

## D 7.3 (D 19) GENDERACTION Dashboard

Project acronym	GENDERACTION
Project name	GENDER equality in the ERA Community To Innovate policy implementatiON
Grant Agreement no.	741466
Project type	Coordination and Support Action
Start date of the project	01 / 04 / 2017
End date of the project	30 / 09 / 2021
Contributing WP	WP7
WP lead partner	ISAS CR
Other partners involved	OPI
Deliverable identifier	D7.3
Contractual delivery date	30 / 09 / 2021
Actual delivery date	30 / 09 / 2021
Deliverable type	Report
Dissemination level	Public



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 741466.

Disclaimer: The views and opinions expressed in this document are solely those of the project, not those of the European Commission.

Revision history			
Version	Date	Created/Modified by:	Comments
0.1	28/09/2021	Aldona Tomczyńska, OPI	Draft deliverable prepared and sent to ISAS
1.0	29/09/2021	Martina Fucimánova, ISAS CR	Deliverable report finalized for submission
2.0	4.1.2022	Aldona Tomczyńska, OPI, Martina Fucimánova, ISAS CR	Deliverable report revised in line with the rejection letter (p. 4)

## **Executive summary**

This deliverable report briefly summarized the work on the Data Dashboard developed in the frame of Work Package 7, task 7.2.6 GENDERACTION Dashboard, providing a technical description of the work delivered.

This dissemination tool has been developed to showcase the results of Work Package 3 analyses of National Action Plans (NAPs) implementation. The Data Dashboard provides in-depth analysis and expert commentary to the reports, in line with European Union documents and GENDERACTION policy briefs.

The Data Dashboard presents visual reports that display important metrics on women in science and higher education. Dashboards help to quickly visualize and drill down on trends and patterns in official statistics in order to answer policy questions. Visualizations and entire data sets can be downloaded and used to create own data summaries.

The Data Dashboard is open to the public, accessed via the project website.

## 1. Introduction

Data Dashboard is a solution consisting of data processing engine written in Java and an interactive visualisation layer developed using modern web technologies (Angular, amCharts). The tool used to create individual reports is reports-engine developed by the National Information Processing Institute (OPI). It is a service supporting the generation of reports based on data mainly from relational databases.

Each report is a complex business object that can consist of sections such as descriptions, charts, maps, tables, and filters. The data presented in a single report can come from many queries from different databases or other sources, such as Eurostat, She Figures. Data are retrieved from these sources with the use of R programming language and analysed in terms of data completeness.

The decision to develop the Dashboard was made as a result of the reallocation of unspent budget of the terminated partner 7 UL and cancelled travels due to the COVID-19 pandemic. The new project beneficiary 19 OPI (former Associated Partner) which has long standing expertise in data mining, visualisation, and development of complex IT systems for science and higher education sector has agreed to develop a new dissemination tool, GENDERACTION Dashboard, to showcase the results of WP 3 analyses of NAPs implementation.

The idea behind the tool was to gather all available data on gender equality in science and higher education and display it in a ready-to-use form. Visualisation of data is an entirely automated process, and the sole role of users is to choose the scope of data they are interested in (countries, group of countries etc.). The main advantages of this approach are:

- statistics from many different sources are available via one webpage: Eurostat, She Figures, indexes regarding countries' characteristics (such as EIGE gender equality index), indexes developed within the GENDERACTION project (NAPs implementation approach), etc.
- The user can easily access data and analyse correlations between quantitative data and qualitative indexes, for example gender equality of researchers in countries without NAPs. Additionally, data dashboard contains expert commentaries on data and links to appropriate reports for further reading.
- Trends in data can be easily depicted as data are cleaned before they are displayed. For example, there are no breaks in time series, as opposed to data downloaded directly from Eurostat.

This approach is complimentary with the project's goal to map and analyse Members States' progress towards implementation of gender equality in R&I.

Apart from this deliverable report, the new task entailed three milestones to monitor progress toward the tool development. The work on this task allowed beneficiary 19 OPI to enhance its expertise in gender equality analyses in R&I. As analyses and statistics are a tool for building arguments for equal and inclusive science, OPI became a reference centre for all scientific institutions in Poland that are interested in developing and implementing gender equality plans.

