/1141).

Sex refers to the 'biological attributes that distinguish male, female and intersex' (European Commission, 2020d).

Gender identity refers to 'each person's deeply felt internal and individual experience of gender, which may or may not correspond to the sex assigned at birth, including the personal sense of the body (which may involve, if freely chosen, modification of bodily appearance or function by medical, surgical or other means) and other expressions of gender, including dress, speech and mannerisms' (https://eige.europa.eu/thesaurus/terms/1179).

While the data collection for She Figures only considers sex-disaggregated data for men and women, it will be important to also consider non-binary gender for data collection in future publications, where possible. **Non-binary** is an umbrella term for gender identities that fall outside the gender binary of man or woman. This includes individuals whose gender identity is neither exclusively man nor woman, a combination of man and woman or between or beyond genders. The United Nations Economics Commission for Europe (UNECE) provides an in-depth review on measuring gender identity conducted by the Bureau of the Conference of European Statisticians (CES) in February 2019. The review published in April 2019 (https://unece.org/statistics/gender-statistics) provides insights into different approaches for statistical measurement of gender identity that have been undertaken so far and examines the issues and challenges which will be important to consider for future data collection for She Figures.

Gender Equality Plan refers to 'a set of actions aiming at conducting impact assessment/ audits of procedures and practices to identify gender bias; identifying and implementing innovative strategies to correct any bias; and setting targets and monitoring progress via indicators (European Commission, 2012).

Equality between women and men (gender equality) refers to the 'equal rights, responsibilities and opportunities of women and men and girls and boys. Equality does not mean that women and men will become the same but that women's and men's rights, responsibilities and opportunities will not depend on whether they are born men or women. Gender equality implies that the interests, needs and priorities of both women and men are taken into consideration, recognizing the diversity of different groups of women and men ...'.

Whilst sex and gender are often used interchangeably, they are not the same. In general, the She Figures project understands sex to be a biological category, whilst gender relates to historical, cultural and social realities. For example, when data are broken down to show the individual data for women and men, these are understood to be *sex-disaggregated* data, and not *gender-disaggregated* data.

Within the fields of education, research and innovation, there are a range of additional terms useful for measuring gender equality (which is understood to be a multi-dimensional concept). Many of these relate to the notion of 'segregation'. Definitions of each of these are given below, based on those discussed by the European Commission's Expert Group on Gender and Employment (EGGE, 2009, p. 30, pp. 40–41):

Gender segregation in the labour market refers to the gendered division of labour in employment. It is a broad term, describing the tendency for women and men to work in different occupations, sectors, fields, etc. It is often associated with potentially negative effects, including narrowed choice for women and men, perpetuation of gender

stereotypes, vertical segregation (see below) and finally, the under-valuing of skills and abilities linked to women's work (affecting their pay). Since the 1960s, a range of additional terms have emerged to understand gender segregation more fully, including horizontal, vertical, sectoral and occupational segregation.

Horizontal segregation relates to the concentration of women and men around different sectors (sectoral segregation) and occupations (occupational segregation) (https://eige.europa.eu/thesaurus/terms/1247). It can occur within both education (e.g., over-/under-representation of one sex in particular subjects) and employment (e.g., over-/under-representation of one sex in particular professions, industries, etc.). Unlike vertical segregation, these occupations and sectors are not ordered by a particular criterion. However, the issue of horizontal segregation may in turn lead to greater vertical segregation. For example, the under-valuing of competencies associated with 'women's work' may limit women's prospects for career advancement.

Vertical segregation refers to the concentration of either men or women in 'top' posts, such as decision-making positions or other positions of responsibility. Such roles are often associated with 'desirable' features, including greater pay, prestige and social security. In the context of research and innovation, the over-representation of men amongst heads of universities is an example of such segregation. Below is the list of positions used in producing the She Figures publication:

A: *The single highest grade / post at which research is normally conducted within the institutional or corporate system.*

B: *All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e., below A and above C.*

C: *The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system.*

D: *Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.*

A table with the qualifications for each grade by country is presented in the Appendix 2 of the main publication of She Figures.

She Figures 2015 introduced new definitions of 'boards' as part of the Women in Science questionnaire, based on consultation with the European Commission and the Statistical Correspondents. These distinguish more clearly between the functions of different boards, by focusing on 'scientific boards' and 'administrative/advisory boards':

Scientific board of research organisation: A publicly or privately managed and financed group of elected or appointed experts that exists to implement scientific policy by, among other things, directing the research agenda, resource allocation and management within scientific research.

Administrative / advisory board of research organisation: A publicly or privately managed and financed group of elected or appointed experts that exists to support the research agenda in a non-executive function by, among other things, administering research activities, consulting and coordinating different actors and taking a general advisory role.

Where boards fall into both categories, this was indicated by Statistical Correspondents. She Figures includes only research boards of umbrella, national-level research performing

organisations (RPOs) and research funding organisations (RFOs), as opposed to all research organisations operating in a particular country. Whilst data were collected separately for the two types of boards, this indicator remained combined in one indicator in the publication.

Gender dimension in research and innovation content

Gender dimension in research and innovation content is a concept regrouping the various aspects concerning biological characteristics and social/cultural factors of both women and men into the development of research policies, programmes and projects' (European Commission, 2014d). The European Commission created the Gendered Innovations/Innovation through Gender Expert Group, defining that research that explores 'gender dimension' integrates sex and gender analysis into research. 'Sex' refers to basic biological characteristics of females and males and 'gender' refers to cultural attitudes and behaviours that shape 'feminine' and 'masculine' behaviours, products, technologies, environments, and knowledge (European Commission, 2020d and European Commission, 2013b).

Annex 3: Index list of indicators

Annex 4: Correspondence of ASJC sub-categories with fields of Research and Development

The Fields of Research and Development (FORD) classifications were mapped with the ASJC sub-categories from Scopus ™ according to the following mapping.
The classifications were mapped at category level (ASJC category to target category).

