



# 1.1 WOMEN IN DIGITAL PLAYBOOK

WP1

BLU



Funded by  
the European Union

## Technical references

Project Acronym	WIDCON
Project Title	Connecting Women in Digital
Document N° and title	D1.1 WIDCON playbook
Work Package	WP1 – Methodology and Planning
Dissemination level	Public
Lead beneficiary	Brendan Rowan (BLU)
Contributing authors	Leonardo Freitas (IDC), Martina Longo (IDC), Simona Ramanauskaitė (VT), Joana Xhemali (EVTA), Cristian Salis (BLU), Jamie Lavin (BLU)
Reviewers	Daniel Djamo (BDVA), Mattia Trino (BDVA)
Due date	31.03.2025
Actual submission date	31.03.2025

## ABSTRACT

The WIDCON Playbook outlines the methodology and strategic framework for the development of the Women in Digital (WiD) Index and Scoreboard. The Index and scoreboard aim to track and benchmark the participation of women across the digital pipeline, from the beginning of their education in STEM to leadership roles in the digital sector. The playbook defines the pipeline stages, research questions, data sources, stakeholder roles, and the different aspects of diversity. It serves as a guide for WIDCON under the EU Digital Europe Programme and aims to aid in impactful policymaking, promote best practices, and introduce and strengthen different communication channels between Member States, industry, and civil society to address gender gaps in digital professions.

## KEYWORDS

Data, Member States, Index, Methodology, Framework, Gender, ICT Specialists, Digital Decade, EU.

## EXTERNAL CONTRIBUTORS

The authors wish to extend their gratitude to the following external contributors who provided inputs and comments on several aspects:

Christel Widliers (FOD Economie), Ghislaine Ihirwe, (FOD Economie), Olivier Bar (FOD Economie), Yael Pelegrin (EC), Oriane Gilloz (EC), Lucía Fernández Martínez (EC), Fabio Nascimbini (ETF), Glenda Quintini (OECD), Elif Bahar (OECD).

## History of changes

Version	Author	Partner	Date
V0.1	Brendan Rowan (BLU)	Table of Contents	18.02.2025
	Brendan Rowan (BLU)	First Draft Sections 2, 3	27.02.2025
	Jamie Lavin (BLU) Cristian Salis (BLU)	Content Sections 1, 2.1	05.03.2025
	Simona Ramanauskaitė (VT)	Content Section 5.3 & 6.3	07.03.2025
	Martina Longo (IDC) Leonardo Freitas (IDC)	Content Section 5.1, 5.2, 6.2	07.03.2025
	Joana Xhemali (EVTA)	Review Sections 1-3	13.03.2025
V0.2	Brendan Rowan (BLU)	First Review	19.03.2025
	Martina Longo (IDC)	Revision S5.1, 5.3	21.03.2024
	Joana Xhemali (EVTA)	S3.5 Gender mainstreaming added	24.03.2025
V0.3	Brendan Rowan (BLU)	Second Review for external consultation	24.03.2025
V0.4	Brendan Rowan (BLU)	Adjustment to S4.1 Research Questions	27.03.2025
	Jamie Lavin (BLU)	Completion of section Addition of Annexes Executive Summary	27.03.2025
	Brendan Rowan (BLU)	Third Review	28.03.2025
	Daniel Djamo (BDVA) Mattia Trino (BDVA)	Internal review	28.03.2025
V0.5	Jamie Lavin (BLU)	Content transfer	31.03.2025
V1.0	Brendan Rowan (BLU)	Final release	31.03.2025

## DISCLAIMER

The information, documentation and figures available in this deliverable are written by the "Connecting Women in Tech" (WIDCON) consortium under EC grant agreement 101195389 and do not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein

## COPYRIGHT NOTICE

© 2025 - 2026 WIDCON Consortium

## EXECUTIVE SUMMARY

The Women in Digital Playbook serves as a comprehensive guide to support the development of a robust, transparent, and replicable WiD Index to assess the representation and progression of women across digital sectors in Europe.

The playbook showcases how the WiD Index and Women in Digital Forum compliment each other, with the index offering data-driven insights while the Forum will work as a stakeholder collaboration and communication tool. It defines key terms, criteria for index design, and a scope that includes EU Member States and benchmark countries (Brazil, USA, India, UK).

In particular the WIDCON playbook:

- Outlines a data model to be built through the use of both secondary desk research and primary research made through the rollout of a survey.
- Highlights a stakeholder-centered approach for validation and engagement across policymaking, industry, academia, and society as a whole.
- Breaks down all aspects of diversity, intersectionality, and gender mainstreaming to allow for a wholly inclusive form of data analysis.

By following the guidelines and methodologies set forth in the WIDCON playbook, the Connecting Women in Digital Consortium, policymakers and researchers should be able to develop a strong response to closing the gender gap in digital professions, making the WiD Index a strategic resource for policy and synergy development across Europe and benchmarked countries.

## Table of Contents

<b>1</b>	<b><i>Introduction.....</i></b>	<b><i>1</i></b>
1.1.	About Connecting Women in Digital.....	1
1.2.	The Women in Digital Playbook.....	1
1.3.	Purpose of the document .....	2
1.4.	Target audience .....	2
<b>2.</b>	<b><i>General Approach to the Women in Digital Index.....</i></b>	<b><i>4</i></b>
2.1.	Context – Women in Digital declaration .....	4
2.2.	About the Women in Digital index .....	4
2.2.1.	What the Women in Digital index is.....	4
2.2.2.	Strategic purpose .....	4
2.2.3.	Identified stakeholders .....	5
2.2.4.	Principal uses and value proposition .....	6
2.3.	Availability of the index .....	8
2.4.	Overview of the Women in Digital Pipeline .....	10
2.5.	Interrelationship between the Women in Digital Index and the Women in Digital Forum	11
<b>3.</b>	<b><i>Core concepts, criteria and scope.....</i></b>	<b><i>12</i></b>
3.1.	Criteria for index definition .....	12
3.2.	Defined scope .....	12
3.2.1.	Geographical.....	12
3.2.2.	Temporal.....	12
3.2.3.	Sectoral .....	13
3.2.4.	Digital roles .....	14
3.3.	Stakeholder participation and validation .....	15
3.4.	Identification and cooperation with existing initiatives .....	16
3.5.	Anticipated challenges.....	17
3.5.1.	Data availability .....	17
3.5.2.	Depth of data and insights.....	17
3.5.3.	Fragmentation of stakeholders.....	18
3.5.4.	Acceptance and adoption .....	18
<b>4.</b>	<b><i>Principal definitions.....</i></b>	<b><i>19</i></b>
4.1.	Definition of Gender.....	19
4.2.	Diversity beyond gender .....	20

4.3.	Components of diversity .....	20
4.4.	Agreed WIDCON approaches .....	24
4.5.	Gender mainstreaming.....	25
5.	<b>Key research questions.....</b>	<b>27</b>
5.1.	Purpose of research questions .....	27
6.	<b>Index structure .....</b>	<b>29</b>
6.1.	Overview.....	29
6.2.	Quantitative data .....	30
6.2.1.	STEM PHASE.....	30
6.2.2.	ICT PHASE.....	31
6.2.3.	DIGITAL PHASE.....	31
6.2.4.	LEADERSHIP PHASE .....	32
6.2.5.	ENABLING ENVIRONMENT.....	33
6.3.	Member State initiatives and mapping .....	34
6.3.1.	Type of actions .....	34
6.3.2.	Selection criteria .....	35
6.4.	Produced index.....	35
7.	<b>Data collection methodologies and tools .....</b>	<b>37</b>
7.1.	Desk research .....	37
7.1.1.	Identified sources .....	37
7.1.2.	Quality assurance .....	38
7.2.	Women in Digital Survey .....	38
7.2.1.	Sample.....	39
7.2.2.	Sections/segments .....	40
7.3.	Member State initiatives .....	40
7.3.1.	Identified sources .....	40
7.3.2.	Automated collection approach .....	41
7.3.1.	Categorisation and labelling .....	42
7.3.2.	Approximations and limitations.....	42
7.4.	Exploratory approaches.....	43
7.5.	Quality control mechanisms .....	44
	<b>ANNEX 1 – Addressing diversity as a whole.....</b>	<b>45</b>
	Holistic view on diversity .....	45
	<b>ANNEX 2 - IDENTIFIED SECONDARY DATA SOURCES.....</b>	<b>49</b>
	<b>ANNEX 2 – WOMEN IN DIGITAL SURVEY - SAMPLE QUESTIONS.....</b>	<b>51</b>



<i>ANNEX 3 – IDENTIFIED DIGITAL / ICT ROLES .....</i>	<i>52</i>
<i>ANNEX 4 EDUCATION LEVELS CLASSIFICATION (POST 2011) .....</i>	<i>53</i>

Pending approval from EC

## LIST OF FIGURES

Figure 1 Representation of the Women in Digital Pipeline and core stages .....	10
Figure 2 Number of ICT Professionals in employment across sectors. ....	14
Figure 3 Extended Loden Wheel of Diversity and OECD Education model for inclusive education .....	21
Figure 4 Proposed model for definition of diversity (adapted from GENDEX study) .....	22
Figure 5 Tiered vision of the Index comparing Member State Initiatives and quantitative data – Mock-up. Source: WiD .....	36
Figure 6 Proposed model for automated initiative data extraction .....	41
Figure 7 Plan for LinkedIn skills mapping .....	43
Figure 4. Proposed model for definition of diversity (adapted from GENDEX study) .....	45