The sustainability of the Dashboard is fully secured for the next three years in case the follow-up project GENDERACTIONplus is supported. If this is not the case, the tool will operate with data that are already provided and its functionality will be secured by the Coordinator, ISAS CR. Data-cleaning and estimation needed for the tool to function properly are work-consuming. Although the process of downloading new data is semi-automated, the expert's supervision is needed to evaluate the outcome. For this reason, the consortium decided not to update the tool with new data.

## 2. Technical description of the tool

Due to the absence of some values in the source data, it was necessary for them to be estimated. Estimations are done with the use of R programming language and according to statistical standards. Countries with insufficient data to perform estimation are excluded from the estimated dataset but included in the original dataset. A user can choose between these two sets of data.

The missing data for particular years in the estimated dataset is completed separately in each group by category intersection according to the following steps:

1. If the median value in any group equals zero, all missing values in that group are replaced by zero.
2. Missing values located between the reported years are replaced by the mean of the adjacent years (if the missing value occurs in two consecutive years, linear interpolation is used). The maximum length of a replaced gap is two consecutive years.
3. Missing extreme values are replaced either by the prediction of linear regression or an ARIMA model (Autoregressive Integrated Moving Average model). This happens only if following step two, data for at least four years is available.
4. Otherwise, missing values are replaced by zero.

The details of the methodologies used are described in each report provided by the Data Dashboard in the methodology section.

## 3. Objectives of the Dashboard

The main goal of the Data Dashboard is to showcase the results of WP3 analyses of the implementation of NAPs and to present data on women in science from external sources (Eurostat, UNESCO, She Figures report).

The main page presents a typology of countries with respect to NAPs and NAP implementation, as well as three indicators of the socio-economic development:

- 1) **Gender Equality Index 2020**, developed by the European Institute for Gender Equality;
- 2) share of women with grade A among academic staff in 2018, according to the **ERA Progress Report 2018**;
- 3) position in the **European Innovation Scoreboard 2021**, developed by the European Commission. The main page links to individual data dashboards (reports) on topics such as:
  - students: new entrants into higher education sector (Eurostat)
  - students: new entrants of social sciences and humanities (Eurostat)
  - students: new entrants of science, technology, engineering and mathematics (Eurostat)
  - graduates (Eurostat)
  - graduates of social sciences and humanities (Eurostat)
  - graduates of science, technology, engineering and mathematics (Eurostat)
  - employees in knowledge-intensive activities (Eurostat)
  - employees in science and technology (Eurostat)
  - employees in research and development (Eurostat)
  - researchers in the government (GOV) and higher education (HES) sectors (Eurostat, UNESCO)
  - researchers in the academic positions (She Figures)
  - researchers with the title of professor (She Figures)
  - Glass Ceiling Index in academia (She Figures)

## **4. Interactive reports**


Each interactive report contains subsections (views) presenting different aspects of data in a form of charts or maps. A user can interact with a set of filters in order to analyse a given dataset in detail. Additionally, a user can download visualisations from each view (as PNG), as well as a set of data used to prepare visualisations in two different data formats (CSV, XLS). These visualisations and datasets can be used in other reports and analyses (i.e. for the purpose of monitoring trends in male/female participation in different disciplines of study over time or to learn how an approach to NAP implementation correlates with a country's standing in innovation performance). Each report contains a brief introduction and methodological clarifications. Moreover, the storytelling approach to data is applied – the Data Dashboard offers insights into statistics and its meaning (expert commentaries), in line with the European Union documents and making use of GENDERACTION policy briefs.

As a result, grasping the broader context of data presented in the Dashboard is possible. Openness to the public and access via the project website, as well as the variety of applied functionalities, makes it a multi-purpose tool. It can be used by many stakeholders: national authorities and policy-makers, researchers, and experts interested in exploring statistics on women in science.