Table 6 Dimensions of the data quality framework

ASJC code	ASJC name	FORD code	FORD Name
1100	General Agricultural and Biological Sciences	4.1	Agriculture, forestry, and fishery
1101	Agricultural and Biological Sciences (miscellaneous)	4.5	Other agricultural sciences
1102	Agronomy and Crop Science	4.1	Agriculture, forestry, and fishery
1103	Animal Science and Zoology	1.6	Biological sciences
1103	Animal Science and Zoology	4.2	Biological sciences | Animal and dairy science
1104	Aquatic Science	1.6	Biological Sciences
1104	Aquatic Science	4.1	Agriculture, forestry, and fishery
1105	Ecology, Evolution, Behaviour and Systematics	1.6	Biological sciences
1105	Ecology, Evolution, Behaviour and Systematics	1.5	Earth and related environmental sciences
1106	Food Science	4.2	Animal and dairy science
1107	Forestry	4.1	Agriculture, forestry, and fishery
1108	Horticulture	4.1	Agriculture, forestry, and fishery
1109	Insect Science	1.6	Biological sciences
1110	Plant Science	1.6	Biological sciences
1111	Soil Science	4.1	Agriculture, forestry, and fishery
1200	General Arts and Humanities	6.5	Other humanities
1201	Arts and Humanities (miscellaneous)	6.5	Other humanities
1202	History	6.1	History and archaeology
1203	Language and Linguistics	6.2	Languages and literature
1204	Archaeology (arts and humanities)	6.1	History and archaeology
1205	Classics	6.2	Languages and literature

ASJC code	ASJC name	FORD code	FORD Name
1206	Conservation	6.5	Other humanities
1207	History and Philosophy of Science	6.5	Other humanities
1208	Literature and Literary Theory	6.2	Languages and literature
1209	Museology	6.5	Other humanities
1210	Music	6.4	Arts (arts, history of arts, performing arts, music)
1211	Philosophy	6.3	Philosophy, ethics and religion
1212	Religious Studies	6.3	Philosophy, ethics and religion
1213	Visual Arts and Performing Arts	6.4	Arts (arts, history of arts, performing arts, music)
1300	General Biochemistry, Genetics and Molecular Biology	1.6	Biological sciences
1301	Biochemistry, Genetics and Molecular Biology (miscellaneous)	1.6	Biological sciences
1302	Aging	1.6	Biological sciences
1302	Aging	3.1	Basic medicine
1303	Biochemistry	1.6	Biological sciences
1304	Biophysics	1.6	Biological sciences
1305	Biotechnology	1.6	Biological sciences
1306	Cancer Research	1.6	Biological science
1306	Cancer Research	3.1	Basic medicine
1307	Cell Biology	1.6	Biological sciences
1308	Clinical Biochemistry	3.2	Clinical medicine
1309	Developmental Biology	1.6	Biological sciences
1310	Endocrinology	1.6	Biological sciences
1311	Genetics	1.6	Biological sciences
1312	Molecular Biology	1.6	Biological sciences
1313	Molecular Medicine	1.6	Biological science
1313	Molecular Medicine	3.1	Basic medicine
1314	Physiology	1.6	Biological science

ASJC code	ASJC name	FORD code	FORD Name
1314	Physiology	3.1	Basic medicine
1315	Structural Biology	1.6	Biological sciences
1400	General Business, Management and Accounting	5.2	Economics and business
1401	Business, Management and Accounting (miscellaneous)	5.2	Economics and business
1402	Accounting	5.2	Economics and business
1403	Business and International Management	5.2	Economics and business
1404	Management Information Systems	5.2	Economics and business
1405	Management of Technology and Innovation	5.2	Economics and business
1406	Marketing	5.2	Economics and business
1407	Organizational Behaviour and Human Resource Management	5.2	Economics and business
1408	Strategy and Management	5.2	Economics and business
1409	Tourism, Leisure and Hospitality Management	5.2	Economics and business
1410	Industrial Relations	5.2	Economics and business
1500	General Chemical Engineering	2.4	Chemical engineering
1501	Chemical Engineering (miscellaneous)	2.4	Chemical engineering
1502	Bioengineering	2.11	Other engineering and technologies
1503	Catalysis	1.4	Chemical Sciences
1503	Catalysis	2.4	Chemical engineering
1504	Chemical Health and Safety	2.4	Chemical engineering
1505	Colloid and Surface Chemistry	1.4	Chemical Sciences
1505	Colloid and Surface Chemistry	2.4	Chemical sciences | Chemical engineering
1506	Filtration and Separation	1.4	Chemical Sciences
1506	Filtration and Separation	2.4	Chemical sciences | Chemical engineering
1507	Fluid Flow and Transfer Processes	1.3	Physical sciences
1507	Fluid Flow and Transfer Processes	1.4	Chemical Sciences

ASJC code	ASJC name	FORD code	FORD Name
1507	Fluid Flow and Transfer Processes	2.4	Chemical engineering
1508	Process Chemistry and Technology	2.4	Chemical engineering
1600	General Chemistry	1.4	Chemical sciences
1601	Chemistry (miscellaneous)	1.4	Chemical sciences
1602	Analytical Chemistry	1.4	Chemical sciences
1603	Electrochemistry	1.4	Chemical sciences
1604	Inorganic Chemistry	1.4	Chemical sciences
1605	Organic Chemistry	1.4	Chemical sciences
1606	Physical and Theoretical Chemistry	1.4	Chemical sciences
1607	Spectroscopy	1.4	Chemical sciences
1700	General Computer Science	1.2	Computer and information sciences
1701	Computer Science (miscellaneous)	1.2	Computer and information sciences
1702	Artificial Intelligence	1.2	Computer and information sciences
1703	Computational Theory and Mathematics	1.2	Computer and information sciences
1704	Computer Graphics and Computer -Aided Design	1.2	Computer and information sciences
1705	Computer Networks and Communications	1.2	Computer and information sciences
1706	Computer Science Applications	1.2	Computer and information sciences
1707	Computer Vision and Pattern Recognition	1.2	Computer and information sciences
1708	Hardware and Architecture	1.2	Computer and information sciences
1709	Human-Computer Interaction	1.2	Computer and information sciences
1710	Information Systems	1.2	Computer and information sciences
1711	Signal Processing	1.2	Computer and information sciences
1712	Software	1.2	Computer and information sciences
1800	General Decision Sciences	5.9	Other social sciences