## LIST OF TABLES

Table 1 Identified ICT roles by sector .....	15
Table 2 Definition of gender: Key components .....	19
Table 3 Quantitative metrics to be included in the STEM phase .....	30
Table 4 Quantitative metrics to be included in the ICT phase .....	31
Table 5 Quantitative metrics to be included in the DIGITAL phase .....	31
Table 6 Quantitative metrics to be included in the LEADERSHIP phase .....	32
Table 7 Quantitative metrics to be included regarding women in the labour force .....	33
Table 8 WiD survey's sampling approach .....	39
Table 9 Data quality control mechanisms .....	44

## ABBREVIATIONS

<b>AI</b>	Artificial Intelligence
<b>API</b>	Application Programming Interface
<b>BDVA</b>	Big Data Value Association
<b>BLU</b>	BluSpecs
<b>CATI</b>	Computer-Assisted Telephone Interviewing
<b>CAWI</b>	Computer-Assisted Web Interviewing
<b>CEMA</b>	Central and Eastern Europe, Middle East, and Africa (IDC region)
<b>CEO</b>	Chief Executive Officer
<b>COO</b>	Chief Operating Officer
<b>DDPP</b>	Digital Decade Policy Programme
<b>DG</b>	Directorate-General (European Commission)
<b>EC</b>	European Commission
<b>EQF</b>	The European Qualifications Framework
<b>EU-27</b>	The 27 European Union Member States
<b>EVTA</b>	European Vocational Training Association
<b>FTSE</b>	Financial Times Stock Exchange
<b>GENDEX</b>	Gender and Diversity Index
<b>GenAI</b>	Generative Artificial Intelligence
<b>ICT</b>	Information and Communication Technology
<b>IDC</b>	International Data Corporation
<b>ISCO</b>	International Standard Classification of Occupations
<b>MBA</b>	Master of Business Administration
<b>ML</b>	Machine Learning
<b>NACE</b>	Statistical Classification of Economic Activities in the European Community
<b>NLP</b>	Natural Language Processing
<b>OECD</b>	Organisation for Economic Co-operation and Development
<b>PISA</b>	Programme for International Student Assessment
<b>RRF</b>	Recovery and Resilience Facility
<b>TIMSS</b>	Trends in International Mathematics and Science Study
<b>VET</b>	Vocational Education and Training
<b>VT</b>	Vilnius Gediminas Technical University
<b>WiD</b>	Women in Digital

# Introduction

## About Connecting Women in Digital

The purpose of Connecting Women in Digital is to drive forward the impact of the investments made throughout the Digital Europe Programme (DIGITAL), leveraging data, knowledge, networks and communities to deliver on the ambitions of the Digital Decade and secure global competitiveness, resilience, and prosperity for European industry. It acts at the core of the Women in Digital Declaration to leverage women's contributions for tackling the shortage of Information and Communication Technology (ICT) Specialists and building an inclusive digital Europe.

As such, the project will source progress across the EU-27 and selected reference countries, identifying best practices and establishing widely accepted benchmarks. It also aims to cultivate a community of experts to address key challenges, validate data, and exchange insights. Knowledge generated in these activities will inform targeted analyses and recommendations to support strategic policy objectives and enhance the recognition of relevant networks.

To achieve the project and wider policy objectives, Connecting Women in Digital will deliver:

- The Women in Digital index with an overview of national initiatives by Member State.
- A sustainable expert Forum.
- A best practice repository accessible by all relevant stakeholders.
- Annual reports on the state of Women in Digital, feeding into an annual summit event.

## The Women in Digital Playbook

The Women in Digital playbook is a document which summarises the approach and methodologies that will be followed for the collection and analysis of primary and secondary data, as part of the project; in particular for the index and scoreboard development. It addresses the underlying framework, i.e. the leaky pipeline of Women in Digital and the key research questions along its conversion points. Designed for relevance, feasibility, and replicability, the playbook will be validated by domain experts to ensure the effectiveness and sustainability of the project's proposed framework.

In the following sections, the playbook provides:

An overview of the adopted approach, covering the leaky pipeline of Women in Digital and key conversion points.

- A description of the interrelationship between the Women in Digital index and scoreboard, and the Forum, explaining how each will be leveraged to best answer different questions regarding the participation of women in ICT.
- An outline of the minimum criteria for the index, alongside its geographical, temporal and sectoral scope.
- The principal definitions of gender, digital roles and intersectionality, and how they will be addressed by the index and project as a whole.
- The approach to stakeholder validation and participation, and gender mainstreaming.
- The definition and purpose of the key identified research questions across the “STEM, ICT, DIGITAL and LEADERSHIP” phases of the leaky pipeline.
- An overview of the index structure, showcasing the quantitative data from every phase, and initial mapping of EU Member State initiative for Women in Digital.
- A description of the quality control processes, including the defined quality criteria and key mechanisms.

## Purpose of the document

The purpose of the playbook is to provide a structured point-of-reference resource that will guide the implementation of the Women in Digital index and scoreboard development. By establishing a clear and rigorous methodology, the playbook aims to support the generation of high-quality, reliable and consistent results, as well as replicable framework that will enable stakeholders to understand and/or implement the proposed approach to achieve consistent outcomes.

It is an important piece of the project’s aspirations, as it will establish a common understanding among the consortium partners, define the scope of all subsequent activities and break down key processes, increasing efficiency and minimising risks associated with the development of the index and scoreboard.

## Target audience

This document has three principal audiences:

- Connecting Women in Digital Consortium members
  - Purpose is to act as a guide for the development of the index and scoreboard. The key definitions, processes and methods described herein will support them in scoping the work and ultimately meeting the pre-determined project objectives.
- Policy Makers
  - This document serves to provide key stakeholder with a reference for the definition of scope, intention and approaches to be taken in the construction of the WiD Index that facilitates their contribution to the activities and results of the Connecting Women in Digital initiative.



- Researchers
  - Relevant researchers from academia, think tanks, consultancies and similar are able to identify relevant outcomes from the Women in Digital approach to their work and propose opportunities for collaboration or further extension of the work.

Pending approval from EC

# General Approach to the Women in Digital Index

## Context – Women in Digital declaration

In 2022, the European Commission launched the ‘Digital Compass & Digital Decade Policy Programme 2030 (DDPP)’ setting specific targets for 2030 in areas like digital skills, digital infrastructure, and making businesses and public services more digital. These initiatives highlight the importance of increasing the number of women in ICT professions, as well as the number of girls and women studying ICT, both at school and university.

The participation of girls and women in digital, is important for Europe to ensure:

- Equal opportunities in the digital world.
- Highly skilled workforce in line with the Digital Decade target of 20 million employed ICT specialists by 2030.
- More diverse and gender-balanced teams in innovation.

Previously, in 2019, 26 EU Member States, along with Norway and the United Kingdom, signed the “Women in Digital Declaration”. They committed to bring together government, businesses, and the community to make sure more girls and women participate on an equal basis in the digital world.

Under the Declaration, the Member States, supported by the European Commission, commit to encouraging and empowering women to play a more active role in the digital age by:

- Promoting digital skills and education.
- Challenging digital gender stereotypes.
- Advocating for more women entrepreneurs.

## About the Women in Digital index

### *What the Women in Digital index is.*

The Women in Digital Index (WiD Index) is a data-driven tool developed by the WIDCON consortium to track how women progress through education and careers in ICT across the EU and selected benchmark countries. It combines primary data, secondary data and expert input to identify where women drop out of the Women in Digital pipeline, starting from ICT and finishing with leadership roles. By comparing national initiatives and outcomes, the index highlights gaps, showcases best practices, and best performers across Member States and benchmarked countries.

### *Strategic purpose*

The data on the participation of women in digital is clear. There is little progress being achieved in the past decade across economies and societies within the EU.



The WiD Index builds on the previous work within the Digital Economy Social Index which publishes an annual Women in Digital Scoreboard based on selected metrics from within their survey.<sup>1</sup> This is now integrated into the monitoring of the progress of the Digital Decade.

There exists, however, a limitation to the data collected and provided within this framework due to various dependencies and resources in the collection of annual surveys and statistical accounts.

The Connecting Women in Digital initiative serves the principal Digital Decade targets of achieving 20 million employed ICT Specialists and greater gender convergence. The achieving of such targets depends on many factors that influence the pathways for women in the form of both push and pull, acting within a societal bias related to gender roles, professions and subjects. For this reason, a whole of pipeline view is required within the index, joining the dots and the data.

Thus, the WiD Scoreboard holds 4 principal strategic purposes:

- Provide a whole pipeline view of progression of women from STEM subjects to ICT and digital roles across the EU delivering a benchmark.
- Support the identification and articulation of key weak points in the pipeline to enable targeted actions.
- Facilitate a facile comparison of progress and progression across Member States to reach common targets.
- Equip decision-makers with the evidence base required to enact change and advocate for areas of underinvestment.