The Data Dashboard is linked to the GENDERACTION project website and its address is: <https://genderaction-data-dashboard.opi.org.pl>

## 5. Attachments

### 1. Main page (Data Dashboard webpage screenshot)



[Homepage](#) [Newsletter](#) [Twitter](#) [YouTube](#)

## DATA DASHBOARDS

[SEE ALL DASHBOARDS](#)

Data dashboards present visual reports that display important metrics on women in science and higher education. Dashboards help to quickly visualize and drill down on trends and patterns in official statistics in order to answer policy questions. Visualizations and entire data sets can be downloaded and used to create own data summaries.

### WHY SHOULD WE MONITOR DATA ON GENDER EQUALITY IN SCIENCE AND HIGHER EDUCATION?

The main goal of the research policy in the European Union is to create the **European Research Area (ERA)**. ERA is a single research area that is open to participation from all over the world and that enables the free movement of scientists, research knowledge and technology. The European Research Area (ERA) priority 4 focuses on gender equality and gender mainstreaming in research and innovation. The objective is to foster scientific excellence and breadth of research approaches by fully utilising gender diversity and equality and avoiding an indefensible waste of talent.

The main instrument for implementing ERA actions at the national level are so-called **National Action Plans (NAPs)**. Within their NAPs European Union member states and associated countries are asked to develop policies which address gender imbalances, particularly at senior levels and in decision making, and which strengthen the gender dimension in research. In the GENDERACTION we have published two reports on NAPs: **on national roadmaps and mechanisms in ERA priority 4** and **on monitoring of ERA priority 4 implementation**. The results of these analysis are used in the dashboards to approach trends in data from official data sources, such as Eurostat.

### WHAT TO LOOK FOR IN DATA?

The level of implementation of gender equality policies differs from country to country. Multiple data sources (results from an analysis of NAP documents, an online survey, interviews with members of the Standing Working Group on Gender in R&I) were used to develop a typology of countries with respect to NAPs and NAP implementation. Dividing countries taking different approaches to the implementation of NAPs into clusters will be used to present statistics regarding the presence of women in the research and innovation sector. Three other indicators of the socio-economic development of the countries are also taken into account in each dashboard.

- Gender Equality Index 2020, developed by the European Institute for Gender Equality
- share of women with grade A among academic staff in 2018, according to the **ERA Progress Report 2018**
- position in the **European Innovation Scoreboard 2021**, developed by the European Commission

You can learn more about each indicator by looking at maps below. To explore data dashboards on students, graduates, researchers or employees, click on links in boxes at the bottom of this page.

NATIONAL ACTION PLANS

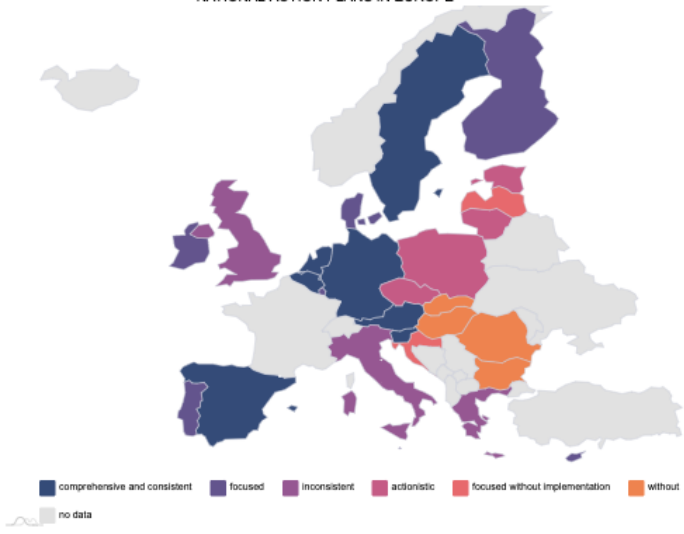
EUROPEAN INNOVATION SCOREBOARD

GENDER EQUALITY INDEX

GRADE A INDEX

NATIONAL ACTION PLANS IN EUROPE

DOWNLOAD



comprehensive and consistent

focused

inconsistent

actionistic

focused without implementation

without

no data

### SIX CLUSTERS OF COUNTRIES:

- comprehensive and consistent:** the NAP covers all gender equality objectives, and corresponds to their implementation
- focused:** two out of three gender equality objectives are addressed in the NAP
- inconsistent:** there are either incompatibilities within the NAP or between the NAP and its implementation
- actionistic:** lack of a context analysis in the NAP but priorities and/or implement measures are formulated
- focused without implementation:** two out of three gender equality objectives are addressed in the NAP, but the implementation is missing
- without:** the NAP either does not exist or it is formulated but gender equality priorities are not included