ASJC code	ASJC name	FORD code	FORD Name
1801	Decision Sciences (miscellaneous)	5.9	Other social sciences
1802	Information Systems and Management	1.2	Computer and information sciences
1803	Management Science and Operations Research	5.9	Other social sciences
1804	Statistics, Probability and Uncertainty	1.1	Mathematics
1900	General Earth and Planetary Sciences	1.5	Earth and related environmental sciences
1901	Earth and Planetary Sciences (miscellaneous)	1.5	Earth and related environmental sciences
1902	Atmospheric Science	1.5	Earth and related environmental sciences
1903	Computers in Earth Sciences	1.5	Earth and related environmental sciences
1904	Earth-Surface Processes	1.5	Earth and related environmental sciences
1905	Economic Geology	1.5	Earth and related environmental sciences
1906	Geochemistry and Petrology	1.5	Earth and related environmental sciences
1907	Geology	1.5	Earth and related environmental sciences
1908	Geophysics	1.5	Earth and related environmental sciences
1909	Geotechnical Engineering and Engineering Geology	1.5	Earth and related environmental sciences
1910	Oceanography	1.5	Earth and related environmental sciences
1911	Palaeontology	1.5	Earth and related environmental sciences
1912	Space and Planetary Science	1.5	Earth and related environmental sciences
1913	Stratigraphy	1.5	Earth and related environmental sciences
2000	General Economics, Econometrics and Finance	5.2	Economics and business
2001	Economics, Econometrics and Finance (miscellaneous)	5.2	Economics and business
2002	Economics and Econometrics	5.2	Economics and business
2003	Finance	5.2	Economics and business
2100	General Energy	1.7	Other natural sciences

ASJC code	ASJC name	FORD code	FORD Name
2101	Energy (miscellaneous)	1.7	Other natural sciences
2102	Energy Engineering and Power Technology	1.7	Other natural sciences
2103	Fuel Technology	1.7	Other natural sciences
2103	Fuel Technology	2.7	Environmental engineering
2104	Nuclear Energy and Engineering	2.11	Other engineering and technologies
2105	Renewable Energy, Sustainability and the Environment	1.5	Earth and related environmental sciences
2105	Renewable Energy, Sustainability and the Environment	1.7	Other natural sciences
2105	Renewable Energy, Sustainability and the Environment	2.7	Environmental engineering
2200	General Engineering	2.11	Other engineering and technologies
2201	Engineering (miscellaneous)	2.11	Other engineering and technologies
2202	Aerospace Engineering	2.3	Mechanical Engineering
2203	Automotive Engineering	2.3	Mechanical Engineering
2204	Biomedical Engineering	2.6	Medical engineering
2204	Biomedical Engineering	3.4	Medical biotechnology
2205	Civil and Structural Engineering	2.1	Civil engineering
2206	Computational Mechanics	1.3	Physical sciences
2207	Control and Systems Engineering	2.11	Other engineering and technologies
2208	Electrical and Electronic Engineering	2.2	Electrical engineering, electronic engineering, information engineering
2209	Industrial and Manufacturing Engineering	2.3	Mechanical Engineering
2209	Industrial and Manufacturing Engineering	2.4	Chemical engineering
2209	Industrial and Manufacturing Engineering	2.11	Other engineering and technologies
2210	Mechanical Engineering	2.3	Mechanical Engineering
2211	Mechanics of Materials	2.5	Materials engineering
2212	Ocean Engineering	2.7	Environmental engineering

ASJC code	ASJC name	FORD code	FORD Name
2213	Safety, Risk, Reliability and Quality	2.11	Other engineering and technologies
2214	Media Technology	2.11	Other engineering and technologies
2215	Building and Construction	2.1	Civil engineering
2216	Architecture	2.1	Civil engineering
2300	General Environmental Science	1.5	Earth and related environmental sciences
2301	Environmental Science (miscellaneous)	1.5	Earth and related environmental sciences
2302	Ecological Modelling	1.5	Earth and related environmental sciences
2303	Ecology	1.5	Earth and related environmental sciences
2304	Environmental Chemistry	1.5	Earth and related environmental sciences
2304	Environmental Chemistry	1.4	Chemical sciences
2305	Environmental Engineering	1.5	Earth and related environmental sciences
2306	Global and Planetary Change	1.5	Earth and related environmental sciences
2307	Health, Toxicology and Mutagenesis	1.6	Biological sciences
2307	Health, Toxicology and Mutagenesis	3.1	Biological sciences | Basic medicine
2308	Management, Monitoring, Policy and Law	5.5	Law
2308	Management, Monitoring, Policy and Law	5.6	Political Sciences
2309	Nature and Landscape Conservation	1.5	Earth and related environmental sciences
2310	Pollution	1.5	Earth and related environmental sciences
2311	Waste Management and Disposal	1.5	Earth and related environmental sciences
2312	Water Science and Technology	1.5	Earth and related environmental sciences
2400	General Immunology and Microbiology	1.6	Biological sciences
2400	General Immunology and Microbiology	3.1	Basic medicine
2401	Immunology and Microbiology (miscellaneous)	1.6	Biological sciences

ASJC code	ASJC name	FORD code	FORD Name
2401	Immunology and Microbiology (miscellaneous)	3.1	Basic medicine
2402	Applied Microbiology and Biotechnology	1.6	Biological sciences
2402	Applied Microbiology and Biotechnology	2.9	Industrial biotechnology
2402	Applied Microbiology and Biotechnology	3.1	Basic Medicine
2403	Immunology	1.6	Biological sciences
2403	Immunology	3.1	Basic medicine
2404	Microbiology	1.6	Biological sciences
2404	Microbiology	3.1	Basic medicine
2405	Parasitology	1.6	Biological sciences
2405	Parasitology	3.1	Basic medicine
2406	Virology	1.6	Biological sciences
2406	Virology	3.1	Basic medicine
2500	General Materials Science	1.7	Other natural sciences
2500	General Materials Science	2.5	Materials engineering
2501	Materials Science (miscellaneous)	1.7	Other natural sciences
2501	Materials Science (miscellaneous)	2.5	Materials engineering
2502	Biomaterials	1.7	Other natural sciences
2502	Biomaterials	2.5	Materials engineering
2502	Biomaterials	2.6	Medical engineering
2503	Ceramics and Composites	1.4	Chemical sciences
2503	Ceramics and Composites	2.5	Materials engineering
2504	Electronic, Optical and Magnetic Materials	1.4	Chemical sciences
2504	Electronic, Optical and Magnetic Materials	2.5	Materials engineering
2505	Materials Chemistry	1.4	Chemical sciences
2505	Materials Chemistry	2.5	Materials engineering
2506	Metals and Alloys	1.4	Chemical sciences