### *Identified stakeholders*

Across the potential stakeholders there exists the following primary stakeholders:

- POLICY MAKERS
  - Member States – Ministries with responsibility for Digitalisation, Economy and Education, Labour and Social Affairs, among others.
  - European Commission – DG CNECT, DG EMPL, DG EAC, DG RTD.
- INDUSTRY
  - Corporates – COO, CEO, HR, Investor Relations
  - Industry Associations – Policy Working Groups
- EDUCATION AND TRAINING
  - Higher Education

---

<sup>1</sup> Women in Digital Scoreboard 2024 (2024) Shaping Europe's digital future



- Public VET providers.<sup>2</sup>
- SOCIETY
  - All citizens

### *Principal uses and value proposition*

#### POLICYMAKERS

USER	PRINCIPAL USES	VALUE PROPOSITION
MEMBER STATES	<ul style="list-style-type: none"> <li>Support to the completion of Women in Digital Strategies.</li> <li>Support to the completion of Digital Decade Strategic Roadmaps.</li> <li>Enable common categorisation of actions and reporting.</li> <li>Sourcing of best practices from peers.</li> <li>Localisation of needs for policy development.</li> <li>Evidence bases for whole of government approaches.</li> <li>Enable cross-departmental or inter-ministerial collaboration (e.g., Education, Labour, Gender Equality)</li> </ul> <p>Supporting cross-border collaboration between governmental entities through its database (in Europe and beyond).</p>	<p>WiD Index provides a consolidated source of data on the topic of women in digital, allowing for easy access and preparation of own strategies.</p> <p>WiD Index provides a harmonised structure that allows individuals to make direct comparisons and identify areas of strength.</p> <p>WiD supports coordinated action across diverse governmental domains like digitalisation, gender equality, and education.</p> <p>WiD will become the go-to platform for governmental bodies and policymakers interested in bridging the gender gap in the digital sector and attracting in women tech in general (in Europe and beyond).</p> <p>(potentially) WiD offers the possibility of updating the content of its website (even after the project is over) to governmental bodies interested in updating and sharing information. User accounts can be created for content providers.</p>
EUROPEAN COMMISSION	<ul style="list-style-type: none"> <li>Integration in annual Digital Decade reporting.</li> <li>Exploration and sourcing of key data for strategies and action plans.</li> <li>Sourcing relevant data on Member States and whole Union progress to contextualise reporting.</li> <li>Sourcing good practice and champions.</li> </ul>	<p>Provides an easy to navigate source of key metrics that covers areas across the EC.</p> <p>WiD Index is a one-stop shop for analysis and sourcing of actions in Member States.</p> <p>WiD Index is objective and provided by external entities.</p> <p>WiD Index offers a thorough analysis of the gender gap in the digital sector</p>

<sup>2</sup> Recognising the role of upskilling and reskilling is an important aspect to the topic; private VET is a key stakeholder but not a primary stakeholder as an index user due to the relative independence from policy but firmly remains an interested party in all activities of Connecting Women in Digital.

	<ul style="list-style-type: none"> <li>• Provide baselines for common actions and joint instruments.</li> </ul> <p>A manual built on strategies that can attract and retain the women workforce in the digital sector.</p>	
--	--	--

## INDUSTRY

USER	PRINCIPAL USES	VALUE PROPOSITION
CORPORATE	<ul style="list-style-type: none"> <li>• Referencing the benchmark for drafting annual reports in compliance with the CRSD.<sup>3</sup></li> </ul> <p>Identification of challenges and common barriers in attracting and retaining women in digital roles.</p>	<p>WiD provides a common definition for digital roles and whole pipeline view to understand dependencies.</p> <p>WiD provides an external view on the challenges and barriers facing organisations, acting as a benchmark.</p> <p>Provides valuable insights on where women are underrepresented and where targeted recruitment efforts are needed.</p>
INDUSTRY ASSOCIATIONS	<ul style="list-style-type: none"> <li>• Access to baselines for supporting member activities.</li> </ul> <p>Identification of challenges and common barriers in attracting and retaining women in digital roles.</p>	<p>WiD provides calls to action for members to develop positions and responses.</p> <p>WiD offers an up-to-date European-level overview of the challenges related to retaining women talent in the field and identifies potential solutions, valuable to members.</p>

## EDUCATION AND TRAINING

USER	PRINCIPAL USES	VALUE PROPOSITION
HIGHER EDUCATION	<ul style="list-style-type: none"> <li>• Referencing key challenges and barriers facing the attraction and completion of ICT and Engineering programmes.</li> </ul> <p>Understand the structural weaknesses of the pipeline to develop mitigation/proactive engagement in secondary education.</p>	<p>WiD Index provides international benchmark for individual institutions.</p>
PUBLIC VET	<ul style="list-style-type: none"> <li>• Identify weaknesses in the VET sector for action development.</li> </ul>	<p>WiD Index provides the VET community with comparable data.</p>

<sup>3</sup> Corporate Sustainability Reporting Directive (CSRD) – Directive (EU) 2022/2464

	Compare progression across Member States to identify good practice.	<p>WiD Index addresses VET as a core of the pipeline view.</p> <p>WiD index allows identifying trends in women enrolment and completion rates, helping to address drop-off points and implement strategies to retain girls and women in digital-focused VET programs.</p> <p>WiD index enables public VET providers to align their curricula/training programmes with the digital sector's evolving needs, ensuring women are better positioned to enter and succeed in digital careers.</p>
--	---	--

## SOCIETY

Across all of society, there is an interest in viewing the progression of women in digital and the fostering of mutual responsibility and accountability for the promotion of women in digital. The WiD Index serves as an important tool for this purpose. It allows individuals, communities, and organisations to easily track how women are advancing in digital careers.

It also encourages reflection on how gender norms and roles in European society contribute to the gender gaps in digital sectors. These societal biases often limit women's opportunities and career choices. The WiD Index helps identify these barriers and informs discussions on how policy and education reforms can address them. The WiD Index will also be an important reference for press coverage of the topic and act as a touchstone for the effectiveness of public and private investment.

### Availability of the index

In order to reach identified users, the following mechanisms are foreseen as part of the Connecting Women in Digital activities:

- **Public dashboard:** Online dashboard that allows users to explore the index and underlying data. It will allow a view on the identified initiatives in individual Member States and where possible, provide EU and international benchmarks.
- **WiD forum:** A platform used to bring together a pan-European community of experts will serve to disseminate and access the key findings and allow the interrogation of the resulting index and results.
- **State of WiD report:** The annual flagship report will provide analysis and commentary on the index and will allow users to contextualise individual metrics and insights.

Together, these actions will work in conjunction to enhance visibility, promote collaboration, and provide lasting resources to help close the gender gap in the digital sector.

Pending approval from EC



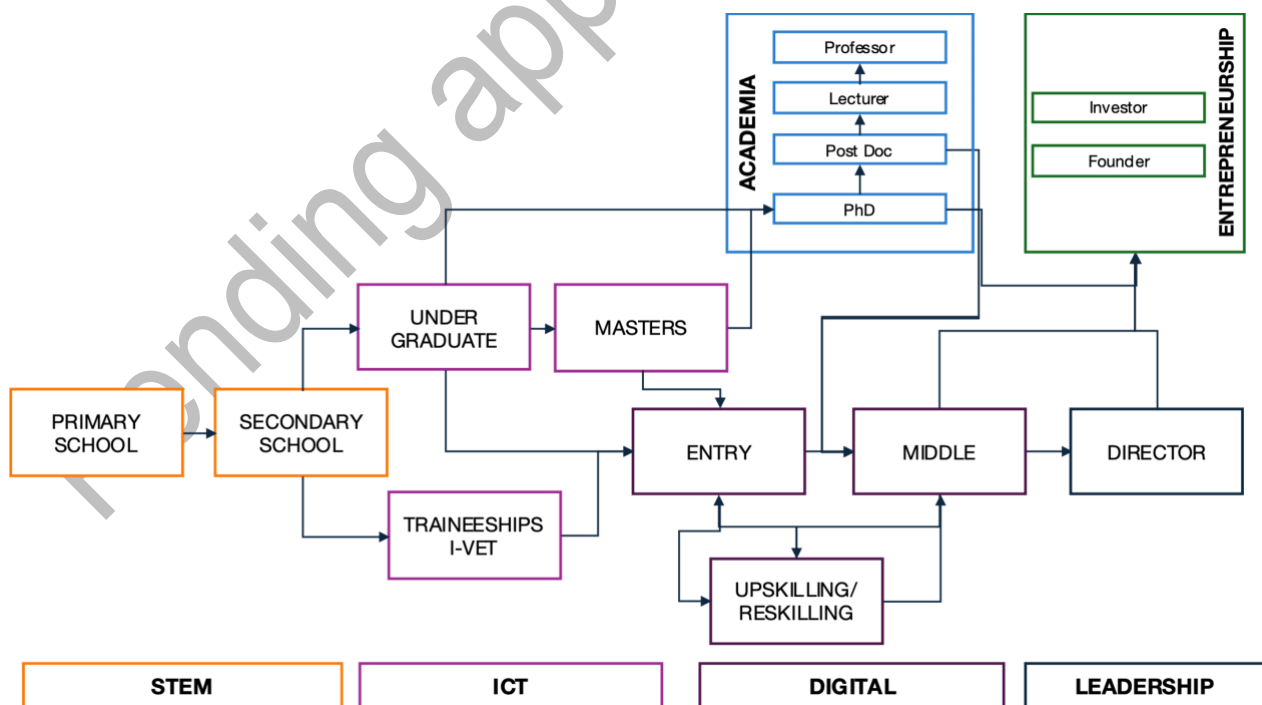
## Overview of the Women in Digital Pipeline

As previously mentioned, the central approach of the Connecting Women in Digital consortium is to address the full pipeline that is responsible for women reaching and remaining in digital roles. On the converse, it supports the identification of weak points, and significant losses of women and girls from the talent pipeline.

The Women in Digital Pipeline has been divided in four principal phases that reflect core stages of an individual's education and career paths. These are:

- **STEM**
  - Stages: Primary and Secondary Education
  - Core points: participation of girls in STEM subjects
- **ICT**
  - Stages: Higher Education, I-VET
  - Core points: participation of women in ICT and related subjects
- **DIGITAL**
  - Stages: Early-to-middle career
  - Core points: participation of women in digital roles
- **LEADERSHIP**
  - Stage: Leadership
  - Core points: participation of women in leadership roles

*Figure 1 Representation of the Women in Digital Pipeline and core stages*



Source: Connecting Women in Digital

This pipeline and principal phases are the framework that the Women in Digital Index will be constructed, combining publicly available statistical datasets, surveys, secondary research and the sourcing of public investments and actions within each phase and individual stage. Through this view, it will be possible to identify poor rates of conversion, high levels of retention, an imbalanced investment across the whole pipeline or opportunities for targeted action.