## DATA DASHBOARDS

### Students

- Students
- Students of social sciences and humanities (SSH)
- Students of science, technology, engineering and mathematics (STEM)

### Graduates

- Graduates
- Graduates of social sciences and humanities (SSH)
- Graduates of science, technology, engineering and mathematics (STEM)

### Researchers

- Researchers in the government (GOV) and higher education (HES) sectors
- Researchers in the academic positions
- Researchers with the title of professor
- Glass Ceiling Index in academia

### Employees

- Employees in knowledge-intensive activities (KIA)
- Employees in science and technology (SAT)
- Employees in research and development (RAD)

## 2. Exemplary report (Data Dashboard webpage screenshot)

### STUDENTS OF SCIENCE, TECHNOLOGY, ENGINEERING AND MATHEMATICS (STEM)

SEE ALL DASHBOARDS

#### New entrants of STEM

This report and its subsequent views present data on female and male tertiary education students (new entrants) in science, technology, engineering and mathematics (STEM). The data pertains to European Union member states and some other countries (mainly associated and candidate countries) for which information is collected by the European Statistical Office (Eurostat), a Directorate-General of the European Commission: Bosnia and Herzegovina, Iceland, Liechtenstein, Montenegro, North Macedonia, Norway, Serbia, Switzerland, Turkey and the United Kingdom.

Due to the absence of some values in the source dataset, it was necessary for some of them to be estimated. Countries with insufficient data to perform estimation have been excluded from the estimated dataset. By default, the estimated dataset is shown. The original data can be accessed using the 'data scope' filter.

Regarding countries' indicators, the main split criterion is the approach to the implementation of National Action Plans (NAPs). General Gender Equality Index, position in the European Innovation Scoreboard and share of female professors among academic staff are also considered. Besides, statistics can be filtered by:

- countries' characteristics (name, date of accession to the EU)
- students' characteristics (sex, education level, discipline)

In this report, the term 'students' covers new entrants - those who, during the course of the reference academic year, entered a programme at a particular level of tertiary education for the first time - irrespective of whether the student entered at the beginning or at an advanced stage of the programme.

For more details, see the methodological clarifications at the bottom of this page.

DATA SCOPE

ORIGINAL VS ESTIMATED DATA

Select

CLEAR

APPLY

OTHER FILTERS

BY SEX

BY EDUCATION LEVEL

BY DISCIPLINE

BY COUNTRY

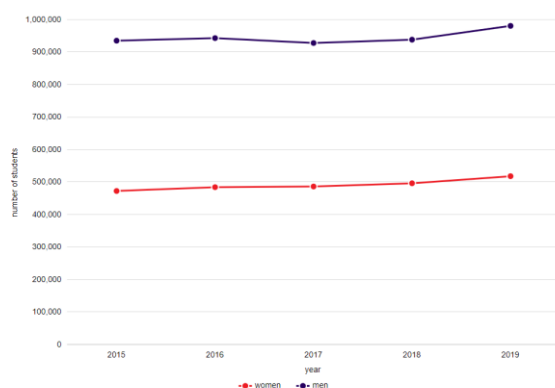
IN TOTAL

The compound annual growth rate of students in STEM from 2015 to 2019 was 1.1% (data available for the EU are considered here) and it was higher than in case of all new entrants, regardless of their education disciplines.

This percentage, however, is much higher among female students (1.9%) than among male ones (0.7%).