ASJC code	ASJC name	FORD code	FORD Name
2506	Metals and Alloys	2.5	Materials engineering
2507	Polymers and Plastics	1.4	Chemical sciences
2507	Polymers and Plastics	2.5	Materials engineering
2508	Surfaces, Coatings and Films	1.4	Chemical sciences
2508	Surfaces, Coatings and Films	2.5	Materials engineering
2600	General Mathematics	1.1	Mathematics
2601	Mathematics (miscellaneous)	1.1	Mathematics
2602	Algebra and Number Theory	1.1	Mathematics
2603	Analysis	1.1	Mathematics
2604	Applied Mathematics	1.1	Mathematics
2605	Computational Mathematics	1.1	Mathematics
2606	Control and Optimization	1.1	Mathematics
2607	Discrete Mathematics and Combinatorics	1.1	Mathematics
2608	Geometry and Topology	1.1	Mathematics
2609	Logic	1.1	Mathematics
2610	Mathematical Physics	1.1	Mathematics
2611	Modelling and Simulation	1.1	Mathematics
2612	Numerical Analysis	1.1	Mathematics
2613	Statistics and Probability	1.1	Mathematics
2614	Theoretical Computer Science	1.1	Mathematics
2614	Theoretical Computer Science	1.2	Computer and information sciences
2700	General Medicine	3.1	Basic medicine
2701	Medicine (miscellaneous)	3.1	Basic medicine
2702	Anatomy	3.1	Basic medicine
2703	Anesthesiology and Pain Medicine	3.1	Basic medicine
2704	Biochemistry (medical)	3.2	Clinical medicine
2705	Cardiology and Cardiovascular Medicine	3.2	Clinical medicine

ASJC code	ASJC name	FORD code	FORD Name
2706	Critical Care and Intensive Care Medicine	3.2	Clinical medicine
2707	Complementary and Alternative Medicine	3.2	Clinical medicine
2708	Dermatology	3.2	Clinical medicine
2709	Drug Guides	3.2	Clinical medicine
2710	Embryology	3.2	Clinical medicine
2711	Emergency Medicine	3.2	Clinical medicine
2712	Endocrinology, Diabetes and Metabolism	3.2	Clinical medicine
2713	Epidemiology	3.2	Clinical medicine
2714	Family Practice	3.3	Health sciences
2715	Gastroenterology	3.2	Clinical medicine
2716	Genetics (clinical)	3.1	Basic medicine
2717	Geriatrics and Gerontology	3.2	Clinical medicine
2718	Health Informatics	3.3	Health sciences
2719	Health Policy	3.3	Health sciences
2720	Haematology	3.2	Clinical medicine
2721	Hepatology	3.2	Clinical medicine
2722	Histology	3.2	Clinical medicine
2723	Immunology and Allergy	3.2	Clinical medicine
2724	Internal Medicine	3.2	Clinical medicine
2725	Infectious Diseases	3.2	Clinical medicine
2726	Microbiology (medical)	3.2	Clinical medicine
2727	Nephrology	3.2	Clinical medicine
2728	Neurology (clinical)	3.2	Clinical medicine
2729	Obstetrics and Gynaecology	3.2	Clinical medicine
2730	Oncology	3.2	Clinical medicine
2731	Ophthalmology	3.2	Clinical medicine
2732	Orthopaedics and Sports Medicine	3.2	Clinical medicine

ASJC code	ASJC name	FORD code	FORD Name
2733	Otorhinolaryngology	3.2	Clinical medicine
2734	Pathology and Forensic Medicine	3.1	Basic medicine
2735	Paediatrics, Perinatology and Child Health	3.2	Clinical medicine
2736	Pharmacology (medical)	3.2	Clinical medicine
2737	Physiology (medical)	3.2	Clinical medicine
2738	Psychiatry and Mental Health	3.2	Clinical medicine
2739	Public Health, Environmental and Occupational Health	3.3	Health sciences
2740	Pulmonary and Respiratory Medicine	3.2	Clinical medicine
2741	Radiology, Nuclear Medicine and Imaging	3.2	Clinical medicine
2742	Rehabilitation	3.3	Health sciences
2743	Reproductive Medicine	3.2	Clinical medicine
2743	Reproductive Medicine	3.3	Health sciences
2744	Reviews and References (medical)	3.2	Clinical medicine
2744	Reviews and References (medical)	3.3	Health sciences
2745	Rheumatology	3.2	Clinical medicine
2746	Surgery	3.2	Clinical medicine
2747	Transplantation	3.2	Clinical medicine
2748	Urology	3.2	Clinical medicine
2800	General Neuroscience	3.1	Basic medicine
2801	Neuroscience (miscellaneous)	3.1	Basic medicine
2802	Behavioural Neuroscience	3.1	Basic medicine
2802	Behavioural Neuroscience	5.1	Psychology and cognitive sciences
2803	Biological Psychiatry	3.2	Clinical medicine
2804	Cellular and Molecular Neuroscience	3.1	Basic medicine
2805	Cognitive Neuroscience	3.1	Basic medicine
2805	Cognitive Neuroscience	5.1	Psychology and cognitive sciences