### Core conversion points

Within the proposed approach to developing the Index, a specific weighting is not foreseen for the first edition, however, this may be under review after data collection and analysis is completed.

What are of interest is to measure the conversion between each stage and grouped phase to provide a simple and direct view on the rate of loss, retention or growth across the EU.

## Interrelationship between the Women in Digital Index and the Women in Digital Forum

Within the Connecting Women in Digital funded actions, there are two key activities; the first that is described in detail in this document is the WiD Index, the second is the establishment of a network of experts within the Women in Digital Forum.

The purpose and structure of the Women in Digital Forum is to convene European (and international) experts to deliver constructive debate and specification of actions for overcoming identified barriers or gaps.

To this end, the developed index will be a source of challenges to seed the Thematic Working Groups of the Forum, and at the same time, the Forum offers the platform to perform deep dives or developed discourse on the conversion points that is not readily described or evident from the quantitative data. In fact, the nuance and the confirmation of particular issues, often anecdotal in nature, is expected to be delivered by the Forum, complementing the index itself.

## Core concepts, criteria and scope

### Criteria for index definition

The design and construction of the WiD Index must meet the following criteria. It must be:

- **TRANSPARENT:** The developed index should be evident and open in its formation; users should be able to readily understand how and where data has been collected and the inherent biases or ambiguity that is present. The index should clearly communicate where approximations are made and inclusion and exclusion data for various metrics.
- **ROBUST:** The collected data and analysis should be of a sound basis, with reliable sources that is verifiable and validated. Applied methods should be grounded in published and recognised research and evidence.
- **RELEVANT:** Provided data and analysis must be specific and relevant to the core mission of supporting the advancement of securing a greater participation of women in digital roles in industry.
- **SUSTAINABLE:** Selected methods and approaches must allow for a reproducible result and, where possible, take into account reliable data sources into the future. The index must be feasible to continue for years to come with reasonable resources and effort.

In the definition of the methodology, and specifically the methods for data collection and validation, the responsible team member must cross-reference the above criteria and provide a justification against the same.

### Defined scope

#### *Geographical*

The index should provide data and coverage in the following fields:

- EU – Whole of EU
- EU-27 – Individual Member States
- Selected international benchmark countries
- Brazil
- USA
- India
- UK

#### *Temporal*

The Connecting Women in Digital project acts across the whole pipeline which has a significant lag from actions taken in STEM phases through to the DIGITAL phases i.e. time

to outcome. In addition, there are regular studies and surveys published from DIGITAL and LEADERSHIP phases, for this reason there is a need to act on two levels:

- STEM/ICT – Data should strive to provide as a minimum a 10-year historical trend
- DIGITAL/LEADERSHIP – Data should be from the previous 4 years as a maximum.

### *Sectoral*

Digital is a horizontal domain where digital roles are found not just in the ICT Sector (NACE code K), but also across all sectors. In particular it is important to account for the significant role of consultancies in developing and deploying digital solutions for various industries.

- Priority sectors:<sup>4</sup>
- K. Telecommunication, Computer Programming, Consulting, Computing Infrastructure and Other Information Service Activities
- 61 Telecommunication
- 62 Computer programming, consultancy and related activities
- 63 Computing infrastructure, data processing, hosting and other information service activities
- N. Professional, Scientific and Technical Activities
- 70.2 Business and other management consultancy activities
- 71.12 Engineering activities and related technical consultancy
- 72.1 Research and experimental development on natural sciences and engineering

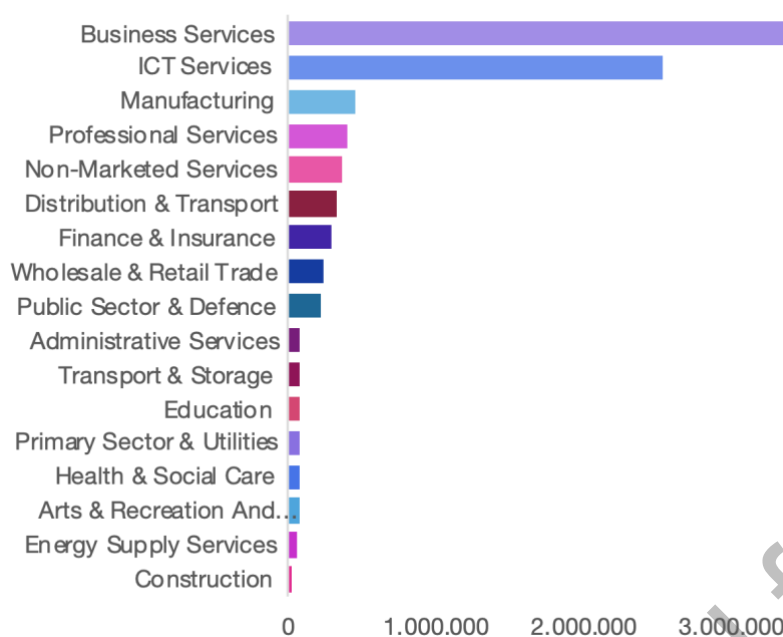
The definition of Digital Roles and the reach beyond ICT Sector are further discussed below. In the following graph, where the participation of ICT Professionals in sectors is mapped, it is of note that the majority are not found within ICT Sector itself but in the Business Services category, which validates the scoping of including the NACE codes 70.2 in the selection of data. There is also a evident need to include a particular emphasis on Manufacturing, Logistics, and Finance sectors which provide a significant employment of ICT Professionals.

---

<sup>4</sup> Current datasets still use older codes from NACE2.0, here are provided for NACE2.1



*Figure 2 Number of ICT Professionals in employment across sectors.*



Source: Employment of ICT professionals in EU in 2022 across sectors - CEDEFOP

Outside the central scope of ICT Specialists, there exists the opportunity to study other sectors such as healthcare, education and legal services that can draw parallels and provide sources of inspiration and comparison to resolve the core challenges of Women in Digital.

### *Digital roles*

Digital roles, while predominantly part of the ICT sector and thus classified under the ICT Specialists,<sup>5</sup> in particular this includes the following ISCO occupations:

- OC25 Information and Communications Technology Professionals
- OC35 Information and Communications Technicians

There exists, however, more occupations not covered within these categories that may be considered:

- OC215 Electrotechnology Engineers
- OC742 Electronics and telecommunications installers and repairers
- OC31 Science and Engineering Associate Professionals
- OC2166 Graphic and multimedia designers
- OC2356 Information technology trainers
- OC2434 ICT sales professionals

<sup>5</sup>Eurostat (isoc\_skslf)

[https://ec.europa.eu/eurostat/cache/metadata/de/isoc\\_skslf\\_esms.htm#shortdata\\_descr](https://ec.europa.eu/eurostat/cache/metadata/de/isoc_skslf_esms.htm#shortdata_descr) Disseminated

- OC3114 Electronics engineering technicians

Further to this, most data related to job titles and roles does not necessarily follow the ISCO framework. Based on the previous work provided by IDC<sup>66</sup>, the following taxonomy, while not exhaustive provides a more market centred definition of roles, independent of sector:

*Table 1 Identified ICT roles by sector*

Applications	IT Infrastructure	Cybersecurity	Data Management/ Analytics	Technical Support	IT Leadership
Software Developer Engineer	Network Engineer/ Network Architect	Cybersecurity / Information Security Analyst	Database Administrator (DBA)	IT Project Manager / Specialist	Chief Information Officer / Director of IT
Web Developer	Infrastructure and Cloud Architect	Security Management Specialists	Business Intelligence Analyst/Developer	User Support Specialists	Enterprise Architect
UX / UI Designer / Developer	Systems Administrator	Information Security Architect/Data Engineer	Data Scientist	Computer User Support Specialist	Information Systems Manager
Mobile Applications Developer	Cloud Solutions Architect	Data Analyst	Network Support Specialist	Chief Technology Officer (CTO)	Technical Writer
ML Designer / Developer		Applications Manager / Business Systems Analyst	IoT Designer / Developer	Systems Analyst	Data Engineer

Source: IDC Job Role Taxonomy

There is a required matching of terms and taxonomies to create coherence across all phases of the Women in Digital framework, but we highlight the need to address roles as well as sectors to provide accurate view on the participation of women in digital.

## Stakeholder participation and validation

Stakeholder engagement is extremely important for the successful implementation of our project, and in order to achieve our expected outcomes. Beginning with the creation of the Women in Digital Index and scoreboard, stakeholders will contribute through structured surveys and expert roundtables. This engagement will allow for us to gather insights so that we may update the 'Women in Digital Index and Scoreboard', ensuring that the data is as up to date as possible, and allowing us to make strategic decisions regarding the promotion of women in the ICT and STEM sectors.

<sup>66</sup> IDCs Worldwide ICT Job Role Taxonomy 2022

In a similar manner, with the establishment of the Women in Digital forum, stakeholder engagement is extremely valuable, not just with the project, but also with each other. One of the main objectives of the forum is to create a platform in which relevant stakeholders can interact and potentially foster collaboration and synergies with other existing initiatives and working groups. The Forum will implement robust mechanisms to enhance collaboration, transparency, and responsiveness, including

- Monitoring and evaluation of engagement activities
- Stakeholder feedback loops
- Clear communication channels
- An adaptive and flexible approach
- A strong post-project sustainability strategy

Through these mechanisms, the forum is committed to maximising the impact and relevance of its stakeholder engagement efforts, ultimately driving meaningful change in advancing gender diversity within the ICT sector.