DOWNLOAD

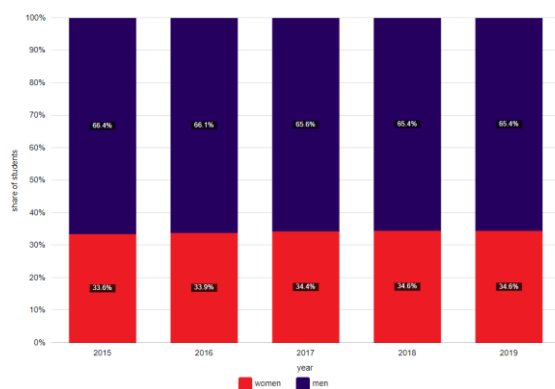
NUMBER OF STEM STUDENTS BY SEX



Source of data: Eurostat [educ\_uoe\_ent02], as of August 3, 2021.

DOWNLOAD

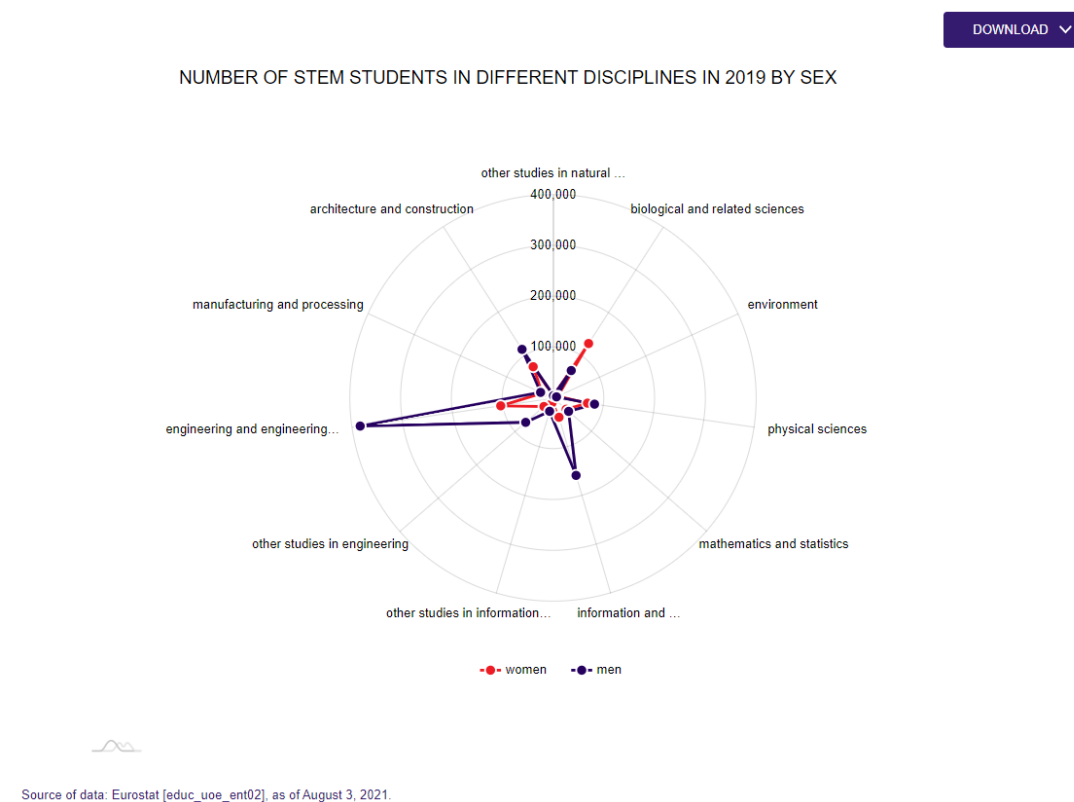
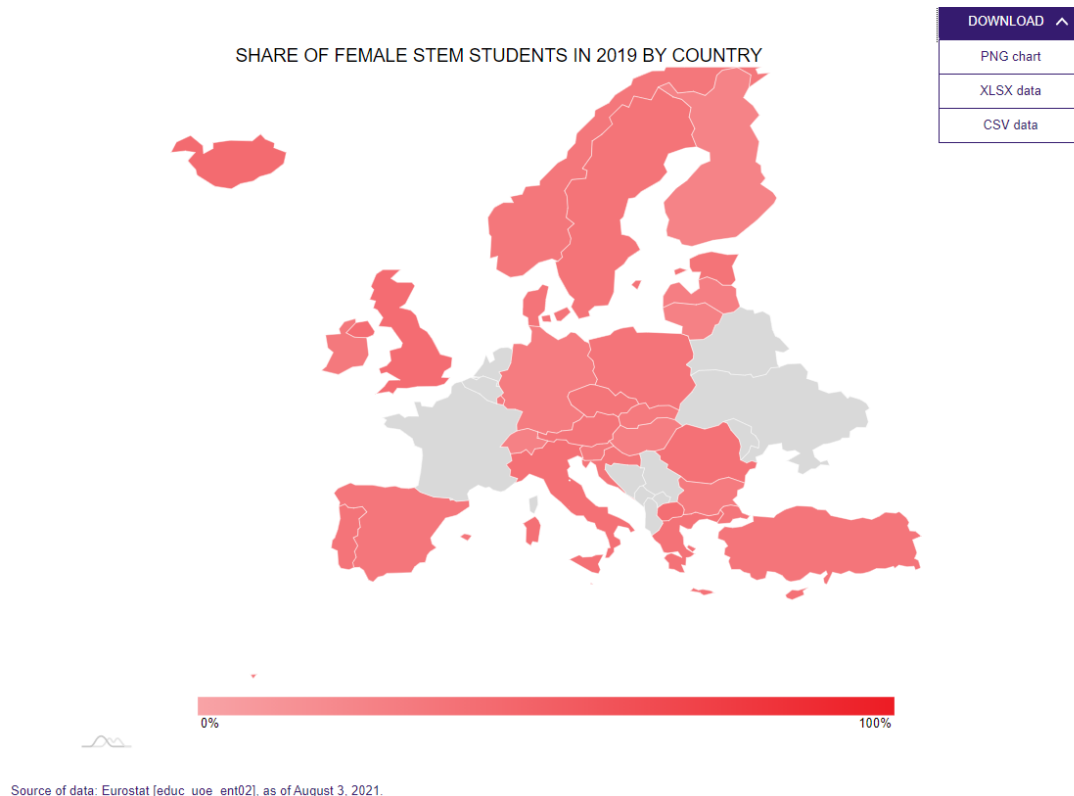
SHARE OF STEM STUDENTS BY SEX