ASJC code	ASJC name	FORD code	FORD Name
2806	Developmental Neuroscience	3.1	Basic medicine
2806	Developmental Neuroscience	5.1	Psychology and cognitive sciences
2807	Endocrine and Autonomic Systems	3.2	Clinical medicine
2808	Neurology	3.2	Clinical medicine
2809	Sensory Systems	3.1	Basic medicine
2809	Sensory Systems	3.4	Medical biotechnology
2900	General Nursing	3.3	Health sciences
2901	Nursing (miscellaneous)	3.3	Health sciences
2902	Advanced and Specialized Nursing	3.3	Health sciences
2903	Assessment and Diagnosis	3.3	Health sciences
2904	Care Planning	3.3	Health sciences
2905	Community and Home Care	3.3	Health sciences
2906	Critical Care Nursing	3.3	Health sciences
2907	Emergency Nursing	3.3	Health sciences
2908	Fundamentals and Skills	3.3	Health sciences
2909	Gerontology	3.2	Clinical medicine
2909	Gerontology	3.3	Health sciences
2910	Issues, Ethics and Legal Aspects	3.3	Health sciences
2911	Leadership and Management	3.3	Health sciences
2912	LPN and LVN	3.3	Health sciences
2913	Maternity and Midwifery	3.3	Health sciences
2913	Maternity and Midwifery	5.1	Psychology and cognitive sciences
2914	Medical and Surgical Nursing	3.2	Clinical medicine
2914	Medical and Surgical Nursing	3.3	Health sciences
2915	Nurse Assisting	3.3	Health sciences
2916	Nutrition and Dietetics	3.3	Health sciences
2917	Oncology (nursing)	3.3	Health sciences

ASJC code	ASJC name	FORD code	FORD Name
2918	Pathophysiology	3.2	Clinical medicine
2919	Paediatrics	3.2	Clinical medicine
2919	Paediatrics	3.3	Health sciences
2920	Pharmacology (nursing)	3.3	Health sciences
2921	Psychiatric Mental Health	3.3	Health sciences
2922	Research and Theory	3.3	Health sciences
2923	Review and Exam Preparation	3.3	Health sciences
3000	General Pharmacology, Toxicology and Pharmaceutics	3.1	Basic medicine
3001	Pharmacology, Toxicology and Pharmaceutics (miscellaneous)	3.1	Basic medicine
3002	Drug Discovery	3.1	Basic medicine
3003	Pharmaceutical Science	3.1	Basic medicine
3004	Pharmacology	3.1	Basic medicine
3005	Toxicology	3.1	Basic medicine
3100	General Physics and Astronomy	1.3	Physical sciences
3101	Physics and Astronomy (miscellaneous)	1.3	Physical sciences
3102	Acoustics and Ultrasonics	1.3	Physical sciences
3103	Astronomy and Astrophysics	1.3	Physical sciences
3104	Condensed Matter Physics	1.3	Physical sciences
3105	Instrumentation	1.3	Physical sciences
3106	Nuclear and High Energy Physics	1.3	Physical sciences
3107	Atomic and Molecular Physics, and Optics	1.3	Physical sciences
3108	Radiation	1.3	Physical sciences
3109	Statistical and Nonlinear Physics	1.3	Physical sciences
3110	Surfaces and Interfaces	1.3	Physical sciences
3200	General Psychology	5.1	Psychology and cognitive sciences
3201	Psychology (miscellaneous)	5.1	Psychology and cognitive sciences

ASJC code	ASJC name	FORD code	FORD Name
3202	Applied Psychology	5.1	Psychology and cognitive sciences
3203	Clinical Psychology	5.1	Psychology and cognitive sciences
3204	Developmental and Educational Psychology	5.1	Psychology and cognitive sciences
3205	Experimental and Cognitive Psychology	5.1	Psychology and cognitive sciences
3206	Neuropsychology and Physiological Psychology	5.1	Psychology and cognitive sciences
3207	Social Psychology	5.1	Psychology and cognitive sciences
3300	General Social Sciences	5.9	Other social sciences
3301	Social Sciences (miscellaneous)	5.9	Other social sciences
3302	Archaeology	6.1	History and archaeology
3303	Development	5.1	Psychology and cognitive sciences
3304	Education	5.3	Education
3305	Geography, Planning and Development	5.7	Social and economic geography
3306	Health (social science)	3.3	Health sciences
3306	Health (social science)	5.9	Other social sciences
3307	Human Factors and Ergonomics	5.1	Psychology and cognitive sciences
3308	Law	5.5	Law
3309	Library and Information Sciences	5.9	Other social sciences
3310	Linguistics and Language	6.2	Languages and literature
3311	Safety Research	2.11	Other engineering and technologies
3312	Sociology and Political Science	5.4	Sociology
3312	Sociology and Political Science	5.6	Political Sciences
3313	Transportation	2.11	Other engineering and technologies
3314	Anthropology	5.4	Sociology
3315	Communication	5.8	Media and communications
3316	Cultural Studies	5.4	Sociology

ASJC code	ASJC name	FORD code	FORD Name
3317	Demography	5.4	Sociology
3318	Gender Studies	5.4	Sociology
3319	Life-span and Life-course Studies	5.4	Sociology
3320	Political Science and International Relations	5.6	Political Sciences
3321	Public Administration	5.6	Political Sciences
3322	Urban Studies	5.4	Sociology
3322	Urban Studies	5.7	Social and economic geography
3400	General Veterinary	4.3	Veterinary science
3401	Veterinary (miscellaneous)	4.3	Veterinary science
3402	Equine	4.3	Veterinary science
3403	Food Animals	4.3	Veterinary science
3404	Small Animals	4.3	Veterinary science
3500	General Dentistry	3.2	Clinical medicine
3501	Dentistry (miscellaneous)	3.2	Clinical medicine
3502	Dental Assisting	3.2	Clinical medicine
3503	Dental Hygiene	3.2	Clinical medicine
3504	Oral Surgery	3.2	Clinical medicine
3505	Orthodontics	3.2	Clinical medicine
3506	Periodontics	3.2	Clinical medicine
3600	General Health Professions	3.3	Health sciences
3601	Health Professions (miscellaneous)	3.3	Health sciences
3602	Chiropractics	3.2	Clinical medicine
3602	Chiropractics	3.3	Health sciences
3603	Complementary and Manual Therapy	3.3	Health sciences
3604	Emergency Medical Services	3.2	Clinical medicine
3604	Emergency Medical Services	3.3	Health sciences
3605	Health Information Management	3.3	Health sciences