The publication of the State of WiD report will offer stakeholders an in-depth view of the biggest challenges and most successful practices across the EU and the four benchmarked countries. The report will combine the best practices that we have collected, along with the white papers developed in the working groups. This report will also highlight the contributions of stakeholders throughout the duration of the project, which should ensure that the community that we engage with throughout the duration of this project will remain strong long after the end of the project.

## Identification and cooperation with existing initiatives

Identifying and cooperating with initiatives that are actively involved in advancing girls' and women's participation in ICT is key to the successful implementation of the WIDCON project, as well as the development of the WIDCON index and scoreboard. As such, a robust stakeholder mapping and engagement process has been put in place. This process will not solely act as a guiding framework but will also act as an operational tool for the identification, classification, recruitment and involvement of key actors throughout the lifecycle of the WIDCON project.

WIDCON will identify key gaps across the "leaky pipeline" of women in ICT, using these insights to inform the mapping and engagement of relevant initiatives and their stakeholders. The consortium will collaborate with a variety of existing efforts, such as #BeDigitalTogether, Women in Tech, She Figures, and many more, with the goal of exchanging knowledge, aligning strategies, and amplifying impact.

These partnerships will involve activities such as data-sharing and validation workshops, co-hosted events to promote women's participation in the digital sector and coordinated outreach via platforms like LinkedIn and initiative partner networks. Insights gained because of these collaborations will help with the design of the Women in Digital Forum, guiding the

themes for its working groups, governance model, and stakeholder engagement. This approach will help ensure WIDCON delivers not just short-term solutions but a sustainable model for the long-term pursuit of our goal of gender equality in the digital sector.

In order to achieve the maximum value possible from the identified initiatives, a detailed database of key stakeholders is to be created, documenting their roles, interests, influence and potential contributions to the objectives of the forum. This database is to be used as a strategic asset, allowing the project to monitor participation, ensure adequate representation, and promote cross-sector engagement and cooperation.<sup>7</sup>

## Anticipated challenges

### *Data availability*

- Some data are collected by different countries and organizations, but in a different format and using different taxonomies
- Data collection frequency:
  - Lag of data collection (e.g. Eurostat)
  - Different time frameworks and frequency of data collection (e.g. every five years for OECD' PISA and every four years for TIMSS)
- Missing data for some (usually smaller) EU Member States (e.g. Eurostat for Malta, but sometimes also bigger states such as Bulgaria in Education phase)
- Missing data availability for the Digital Phase in WiD Index specifically leads to primary research. Survey results provide a trend, but are not representative of entire population, and also provide a snapshot in time (commencing 2025)
- Alternative data collection methodologies are being considered (e.g. LinkedIn and online job portals), but technical limitations may hinder viability, resulting in relying on other data collection methodologies (e.g. primary research)
- Desk research activities for collecting data from company's reports may limit the sample of analysis to only the top ICT companies operating in EU

### *Depth of data and insights*

- Current data availability for Member States will allow for having at least the EU27 average for each indicator in the WiD index, but complete representativeness of smaller countries is not granted
- Given the point, analysis will not take into account single country fragmentation in different regions/federal states (though acknowledging relevant differences may be present)

---

<sup>7</sup> WIDCON – 'D3.1 Stakeholder mapping – Database of stakeholders actively engaged in advancing girls and women's participation in ICT, categorised based on roles, interests, influence, and potential contributions to the forum objectives.'

- This accounts also for WiD survey results where statistical representation needs to be balanced with limitations on primary filed activities (time/cost/efforts/sample availability)

### *Fragmentation of stakeholders*

- The topic of the participation of women spans many areas, by sector, by technology field, by type of activity and across the whole pipeline. This translates to a community of active stakeholders that exist in silos, with each addressing one aspect of the Women in Digital challenge.
- The scope of the EU level present geographical limitations on the activities on national level and may not be feasible to engage all in the activities of the forum or the collection of data.

### *Acceptance and adoption*

- One of the most prominent challenges, particularly in developing the index centres on the acceptance by target users and stakeholders. For Member States, there is a risk that the index is perceived as an external scorecard rather than a support mechanism to their own strategies and policy developments. There is a need to ensure that Member States are engaged early on, made aware of the activities and provided with initial drafts of reports to correct any gaps.
- The scale of the task is ambitious, and the development of any index will have limitations towards the comprehensiveness of the data aggregated. The approach to mitigate negative reception of the results will lie in a fully documented methodological approach, reference reports and transparent commentary on the limitations, such as provided within this document. Furthermore, the data will be made available for exploration through an online dashboard to enable a self-discovery of the metrics and indicators.
- Finally, it is important that the Connecting Women in Digital as a whole, proactively engages key communities and stakeholders to achieve buy-in, collaboration and validation of insights and outcomes.

## Principal definitions

The scope of Connecting Women in Digital is to primarily address the gender divide in digital roles. It does, however, require a view of diversity and identity to appreciate the context and complexity of identity and the impact on data collection and analysis.

### Definition of Gender

*In the definition of gender, it is specified as the social attributes and opportunities associated with being male and female and the relationships between women and men and girls and boys, as well as the relations between women and those between men.<sup>8</sup>*

The common reference for gender is man and woman. In total the below characteristics and labels can be applied. In the scoping of this index and study, however, it is necessary to focus on those that are most objectively defined and reported, providing reliability in reported data, the size of the sample currently, and the relevance for the purpose of the Connecting Women in Digital mandate.

It is not a comment on any of the use or definition of any parameters, it is rather an assessment of the capacity for this pilot action to include data into an index that is feasible at the point of collection and assessment – i.e. provision of a robust and reliable index.

Key components considered for the aspect of gender are provided in the table below and assessed for their potential to be included in a data-centred index.

*Table 2 Definition of gender: Key components*

	OBJECTIVITY	RELIABILITY	RELEVANCE <sup>9</sup>
Woman	High	High	High
Man	High	High	High
Trans-woman	Medium	Low	Low
Trans-man	Medium	Low	Low
Non-binary or fluid	Low	Low	Low
Other	Low	Low	Low

Source: Adapted from GENDEX Study

Thus, while definitions of categories are clear and self-identification and assessment are feasible, it is unlikely that statistically relevant and actionable insights will be drawn from classification of founders or investors outside of the mainstream roles of woman and man. Where data is collected, the open option of other should be collected and included in datasets.

---

<sup>8</sup> UN Women: Office of the Special Adviser to the Secretary-General on Gender Issues and Advancement of Women (OSAGI)

<sup>9</sup> Where relevance is considered low, this is based on the proposed sample size and currently availability of reports and data to this effect.

While it is not foreseen that it will have significant impact on the index itself, it may contribute to other areas of study.

## Diversity beyond gender

There is a significant body of literature, particularly in a European context, which addresses diversity from defining a legal position of discrimination and the practices of national statistic bodies. From these arise a common set of standard principals to be taken into account for the definition of any index and selection of scope of study.<sup>10</sup>

- Principle 1: Do no harm
  - Disaggregation of the data into characteristics must have a purpose – just because we can doesn't mean we should.
  - Measure only if we can take action or effect change.
- Principle 2: An individual's identity cannot be reduced to a single characteristic; we are the sum of our experiences and backgrounds.
- Principle 3: Diversity is contextual: What is diverse in one location or field is common in others.
  - Each of the EU27 countries has its own definitions, interpretations, level of cultural openness, risks; certain concept may be biased towards anglophonic standards.
- Principle 4: Individuals should not be made to provide or accept labels. At its core, the concept of Diversity includes anyone who does not appear, act or sound like the majority.

## Components of diversity

In reviewing the definitions of diversity, with the goal of developing specific parameters, there are two frameworks and approaches that form the basis. The first, refers to the Loden Wheel of Diversity, which provides a layering of characteristics that make up one's identity with contributing internal and external dimensions, those that are intrinsic and that which are environmental or cultural.<sup>11</sup> This was further extended to incorporate the world of work which looked at roles and functions within an organisational structure.<sup>12</sup>

From the management of diversity and ensuring the equality of opportunities in education, the OECD provided a complementary framework that incorporated in a continuous spectrum both internal and external dimensions, that included capacity in terms of giftedness or special education needs, that are not relevant to the current purpose.

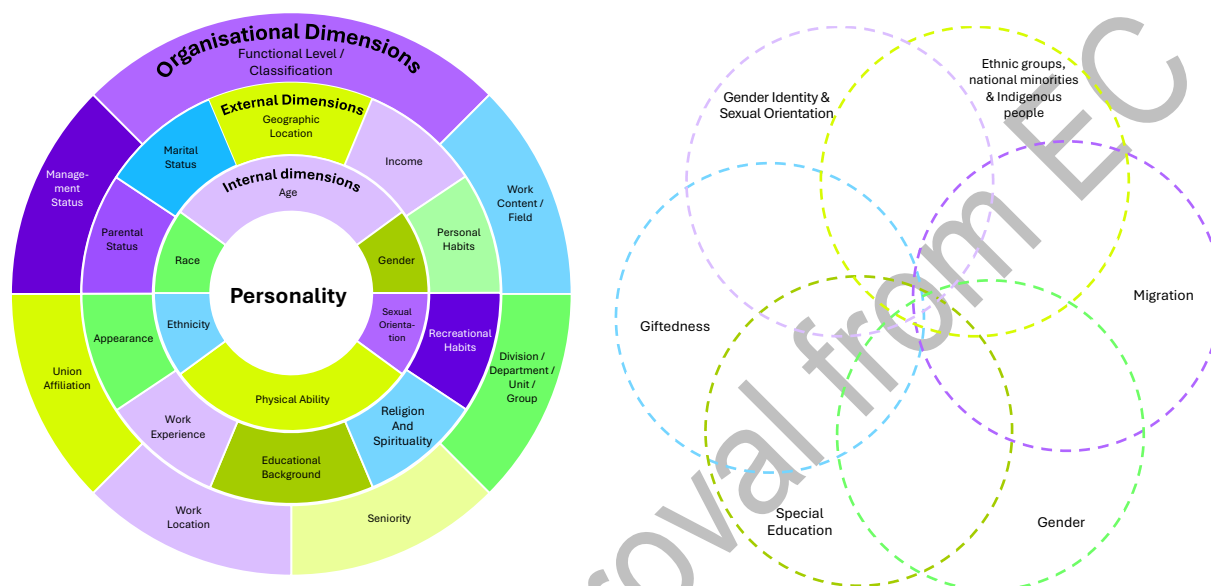
---

<sup>11</sup> Loden, M.; & Rosener, J. B: Workforce America: Managing Employee Diversity as a Vital Resource, 1991