Source of data: Eurostat [educ\_uoe\_ent02], as of August 3, 2021.



### 3. Different kinds of visualisations with the download functionality (Data Dashboard webpage screenshot)



## 4. Expert commentary and methodological clarification (Data Dashboard webpage screenshot)

EXPERT COMMENTARY

New technologies continue to transform the world. Modern economies pay particular attention to the areas of STEM (science, technology, engineering and mathematics), with special emphasis on engineering and ICT. Culturally, these fields remain associated with men; they are stereotypically bonded with male skills and competences.

The underrepresentation of women remains evident, both in the 'old' (EU-15) and the 'new' (EU-13) member states (the 'old' EU presents slightly worse indicators here). Within STEM, women in the European Unions' member states analysed here comprise around 34% of students at all levels of study; if, however, we narrow the analysis to engineering and ICT, that share drops to around 20%.

Particularly in Slovakia, Lithuania, and Slovenia this gap in engineering and ICT is striking – females account for 12-16% of students; whereas in Romania, Sweden, and Denmark this indicator amounts to 25% or higher. This imbalance can impact the shape of the labour market negatively, and translate into reproductions of gender bias and gender stereotypes in the design of robotics, artificial intelligence and ICT technologies ([GENDERACTION 2018](#)). Undoubtedly, actions towards attracting larger numbers of girls and young women should continue regionally, nationally and transnationally. Examples of good practices from different countries can be found in the [GENDERACTION report on monitoring of ERA priority 4 implementation](#).

The compound annual growth rate of students in STEM from 2015 to 2019 was 1.1% (data available for the EU are considered here) and it was higher than in case of all new entrants, regardless of their education disciplines (see [Students](#)). This percentage, however, is much higher among female students (1.9%) than among male ones (0.7%). Within the 'new' member states, this trend is reversed: the number of students is falling (-3.6% between 2015 and 2019), which is particularly apparent in the case of female students (-3.8%).