ASJC code	ASJC name	FORD code	FORD Name
3606	Medical Assisting and Transcription	3.3	Health sciences
3607	Medical Laboratory Technology	2.6	Medical engineering
3608	Medical Terminology	3.3	Health sciences
3609	Occupational Therapy	3.3	Health sciences
3610	Optometry	3.2	Clinical medicine
3611	Pharmacy	3.3	Health sciences
3612	Physical Therapy, Sports Therapy and Rehabilitation	3.2	Clinical medicine
3612	Physical Therapy, Sports Therapy and Rehabilitation	3.3	Health sciences
3613	Podiatry	3.2	Clinical medicine
3614	Radiological and Ultrasound Technology	2.6	Medical engineering
3615	Respiratory Care	3.2	Clinical medicine
3616	Speech and Hearing	3.2	Clinical Medicine
3616	Speech and Hearing	5.1	Psychology and cognitive sciences

Annex 5: Identifying good practices

In addition to analysing trends in quantitative data and providing EU policy contextualisation, the She Figures 2021 publication identifies national policy measures relevant to gender equality implementation in the European Research Area. These good practices are presented throughout the main study and in the policy briefs. Good practices were identified through literature and policy review and direct suggestions from participants in online policy workshops held in 2020 and 2021.[35]

Process for identifying good practices

Outlined below is the process followed to identify good practices. Ultimately not all information described below is available for all policies and practices, therefore suggested good practices identified by the European Commission and the Standing Working Group on Gender in Research and Innovation under European Research Area and Innovation Committee were prioritised for inclusion.

A two-stage process to undertake policy analysis was established. 'Stage 1' gathered a wide range of policy details to identify some initial good practice policy selections. Information that gathered in the first stage of policy mapping was:

- Policy name (native language)
- Policy name (EN)
- Policy Description
- Specific policy objectives
- Nature of the policy: - legislative: no time limit- legislative: time limited- non-legal measures
- Target groups
- Is it national?
- Is it a permanent measure?
- Year established & Year of significant amendments (if applicable)
- Detail of significant amendments (if applicable)
- She Figures theme
- Relevant She Figures indicators
- Evidence of impact of the policy in sources consulted.

Stage 2 of the data collection sought to gather more detailed information for a small number of polices / practices to explain the intervention logic and identify evidence of positive impact identified during Stage 1 of the review. The information sought was:

- Details of the intervention logic – (planned or actual)
    - Baseline and policy objectives
    - MS inputs (financial)
    - European funding inputs (financial)
    - Activities
    - Outputs
    - Impacts
    - Explanation of any innovative elements identified

To develop the collected information and determine which policies and practices could be considered to be good practices, a range of assessment criteria for potential good practices were developed during the process of consultation on the content of She Figures 2021. Suggested criteria for good practices were collected during policy workshop 1 and further

[35] Two policy workshops were held in 2020 and 2021 to develop the topics and content of policy briefs for She Figures 2021. Participants included members of the European Research Area and Innovation Committee Standing Working Group on Gender in Research and Innovation and other pan-European key stakeholder organisations with expertise on the proposed policy brief topics.

discussed and validated during policy workshop 2. The box below contains the criteria agreed as suitable for assessing good practice.

Criteria to identify and assess good practices

A good practice should:

- Be specific and based on an empirical baseline assessment
- Aim to contribute to a particular gender equality objective in research and innovation, ideally one or more of the three main ERA priorities
- Have clear targets and target groups
- Reflect a theory of change/programme theory
- Guarantee that key stakeholders have a role in its development
- Aim at systematic and structured change (as opposed to being one-off or time-limited)
- Be accompanied by evidence of positive impact

Ideally, it should also:

- Be supported in national law or by a wider gender equality strategy
- Show signs of being replicable to other contexts (either other institutions or countries)
- (For institutional measures) Reflect evidence of institutional leadership support (for actions within RPOs)

REFERENCES

Aad, G., et al., *Combined Measurement of the Higgs Boson Mass in pp Collisions at* $\sqrt{s} = 7\ and\ 8$ *TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. 114, article 191803, 2015. Available at: https://arxiv.org/pdf/1503.07589.pdf

All Science Journal Classification Codes (ASJC). Available at: https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/

Bagihole, B., *NORFACE gender equality workshop address*, 2005, Reykjavik, Iceland.

Bar-Ilan, J., *Informetrics at the beginning of the 21st century—A review*, Journal of Informetrics, 2 (1), 2008, Pages 1-52.

Berry, A.K. and Zucker, I., *Sex bias in neuroscience and biomedical research*, Neuroscience & Biobehavioural Reviews, Volume 35, Issue 3, January 2011, Pages 565-572.

Burri, S. and Aune, H., *European Network of Legal Experts in the Field of Gender Equality*, *Sex Discrimination in Relation to Part-Time and Fixed-Term Work*, 2013. Available at: https://publications.europa.eu/en/publication-detail/-/publication/83fb80e9-9444-42bb-b35d-f9fa979b2900/language-en

Cameron, I., Synnott, J., Beisiegel, U., O'Carroll, C., Esposito, F., Harrap, K. A., Israel, N., Modjeska, N., Predescu, R., Prijic-Samarzija, S. and Vandevelde, K., *Shaping the future of the Human Resources Strategy for Researchers – HRS4R*, Brussels, 2015. Available at: https://cdn1.euraxess.org/sites/default/files/policy_library/experts-report-strengthened-hrs4r-9-2015_0.pdf

Council of the European Union, *Council conclusions on women and the economy: Economic independence from the perspective of part-time work and self-employment* [pdf], 2014. Available at: http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/lsa/143269.pdf

Council of the European Union, *Council conclusions on enhanced measures to reduce horizontal gender segregation in education and employment* [pdf], 2017. *Available at:* http://data.consilium.europa.eu/doc/document/ST-15468-2017-INIT/en/pdf

DG for Internal Policies, *European Parliament, Precarious Employment in Europe Part 1: Patterns, Trends and Policy Strategies*, Brussels, 2016. Available at: http://www.europarl.europa.eu/RegData/etudes/STUD/2016/587285/IPOL_STU(2016)587285_EN.pdf