<sup>12</sup> Gardenswartz, L.; & Rowe A: Diverse Teams at Work: Capitalizing on the Power of Diversity. Capitalizing on the Power of Diversity. Alexandria, VA: Society for Human Resource Management, 2003



*Figure 3 Extended Loden Wheel of Diversity and OECD Education model for inclusive education*



Source: Adapted from Brunner et al (2021) and Cerna et al (2016)

Contributing further to these frameworks, which attempt to provide a comprehensive view of diversity, was the open question on definition of diversity from directors of FTSE 350 companies.<sup>13</sup>

The resulting characteristics or parameters were, in order of references made:

- Personal/Neuro/Personality
- Gender
- Race/Ethnicity
- Other
- Functional
- Age/Experience
- Nationality/Geographic
- Sexual Orientation

<sup>13</sup> Board Diversity and Effectiveness in FTSE 350 Companies (2021) The Financial Reporting Council Limited

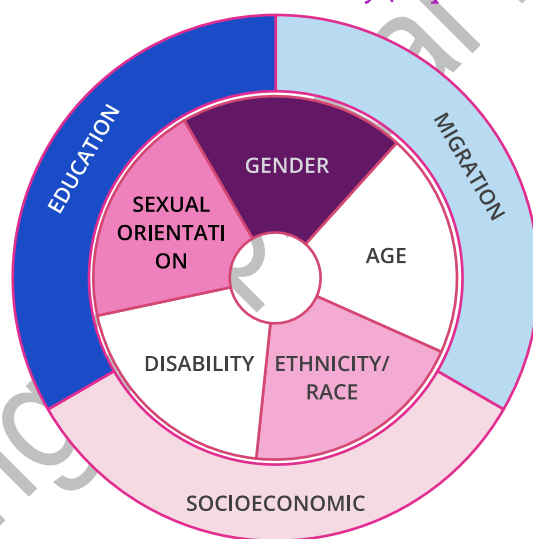


Below is captured a holistic model for the EU context, which has been adapted from the GENDEX study<sup>14</sup> and encapsulates the core characteristics which compose one's identity.

It is composed of two key layers:

- Internal components – immutable characteristics that remain fixed over time<sup>15</sup> or cannot be readily changed these include:
  - Gender.
  - Ethnicity and Race.
  - Age.
  - Sexual orientation.
  - Disability.
- External components – those characteristics that are related to the past or current experiences and environments of the individual which includes:
  - Educational background.
  - Migration status.
  - Socioeconomic status.

*Figure 4 Proposed model for definition of diversity (adapted from GENDEX study)*



Source: Adapted from GENDEX Study

Below, and further in Annex 1, the additional individual components are described in detail, providing existing nuance. In the review of the parameters, we cannot as a

<sup>14</sup> Rowan, B. (2024). GENDEX Review of the State of the Art. GENDEX Consortium. <https://doi.org/10.5281/zenodo.12529560>

<sup>15</sup> On the use of the term immutable – it is understood by the WIDCON consortium that aspects like gender, ethnicity and race are social constructs and may be externally presented in one form or another and can be fluid in practice but are stable core aspects of an individual's identity. Age is also central to an identity that cannot be changed readily by the individual.

consortium, make a conclusive statement on all fields and aspects of diversity but attempt to provide an overview.

### A NOTE ON INTERSECTIONALITY AND ADDITIONALITY

It is important to recognise that individuals cannot be reduced to a single identity marker, as their identities are multifaceted and interconnected. The concept of intersectionality provides a critical framework for understanding how various aspects of an individual's identity such as sex, gender, race, ethnicity, age, class, and other social categories interact and shape unique experiences of both discrimination and privilege. This perspective emphasises that one's position in multiple social categories (e.g., gender, race, ability) can significantly influence the extent to which they encounter discrimination or benefit from privilege.

The European Institute for Gender Equality<sup>16</sup>, alongside feminist scholars,<sup>17,18</sup> defines intersectionality as an analytical tool that explores how sex and gender intersect with other social characteristics, contributing to complex and varying experiences of inequality. As Collins<sup>19</sup> notes, “race, class, gender, sexuality, ethnicity, nation, ability, and age operate not as unitary, mutually exclusive entities, but as reciprocally constructing phenomena”. This highlights the need to recognise that social identities are interdependent, and understanding inequalities requires an awareness of how these factors intersect.

By acknowledging these intersections, we gain a more comprehensive understanding of the mechanisms that perpetuate inequality, including both direct and indirect discrimination, stereotypes, prejudices, and historical patterns of marginalisation. For instance, in the digital sector, different social categories can significantly shape women's experiences. Consider the differing career paths of two women in Paris: one a European, middle-class MBA graduate, and the other a non-European, working-class MBA graduate. Despite similar educational backgrounds, their opportunities and challenges may diverge significantly due to their intersecting identities, illustrating the importance of considering intersectionality when analysing diversity and inequality.

Further to this is the definition of additionality<sup>20</sup>, where a cumulative positive or negative effect takes place based on several grounds operating separately. This is to say, there may not

---

<sup>16</sup> European Institute for Gender Equality (EIGE) (n.d.) Gender mainstreaming.

<sup>17</sup> Carastathis, A the concept of intersectionality in feminist theory., 2014

<sup>18</sup> Crenshaw, K. 'Mapping the margins: Intersectionality, identity politics, and violence against women of color', Stanford Law Review, 43(6), pp. 1241–1299., 1991

<sup>19</sup> Collins Patricia Hill. 2015. Intersectionality's definitional dilemmas. Annual Review of Sociology 41: 1–20.

<sup>20</sup> Mehrdimensionale Diskriminierung – Begriffe, Theorien und juristische Analyse (2010) Antidiskriminierungsstelle des Bundes

be an observable interdependency like in intersectionality, but rather various advantages and disadvantages can be experienced.

Additionality refers to the cumulative effects of separate identity characteristics, where advantages or disadvantages accumulate independently across different identity markers. Unlike **intersectionality**, in this framework, there is no observable interdependency; instead, each axis of differentiation impacts the outcome separately, leading to distinct experiences of privilege or marginalisation. For example, being a woman (gender) and being older (age) might each result in distinct disadvantages in the workplace—such as lower wages or fewer opportunities—but these disadvantages do not necessarily intersect or compound in the same way intersectionality suggests. Instead, the impact of each characteristic is experienced independently, leading to separate but significant effects on the individual's experiences of privilege or marginalisation.

For the purpose of WIDCON, these informal definitions are provided here to permit reflection in any interpretations made upon the resulting index. They do not, however, influence the definition of the component metrics or structure of the index. While practical constraints, such as the lack of comprehensive disaggregated data and resource limitations, prevent the direct integration of intersectionality into the quantitative data collection, we may seek opportunities to incorporate an intersectional perspective where feasible. This could potentially be achieved through cross-analysis of the data when possible. Additionally, intersectionality may be considered in the discussions and activities of the thematic working groups, where applicable, to reflect on its relevance within the context of the WiD Index.

## Agreed WIDCON approaches

In the context of the relevance intersectionality, the following is agreed:

- Intersectionality is an important dynamic that defines the experiences and choices of an individual – one is not just a woman. Women and girls are not a homogeneous group; factors such as age, disability, sexual orientation, ethnicity, migration status, and socioeconomic background influence their access to power, resources, and opportunities in the digital/ICT sector.
- Other key inequalities related to social categories such as socioeconomic background are strongly contributing factors to the lack of diversity in the digital sector, particularly in access and participation from lower socioeconomic groups. However, the limited resources of WIDCON and given that sex is the primary variable used in the existing datasets dictate that we must presume that in terms of participation in STEM and ICT, all other inequalities remain fixed.
- While WIDCON cannot perform intersectional analyses directly due to limitations in available data and resources, it will seek opportunities to gather relevant data where feasible. This will be achieved through cross-analysis of the data when feasible and by exploring intersectionality as a key theme in the qualitative insights

gathered through Women in Digital Forum activities. This approach ensures that WIDCON reflects the complexity of diversity in the digital sector within the constraints of available data

On the handling of statistical and external datasets and gender:

- WIDCON recognises that in the vast majority of cases, statistical data captures assigned sex rather than declared gender. While, for most people, their assigned sex at birth aligns with their gender identity, this is not always the case. Automatically linking the sex variable in a dataset to gender identity can lead to misclassification errors, potentially misrepresenting individuals who are transgender, non-binary, or gender nonconforming. WIDCON assumes that the scale of misgendering contained within is statistically small and will look to third party studies to estimate this more accurately.
- When presenting data collected from sources using sex as a variable of sex, WIDCON will include a brief explanatory note to acknowledge the use of sex instead of gender due to limitations of data sources.
- WIDCON will, when collecting data, provide the following options to ensure it receives accurate representation: Man, Woman, Other (Open Field), Prefer not to answer. It understands that were an individual is asked to categorise themselves; this reflects their gender. WIDCON will take privacy and ethical considerations into account to ensure that individuals feel comfortable providing their gender identity, and that all data is collected, stored, and reported responsibly.