A correlation can be observed between innovativeness and the growth rate in female students. This relationship is stronger among countries named as innovation 'leaders', or 'strong' or 'moderate' innovators, among 'modest' innovators, the growth is absent. When it comes to NAPs, countries that are actionistic or did not formulate NAPs are also those that present decreasing numbers of STEM students. The same is true for the countries that have the lowest rates in the Gender Equality Index.

METHODOLOGICAL CLARIFICATIONS

The original source of data is the [European Statistical Office \(Eurostat\)](#), a Directorate-General of the European Commission. Its main responsibilities are to provide statistical information to the institutions of the European Union and to promote the harmonisation of statistical methods across its member states, as well as candidates for accession and EFTA countries. The data is presented in line with the international standards of statistics on education, set by the three international organisations that jointly administer annual data collection: Eurostat, UNESCO and the OECD (UOE). Countries compile their data according to the concepts and definitions of the UOE data collection manuals on education systems statistics. The results of UOE data collection on education statistics are compiled on the basis of national administrative sources reported by ministries of education or national statistical offices. Countries provide data from administrative records, on the basis of commonly agreed definitions. The data is collected via data collection tables in electronic questionnaires.

The data pertains to the European Union member states and other countries (mainly associated and candidate countries) for which information is collected by Eurostat: Bosnia and Herzegovina, Iceland, Liechtenstein, Montenegro, North Macedonia, Norway, Serbia, Switzerland, Turkey and the United Kingdom.

Education systems differ between countries. National ISCED mappings that present classification of education programmes and qualifications are available on the [CIRCABC website](#). Detailed methodology for particular countries is available on the Eurostat website in the [National metadata](#) section. Country-specific notes to the data are available in the [Annexes](#) section.

The estimated dataset prepared by the National Information Processing Institute team is shown by default. Due to the absence of some values in the source dataset, it was necessary for them to be estimated. Countries with insufficient data to perform estimation are excluded from the estimated dataset. The missing data for particular years is completed separately in each group by category intersection (in this report: sex, country, education level, discipline) according to the following steps:

1. If the median value in any group equals zero, all missing values in that group are replaced by zero.
2. Missing values located between the reported years are replaced by the mean of the adjacent years (if the missing value occurs in two consecutive years, linear interpolation is used). The maximum length of a replaced gap is two consecutive years.
3. Missing extreme values are replaced by the prediction of linear regression. This happens only if following step two, data for at least four years is available.
4. Otherwise, missing values are replaced by zero.

In the case of disciplines, if the Eurostat total value for a group defined by intersecting sex, country, education level and education field (in this report: education; arts and humanities; social sciences, journalism and information; business, administration and law) exceeds the sum of values for each discipline within that group, the difference is added to one of the disciplines marked as 'other studies in...' (see the definition of education field below). This process was performed both for the original and the estimated datasets.

Definitions used in the report:

- **students**: new entrants to tertiary education (excluding short-cycle programmes) who, during the course of an academic year, entered any programme at a given level of education for the first time – irrespective of whether a student entered a programme at the beginning or at an advanced stage, according to the [OECD Handbook for Internationally Comparative Education Statistics](#). The data on students generally pertains to the count on a given date at the beginning of an academic year – preferably at the end of the first month. Exceptions do exist, however, where enrolments can be counted as averages over multiple counting dates, or as the count of individuals registered during the reference period.
- **science, technology, engineering and mathematics (STEM)**: 1) natural sciences, mathematics and statistics; 2) information and communication technologies; 3) engineering, manufacturing and construction, according to the ISCED Fields of Education and Training classification ([ISCED-F-2013](#)).
- **education levels**: 1) **bachelor's or equivalent** - programmes designed to provide intermediate academic and/or professional knowledge, skills and competencies that lead to a first tertiary degree or equivalent qualification; 2) **master's or equivalent** - programmes designed to provide advanced academic and/or professional knowledge, skills and competencies that lead to a second tertiary degree or equivalent qualification; 3) **doctoral or equivalent** - programmes designed primarily to lead to advanced research qualifications, usually concluding with the submission and defence of a substantive dissertation of publishable quality based on original research, according to the International Standard Classification of Education 2011 ([ISCED-2011](#)).
- **disciplines**: 1) biological and related sciences (in this report, 'biology'); 2) environment; 3) physical sciences (in this report, 'physics'); 4) mathematics and statistics; 5) natural sciences, mathematics and statistics not further defined (in this report, 'other studies in natural sciences'); 6) interdisciplinary programmes and qualifications involving natural sciences, mathematics and statistics (in this report, 'other studies in natural sciences'); 7) natural sciences, mathematics and statistics not elsewhere classified (in this report, 'other studies in natural sciences'); 8) information and communication technologies (in this report, 'ICT'); 9) interdisciplinary programmes and qualifications involving information and communication technologies (in this report, 'other studies in ICT'); 10) engineering and engineering trades (in this report, 'engineering'); 11) manufacturing and processing; 12) architecture and construction; 13) engineering, manufacturing and construction not further defined (in this report, 'other studies in engineering'); 14) interdisciplinary programmes and qualifications involving engineering, manufacturing and construction (in this report, 'other studies in engineering'); 15) engineering, manufacturing and construction not elsewhere classified (in this report, 'other studies in engineering'), according to the ISCED Fields of Education and Training classification ([ISCED-F-2013](#)).
- **compound annual growth rate (CAGR)**: the yearly average rate of growth for a given period.
- **National Action Plan (NAP)**: 1) comprehensive and consistent - the NAP covers all gender equality objectives, and corresponds to their implementation; 2) focused - two of the three gender equality objectives are addressed in the NAP; 3) inconsistent - there are either incompatibilities within the NAP, or between the NAP and its implementation; 4) actionistic - context analysis is absent in the NAP, but priorities and/or implementation measures are formulated; 5) focused without implementation - two of the three gender equality objectives are addressed in the NAP, but have not been implemented; 6) no National Action Plan - the NAP either does not exist, or it is formulated without the inclusion of gender equality priorities: a typology of countries with respect to NAPs and their implementation, according to the assessment performed within Work Package 3 of GENDERACTION project.
- **gender equality index**: a composite indicator developed by the [European Institute for Gender Equality \(EIGE\)](#) to measure the progress of gender equality in the European Union. The indicator consists of eight domains. The six key domains: work, money, knowledge, time, power and health are combined into a core index that is complemented by two additional domains: violence and intersecting inequalities. Each domain is divided into subdomains, which cover the key issues within their respective thematic areas. As the EIGE awards points without constructing any typology of countries, our own original classification of the indicator for 2020 is applied for the purposes of this report: leaders (80 points or more), strong (70-79 points), moderate (60-69 points) and modest (60 points or fewer). For Montenegro, North Macedonia and Serbia, their own internal gender equality indexes are considered.
- **innovation index**: a comparative indicator developed by the European Commission (European Innovation Scoreboard) to measure research and innovation performance in the European Union. The indicator is based on scores for 27 separate indicators, including innovation activities in companies, investment in research and innovation, and human resource and employment elements. Member states fall into four performance groups: leaders, strong innovators, moderate innovators and modest innovators. The scoreboard for 2021 is considered.
- **grade A index**: a headline indicator developed by the European Commission ([ERA Progress Report 2018](#)) to assess the representation of women in the highest echelons of research, as well as identifying the degree to which the glass ceiling continues to limit the professional advancement of women. This indicator can be described as the share of women in grade A research positions (equivalent to full professors in most countries) in the higher education sector; as a percentage of all such positions.

DATA DASHBOARDS

Students

- Students
- Students of social sciences and humanities (SSH)
- Students of science, technology, engineering and mathematics (STEM)

Graduates

- Graduates
- Graduates of social sciences and humanities (SSH)
- Graduates of science, technology, engineering and mathematics (STEM)

Researchers

- Researchers in the government (GOV) and higher education (HES) sectors
- Researchers in the academic positions
- Researchers with the title of professor
- Glass Ceiling Index in academia

Employees

- Employees in knowledge-intensive activities (KIA)
- Employees in science and technology (S&T)
- Employees in research and development (R&D)

 This project is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 741466.

Disclaimer —The views and opinions expressed in this document are solely those of the project, not those of the European Commission. The European Commission is not responsible for any use that may be made of the information it contains.

10