DG Research, *Mapping the maze: Getting more women to the top in research* [pdf], Office for Official Publications of the European Communities, Luxembourg, 2008. Available at: http://ec.europa.eu/research/science-society/document_library/pdf_06/mapping-the-maze-getting-more-women-to-the-top-in-research_en.pdf

DG Research, *Women in science and technology: Creating sustainable careers* [pdf], Office for Official Publications of the European Communities, Luxembourg, 2009. Available at: http://ec.europa.eu/research/science-society/document_library/pdf_06/wist2_sustainable-careers-report_en.pdf

DG Research and Innovation, *Towards a European Framework for research careers*, 2011. Available at: https://cdn5.euraxess.org/sites/default/files/policy_library/towards_a_european_framework_for_research_careers_final.pdf

DG Research and Innovation, *Researchers' Report 2012*, 2012. Available at: http://ec.europa.eu/euraxess/pdf/research_policies/121003_The_Researchers_Report_2012_FINAL_REPORT.pdf

DG Research and Innovation, *HORIZON 2020 in Brief* [pdf], Publications Office of the European Union, Luxembourg, 2014. Available at: http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_inBrief_EN_FinalBAT.pdf

DG Research and Innovation, *She Figures 2015* [pdf], Publications Office of the European Union, Luxembourg, 2016. Available at: https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-final.pdf

DG Research and Innovation, *Horizon Europe; The next EU Research and Innovation investment programme (2021-2027)*, 2019a. Available at: https://ec.europa.eu/info/sites/info/files/research_and_innovation/strategy_on_research_and_innovation/presentations/horizon_europe_en_investing_to_shape_our_future.pdf

DG Research and Innovation, *She Figures 2018* [pdf], Publications Office of the European Union, Luxembourg, 2019b. Available at: https://op.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1/language-en

DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)* , Luxembourg, 2020 https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

Elsevier, *Gender in the Global Research Landscape*, 2017. Available at: https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

European Institute for Gender Equality (EIGE), *Gender Statistics Database* https://eige.europa.eu/gender-statistics/dgs

European Institute for Gender Equality (EIGE), *Gender Equality Index*. Available at: http://eige.europa.eu/content/gender-equality-index#/domain/power

European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research* [pdf], Publications Office of the European Union, Luxembourg, 2016. Available at: https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

Eurofound, *Women, men and working conditions: Executive summary* [pdf], Dublin, 2013. Available at: http://www.eurofound.europa.eu/sites/default/files/ef_files/pubdocs/2013/491/en/1/EF13491EN.pdf

European Commission's Expert Group on Gender and Employment (EGGE), *Gender segregation in the labour market: Root causes, implications and policy responses in the EU* [pdf], Publications Office of the European Union, Luxembourg, 2009. Available at: http://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=364&furtherPubs=yes

European Commission, *Supporting growth and jobs – an agenda for the modernisation of Europe's higher education systems*, 2011, COM(2011)567 final. Available at: http://ec.europa.eu/dgs/education_culture/repository/education/library/policy/modernisation_en.pdf

European Commission, *Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of The Regions, A Reinforced European Research Area Partnership for Excellence and Growth*, 2012, COM(2012)0392. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52012DC0392

European Commission, *Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of The Regions Entrepreneurship 2020 Action Plan*, 2013a, COM(2012)0795. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012DC0795&from=EN

European Commission, *Gendered Innovations – How gender analysis contributes to research: report of the expert group 'Innovation through gender'*, 2013b https://op.europa.eu/en/publication-detail/-/publication/d15a85d6-cd2d-4fbc-b998-42e53a73a449

European Commission, *Education statistics*, 2014a. Available at: http://ec.europa.eu/eurostat/statistics-explained/index.php/Education_statistics#Data_sources_and_availability

European Commission, *Report on progress on equality between women and men in 2013* [pdf], SWD (2014) 142 final, 2014b. Available at: http://ec.europa.eu/justice/gender-equality/files/swd_2014_142_en.pdf

European Commission, *Human resources in science and technology*, 2014c. Available at: http://ec.europa.eu/eurostat/cache/metadata/en/hrst_esms.htm

European Commission, *Gender Equality in Horizon 2020* [pdf], 2014d. Available at: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/gender/h2020-hi-guide-gender_en.pdf

European Commission, *Education statistics at regional level*, 2015a. Available at: http://ec.europa.eu/eurostat/statistics-explained/index.php/Education_statistics_at_regional_level

European Commission, *High-tech statistics*, 2015b. Available at: http://ec.europa.eu/eurostat/statistics-explained/index.php/High-tech_statistics#Database

European Commission, *Human resources in science and technology: Statistics explained*, 2015c. Available at: http://ec.europa.eu/eurostat/statistics-explained/index.php/Human_resources_in_science_and_technology#Women_in_science_and_technology

European Commission, *Science, technology and innovation: Overview*, 2015d. Available at: http://ec.europa.eu/eurostat/web/science-technology-innovation/overview

European Commission, *Strategic Engagement for Gender Equality 2016-2019*, 2016. Available at: https://ec.europa.eu/anti-trafficking/sites/antitrafficking/files/strategic_engagement_for_gender_equality_en.pdf

European Commission, *Social Dimension of Europe (Reflection paper)*, 2017a. Available at: https://ec.europa.eu/commission/sites/beta-political/files/reflection-paper-social-dimension-europe_en.pdf

European Commission, *MORE3 Survey: Final report (D5)*, European Commission, Brussels, 2017b. Available at: https://cdn1.euraxess.org/sites/default/files/policy_library/final_report_2.pdf

European Commission, *MORE3: EU Higher Education survey results (Second interim report (D2 part 1))*, European Commission, Brussels, 2017c. Available at: https://cdn1.euraxess.org/sites/default/files/policy_library/survey_on_researchers_in_european_higher_education_institutions.pdf

European Commission, *MORE4 Survey* https://www.more-4.eu/

European Commission, *Towards a sustainable Europe by 2030 (Reflection paper)*, 2019. Available at: https://ec.europa.eu/commission/sites/beta-political/files/rp_sustainable_europe_30-01_en_web.pdf