On the application of data collection methods and assignment of gender:

- Similar to above, when WIDCON seeks to perform analysis on other datasets to attempt to assign gender, it will deploy tools such as libraries of names and their gender that will inherit biases but that these biases are within statistically acceptable limits for the purpose of WIDCON.

## Gender mainstreaming

The integration of a gender perspective into the methodology used for the WIDCON collection and analysis of both primary and secondary data is essential for ensuring that the WIDCON Index and Scoreboard accurately reflect women's participation in digital education, employment, and leadership. This approach goes beyond simply reporting numerical disparities, focusing on the structural and systemic factors that influence career trajectories, access to opportunities, and barriers to advancement. Without it, gender inequalities could be underrepresented or misinterpreted, limiting the effectiveness of policy recommendations and intervention strategies.

The methodology for developing the Women in Digital Index and Scoreboard is designed to incorporate gender considerations at every stage of the research process, from formulating research questions to interpreting results. This ensures that key indicators not only capture

participation rates but also the qualitative factors influencing women's entry and engagement in digital professions such as motivations for selecting ICT subjects in higher education, attitudes towards ICT traineeships or impact of women role models. The data collected will reflect both statistical representation and the broader socio-economic conditions shaping gendered career outcomes.

Gender-sensitive survey design is crucial for gathering meaningful insights into women's experiences in digital careers. The Women in Digital survey will prioritise gender-responsive language and question phrasing to avoid reinforcing biases. The sampling framework will ensure the survey reaches a diverse cross-section of women in digital professions, accounting for variations in educational background, career level, and geographic location.

## Key research questions

### Purpose of research questions

In order to frame and focus the activities of the Connecting Women in Digital team within this far reaching and intertwined challenge, we aim to address a set of guiding questions that will seek to provide the strongest outcomes in terms of concrete positions and actions to effect real change. These are categorised across the four key phases of the pipeline.

#### STEM Phase

Nº	Question	Qual/Quant	Data Source
1.1	Are girls less likely to participate in STEM subjects?	Quant	Statistics
1.2	Is there a notable change between key stages; primary, secondary, upper secondary	Quant	Statistics
1.3	What factors affect subject choices by girls in upper secondary?	Qual	Forum

#### ICT Phase

Nº	Question	Qual/Quant	Data Source
2.1	What is the proportion of women participating in ICT subjects in higher education?	Quant	Statistics
2.2	What motivates women to select ICT subjects in higher education?	Qual	Forum
2.3	Do women participate to ICT traineeships or apprenticeships at same levels in higher education?	Quant	Statistics
2.4	What is the impact of women teachers/lecturers in achieving more women learners?	Quant/Qual	Forum
2.5	What specific career guidance programs are available to women pursuing ICT education?	Quant/Qual	Survey/Forum

#### DIGITAL Phase

Nº	Question	Qual/Quant	Data Source
3.1	What is the retention rate of women in technical digital roles throughout their career?	Quant	Survey
3.2	What factors lead women to take on non-technical roles?	Qual	Forum
3.3	Are women in digital roles more likely to participate in upskilling?	Quant	Survey
3.4	Are women without ICT education or training less likely to enter digital roles compared to men?	Quant	Survey
3.5	Are there differences in the types of projects or tasks assigned to women?	Qual	Forum
3.6	Are there other occupations that were previously male dominated in the past that we can learn from?	Quant	Statistics
3.7	Are women less likely to apply for internal promotions at early stages in their career?	Qual	Forum

#### LEADERSHIP Phase

Nº	Question	Qual/Quant	Data Source
4.1	What is the proportion of women leaders in digital businesses or business units?	Quant	Annual Reports
4.2	Do more women in leadership positions increase the participation of women in digital roles?	Quant	Annual Reports

## Index structure

### Overview

The Women in Digital Index aims to develop a robust and replicable framework for monitoring the conversion rates of female students and professionals throughout their education and career paths, and to highlight where the major gaps are in retaining women's participation in digital and ICT, from the STEM phase to the leadership phase, through various indicators (as detailed in the following paragraph):

- **STEM Phase:** *Girls' participation and achievement in STEM education and subjects in secondary education, including vocational education and training. This stage measures early engagement, growth of interest and achievement as a foundation for later ICT and digital higher education and careers. Note: Due to a low availability of quantitative data, STEM performance in primary education will be covered through qualitative insights via the Forum and dedicated workshops.*
- **ICT Phase:** Participation of women in ICT and related subjects in tertiary, initial vocational training (I-VET), postgraduate and doctoral education. These indicators assess the availability of highly qualified women candidates for specialised roles in ICT and digital.
- **Digital Phase:** Women's participation in ICT and digital roles, focusing on employment rates, classification by role (technical vs. non-technical), gender pay gap and contractual conditions during early to mid-career. IP ownership and ICT entrepreneurship are also covered as proxies for women's innovation in ICT and digital. This phase highlights whether women are progressing beyond entry-level positions, gaining the necessary experience and training opportunities, potential barriers to progression and finally the retention rate of women in digital.
- **Leadership Phase:** Women's participation in ICT and digital leadership, both in technical leadership roles and as top executives or board members in ICT companies. Indicators at this stage are crucial for understanding the extent to which women have broken through the “glass ceiling” and achieved leadership parity, influencing strategic decision-making and innovation.
- **Enabling Environment:** This underlying additional phase outside the ICT and digital pipeline includes metrics linked to broader societal and structural factors that influence women's participation and success beyond - but also directly influence - ICT and digital roles by country, from basic digital skills to overall women participation in the workforce.

The Women in Digital index focuses on monitoring the path of women from primary education to industry participation. Two complementary pathways, though not directly measured by this index, are crucial for context and must be monitored and referenced throughout its development.:



- **Academia and research:** as monitored for the EU27 by SheFigures. This publication serves as a key source of comparable, Europe-wide statistics on gender equality in research and innovation (R&I). The 2024 addition added a pilot "She Figures Index" that allows for a more comprehensive tracking of gender equality progress across the EU.<sup>21</sup>
- **Technology entrepreneurship:** as detailed by the pilot Gender & Diversity index within GENDEX for the EU27 and UK. The goal is to provide data and insights to improve gender and diversity within the European innovation ecosystem, particularly for startups, investors, and policymakers. The project scope is gender participation in deep tech, which can serve as a proxy for general ICT.

## Quantitative data

The tables below list the different indicators for each identified phase of the Women in Digital pipeline, detailing the name of the indicator, a synthetic definition of the indicator, the main primary and secondary sources for these indicators, and any associated limitations (e.g. sampling, frequency of data collection).

### STEM PHASE

*Table 3 Quantitative metrics to be included in the STEM phase*

Metric	Stage	Definition	Data source	Limitations
ST1. Girls' performance in STEM subjects in primary education.	Primary education	Scale and level of competences of girls in STEM subjects (Grade 4)	TIMSS (IEA)	Four years sampling frequency. Sample is consistent but may not represent total population. Lag from date of data collection to access of around 36 months.
ST2. Girls' performance in STEM subjects in secondary education.	Secondary Education	Scale and level of competences of girls in STEM subjects (Secondary education)	OECD's Programme for International Student Assessment (PISA)	Five years sampling frequency. Sample is consistent but focuses on 15-year-old students only, which may not adequately represent total population.
ST3. Girls in upper secondary STEM IVET (ISCED 3) programmes	I-VET	Participation of girls in upper secondary vocational STEM programme education (ISCED 3).	Eurostat: educ_uoe_ent02	Lag from date of data collection to access of around 36 months.

## ICT PHASE

*Table 4 Quantitative metrics to be included in the ICT phase*

Metric	Stage	Definition	Data source	Limitations
IT1. Women ICT graduates (ISCED Level 6)	Higher Education	Proportion of women ICT graduates	Eurostat: educ_uoe_grad10	Lag from date of data collection to access of around 36 months
IT2. Women enrolled in ICT-related vocational trainings (ISCED Level 4-5)	I-VET	Proportion of women with an ICT related post-secondary non-tertiary vocational education and/or short-cycle tertiary education, vocational/professional	Eurostat: educ_uoe_grad02	Lag from date of data collection to access of around 36 months
IT3. Women with an ICT master of science (ISCED Level 7)	Higher Education	Proportion of women with a master degree in ICT	Eurostat: isoc_sks_itsps	Lag from date of data collection to access of around 24 months
IT4. Women with a PhD in ICT-related fields (ISCED Level 8)	Higher Education	Proportion of women among ICT Doctoral graduates	Eurostat: educ_uoe_grad02	Lag from date of data collection to access of around 36 months

## DIGITAL PHASE

*Table 5 Quantitative metrics to be included in the DIGITAL phase*

Metric	Stage	Definition	Data source	Limitations
DG1. Employed women with an ICT education	Early-to-middle career	Number of employed women with an ICT education	Eurostat: isoc_ski_itsex	Lag from date of data collection to access of around 24 months
DG2. Women ICT specialists	Early-to-middle career	Proportion of women ICT specialists	Eurostat: isoc_sks_itsps	Lag from date of data collection to access of around 24 months.
DG3. Women in specific ICT roles	Early-to-middle career	Proportion of women in specific technical ICT roles	Women in Digital proprietary survey	Sample is consistent to show a trend but may not represent total population.
DG4. Women with an ICT background in non-technical roles	Early-to-middle career	Proportion of women with an ICT background in non-technical roles	Women in Digital proprietary survey	Sample is consistent to show a trend but may not represent total population.