European Commission, *Women in Digital - Shaping Europe's Digital Future*, 2020a. Available at: https://ec.europa.eu/digital-single-market/en/women-ict

European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025*, 2020b, COM(2020)0625 final . Available at: https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation*, 2020c, COM(2020)0628 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

European Commission, *Gendered Innovations 2: How Inclusive Analysis contributes to Research and Innovation*, 2020d. Available at: https://op.europa.eu/en/publication-detail/-/publication/33b4c99f-2e66-11eb-b27b-01aa75ed71a1/language-en

European Commission, *Innovation Union Scoreboard 2021*, 2021. Available at: https://op.europa.eu/en/publication-detail/-/publication/bd128c04-0b93-11ec-adb1-01aa75ed71a1

European Commission/EACEA/Eurydice, *Modernisation of Higher Education in Europe: Academic Staff – 2017. Eurydice Report*, 2017. Luxembourg: Publications Office of the European Union. Available at: https://ec.europa.eu/info/sites/default/files/209_academic_staff_2017.pdf

European Parliament, *Commission Implementing Regulation on statistics science and technology No 995/2012*, 2012. Available at: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:299:0018:0030:EN:PDF

Eurostat, *Nace Rev.2 – Statistical classification of economic activities in the European Community*, Methodologies and Working papers, 2008. Available at: http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

Eurostat, *Key figures on Europe*, 2017 Edition, 2017. Available at: https://op.europa.eu/en/publication-detail/-/publication/ec0bf3c8-bac1-11e7-a7f8-01aa75ed71a1/language-en/format-PDF/source-199900535

Eurostat, *Statistics Explained: Tertiary education statistics*, 2020a https://ec.europa.eu/eurostat/statistics-explained/pdfscache/1152.pdf

Eurostat, *Statistics Explained: Europe 2020 indicators – R&D and innovation*, 2020b https://ec.europa.eu/eurostat/statistics-explained/pdfscache/50448.pdf

Fox, M. F., Whittington, K. B., and Linkova, M. *Gender, (In)equity, and the Scientific Workforce*, 2017. In Felt, U., Fouche, R. Miller, C. A. and Smith-Doerr, L. (ed.), Handbook of Science and Technology Studies (p. 50). MIT Press.

Gvozdanović, J. and Maes, K., *Implicit bias in academia: A challenge to the meritocratic principle and to women's careers - And what to do about it*, League of European Research Universities (LERU) Advice Paper No 23, Leuven, 2018. Available at: https://www.leru.org/files/implicit-bias-in-academia-full-paper.pdf

International Labour Organization, *International classification of occupations (ISCO-08)* [pdf], 2012. Available at: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_172572.pdf

Jappelli, T., Nappi, C.A., Torrini, R., *Gender effects in research evaluation*, Research Policy, 46 (5), 2017, 911-924.

Larivière, V., Ni, C., Gingras, Y., Cronin, B. and Sugimoto, C. R., *Global gender disparities in science*, Nature, No. 504, 2013, pp. 211–213.

NamSor API, *Name Ethnicity and Gender Classifier* https://www.namsor.com/

OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. Available at: http://dx.doi.org/10.1787/9789264239012-en

SerpApi, *Google search API* https://serpapi.com

Snieska, V., Valodkiene, G., Daunoriene, A. Draksaite, A. *Education and unemployment in European Union Economic Cycles, 2015.* Available at: https://core.ac.uk/download/pdf/82041268.pdf

Stanford University, '*Disparities between women and men*'. Available at: https://genderedinnovations.stanford.edu/institutions/disparities.html

Trans Equality Network Ireland. https://www.teni.ie/

UN Women, *Concepts and definitions*. Available at: http://www.un.org/womenwatch/osagi/conceptsandefinitions.htm

UNESCO, *International standard classification of education* [pdf], 2012. Available at: http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

UNESCO, *ISCED Fields of Education and Training 2013* [pdf], 2014. Available at:

http://uis.unesco.org/sites/default/files/documents/isced-fields-of-education-and-training-2013-en.pdf

Vinkenburg, C.J., van Engen, M., Peters, C. P., *Promoting new norms and true flexibility: Sustainability in combining career and care*. In: de Vos, A., & van der Heijden, B., Handbook of research on sustainable careers, 131-145, London: Elgar, 2015. Available at:
https://ideas.repec.org/h/elg/eechap/15416_9.html

World Intellectual Property Organization (WIPO), *International Patent Classification. Version 2015.01*, 2015. Available at:
http://www.wipo.int/classifications/ipc/en/

World Intellectual Property Organization (WIPO), *Identifying the gender of PCT inventors*, 2016. Available at:
http://www.wipo.int/publications/zh/details.jsp?id=4125

Getting in touch with the EU

IN PERSON
All over the European Union there are hundreds of Europe Direct Information centres.
You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

ON THE PHONE OR BY EMAIL
Europe Direct is a service that answers your questions about the European Union.
You can contact this service
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by email via: https://europa.eu/european-union/contact_en

Finding information about the EU

ONLINE
Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU PUBLICATIONS
You can download or order free and priced EU publications at:
https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en)

EU LAW AND RELATED DOCUMENTS
For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

OPEN DATA FROM THE EU
The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

The She Figures Handbook (2021) provides methodological guidance on the calculation of indicators included in the She Figures 2021 publication, the seventh iteration of the European Commission's She Figures publication since the release of its seminal version in 2003.

Organised by data source, information provided on each indicator includes a brief definition, rationale, computation method and any comments or critical issues for the reader to note. The handbook also includes a section on the verification and validation of data that outlines coherence checks and additional data considerations to be taken into consideration in the computation and interpretation of indicators. Finally, the annexes outline important information regarding international classification standards (e.g., ISCED, ISCO) to which data for several of the indicators are tied, as well as key termino-logy and definitions.

The release of the 2021 version of the handbook beyond the groups directly involved in the production of the She Figures publication is intended to strengthen the capacity of other stakeholders to systematically produce meaningful, systematic data on gender in research and innovation.

Studies and reports

Publications Office
of the European Union