DG5. Gender pay gap in ICT	Early-to-middle career	Percentage difference between average gross hourly earnings of men and women ICT professionals, expressed as a percentage of men's earnings.	Eurostat: earn_gr_gpgr2	Lag from date of data collection to access of around 24 months.
DG6. Contractual conditions of women in ICT roles	Early-to-middle career	Proportion of women in different contractual conditions of women in ICTs (e.g. FTES, Part time, flexible and remote workers)	Women in Digital proprietary survey	Sample is consistent to show a trend but may not represent total population.
DG7. Women in ICT continuous vocational trainings (CVET)	Early-to-middle career	Proportion of women enrolled in ICT continuous vocational trainings (CVT) (e.g. upskilling and reskilling programmes)	Women in Digital proprietary survey	Sample is consistent to show a trend but may not represent total population.

## **LEADERSHIP PHASE**

*Table 6 Quantitative metrics to be included in the LEADERSHIP phase*

Metric	Stage	Definition	Data source	Limitations
LD1. Women in technical leadership roles	Leadership	Proportion of women with Director level and above in ICT roles	Estimations on secondary database (IDC's job role forecasting) and Women in Digital proprietary survey questionnaire.	Estimation of gender split is applied to the core IDC dataset with a lag of 24 months.
LD2. Women CEOs in top technology companies	Leadership	Proportion of women with CEO level in ICT companies	Estimations on secondary database (IDC's job role forecasting) and Women in Digital proprietary survey questionnaire.	Estimation of gender split is applied to the core IDC dataset with a lag of 24 months.
LD3. Women on the Board of Directors of top	Leadership	Proportion of women on the Board of Directors of top European	Desk research, PitchBook	Total sample of companies considered could be not representative

technology companies		technology companies		of the entire population.
LD4. Time taken for women to reach leadership positions	Leadership	Average duration (in years) required for women to advance from ICT entry-level positions to ICT leadership roles	Women in Digital proprietary survey	Sample is consistent to show a trend but may not represent total population.
LD5 Retention rate of women in ICT executive positions	Leadership	Proportion of women who remain in ICT executive-level roles over a specified period (e.g., annually, bi-annually)	Women in Digital proprietary survey	Sample is consistent to show a trend but may not represent total population.

## ***ENABLING ENVIRONMENT***

***Table 7 Quantitative metrics to be included regarding women in the labour force***

<b>Metric</b>	<b>Definition</b>	<b>Data source</b>	<b>Limitations</b>
EE1. Women with basic or above basic overall digital skills	Proportion of women with basic or above basic overall digital skills over total population	Eurostat: eq_dskl07	Lag from date of data collection to access of around 24 months
EE2. Women currently unemployed	Proportion of women between 18-55 currently not employed	Eurostat: lfsi_emp_a	
EE3. Women with children employment rate	Proportion of women in children employed in households	Eurostat: lfst_hheredty	Lag from date of data collection to access of around 24 months
EE4. Contractual conditions of women	Proportion of women with full-time and part-time employment	Eurostat: lfsa_epgais	Lag from date of data collection to access of around 24 months

## Member State initiatives and mapping

### *Type of actions*

While analysing women in digital, it is important to monitor the gender distribution in different ICT roles and existing challenges in achieving and/or following it. At the same time, the investigation of reasons behind the numbers is even more important as it allows for estimating the efficiency of some initiatives and generating recommendations for possible changes by initiating new initiatives. Individual reasons apply to each person, however estimating a country's status quo, the systematic approach to equal rights, and promotion of diversity in digital areas are highly relevant, illustrating the country's dedication and support to the topic. A summarised list of initiatives in each country not only indicates a country's active role in increasing women's interest and participation in the digital sector but can also serve as a knowledge resource for further recommendations, actions to join by individual person.

Seeking to provide a wider view of activities, the search and summary of initiatives will be organised into types, based on different categories:

- By scope. Phase and stage. As defined in the Women in Digital Pipeline, the target audience can be divided into four groups and respective sub-groups. Based on the pipeline, the same target groups of the initiatives will be analysed:
- STEM phase. All initiatives, dealing with girls STEM education will be associated with this phase.
- ICT phase. All initiatives, which are oriented on ICT education will be marked as ICT phase.
- Digital phase. All initiatives, which are oriented on technical ICT roles, reskilling for it will be marked as matching digital phase.
- Leadership phase. All initiatives, which are oriented on leadership skills in ICT, support achieving leading positions in digital areas are relevant to the leadership phase.
- By funding source. The funding source of initiatives can help to highlight the country's maturity level on the topic question and readiness to take their own responsibility for properly dealing with the topic. The main three main types of initiative funding are estimated:
  - European Commission funding. Member State representatives' participation in EC-funded projects, which include women in digital questions, indicate a country's active position and willingness to contribute to this challenge. At the same time, it will highlight the interest of the European Commission on the topic and the dedication to help bridge the gender gap.
  - National funding. National funding highlights a country's dedication to the women in digital topic. Both this topic-oriented call for projects as well as the

project itself, which additionally incorporates gender balance in the digital field indicate an increased interest in the field.

- Other funding. The existence of local initiatives or joining some global initiatives might be a very good indicator illustrating the Member Country's maturity and social responsibility to act in the direction of promoting women in digital.

The collected set of initiatives should allow the understanding of activities related to the women in digital topic in different countries and phases. Linking the data with the ratio of women participation in digital could help to understand the efficiency of existing initiatives and bridge the path to finding the most effective initiatives.

### *Selection criteria*

The selection of initiatives for analysis will be estimated by several factors – data availability, country involvement and initiative relation to women in digital area.

The data sources with clearly identified initiatives belonging to women in digital relation will be prioritised as it went through the formal justification processes. However, the number of such data sources is limited, therefore data sources, which have a textual description of the initiatives will be included for the consideration.

Country involvement will be estimated by identifying at least one country from EU-27 or the analysed countries for comparison. At the same time initiatives, which mention global, European scope, not oriented to some specific countries will be included.

The initiative's relevance to women in digital will be estimated by analysing the initiative description text and identifying main keywords, mentioning women, gender diversity in digital, ICT, etc.

### *Produced index*

A key scope of the Women in Digital index is to connect with and measure the effectiveness of the different Member States initiatives collected, comparing the number of initiatives and level of funding with quantifiable metrics, such as the increase the number of women ICT graduates by country, along a desirable time span of ten years. This will constitute the foundational baseline to identify which best practices are the most efficient when addressing the ICT and digital gender inequality in the European Union, by also including the benchmarked countries previously listed outside the EU for comparison purposes (Brazil, USA, India, UK).

Figure 5 Tiered vision of the Index comparing Member State Initiatives and quantitative data – Mock-up.  
Source: WiD

	STEM Phase	ICT Phase	Digital Phase	Leadership Phase	Enabling Environment	Overall Positioning
Country 1	LOW	MEDIUM	MEDIUM	LOW	LOW	LOW – Improve ICT education and leadership appeal
Country 2	HIGH	LOW	MEDIUM	HIGH	MEDIUM	MEDIUM – Increase initiatives to attract girls in ICT
Country 3	HIGH	HIGH	MEDIUM	HIGH	HIGH	HIGH – Strong initiatives at various conversion points
...	MEDIUM	LOW	HIGH	MEDIUM	MEDIUM	MEDIUM – Improve ICT education and leadership appeal

● LOW – Small number of initiatives, lack of funding and low impact and reach  
● MEDIUM – Limited number of initiatives but some measurable impact and funding  
● HIGH – Strong initiatives and measurable results at various conversion points

Source: Connecting Women in Digital

As previously noted, within the current proposed approach to developing the Index a specific weighting is not foreseen for the first edition, however, this may be under review after data collection and analysis is completed.

The Women in Digital Index will also be published both as an interactive data dashboard and as an API or similar for connection with ongoing Data Spaces and other related initiatives. It would allow filtering and the generation of individual reports by the user. We will use a hierarchical visualisation, where the EU index can be expanded based on different factors, but to different stages and linked to existing initiatives roadmap, if available and feasible. The roadmap itself will have an interactive survey, filtering and results could be visualised not only as a list, but also as a graph, indicating further steps/initiatives, similar initiatives, and best practices and best performers.



## Data collection methodologies and tools

### Desk research

The Women in Digital Index will rely on robust and reliable publicly available secondary sources as one of the key information bases for its development, allowing data collection of longitudinal data for a desirable ten-year timeframe for the STEM and ICT metrics, and four years for the Digital and Leadership metrics, allowing for proper comparison and benchmarking with the different Member State initiatives collected. Pending possible data gaps, a longer timeframe will allow the impact of policies and initiatives to promote gender equality in the digital sector to be assessed, will allow meaningful comparisons to be made between different countries or regions over time, and will facilitate the identification of best practices and areas for improvement.

### *Identified sources*

Throughout the first assessment of potential sources identified, we list:

- **OECD:** Providing consistent data on STEM education and skills development, such as the PISA studies. The OECD utilises standardised methodologies and definitions, facilitating cross-country comparisons and benchmarking.
- **Eurostat:** Providing relevant dataset related to educational attainment and labour market statistics and gender equality specific datasets. Eurostat also maintains longitudinal data and also employs standardised methodologies and definitions, ensuring data comparability across countries and over time.
- **TIMSS (Trends in International Mathematics and Science Study, IEA):** Providing an international assessment framework and data sets designed to evaluate the mathematics and science knowledge of students in the fourth and eighth grades internationally. It also employs standardised methodologies and definitions, ensuring data comparability across countries and over time to monitor educational achievements.
- **Specialised company databases such as PitchBook:** Covering detailed data on entrepreneurship and investment levels by gender, allowing for the analysis of trends of women leadership and funding for women-led companies.
- **Other existing research from other skills, talent and gender equality related projects such as GENDEX:** Leveraging cross-indicators such as women ownership of IPs to highlight women's contribution to innovation.
- **Proprietary datasets such as the IDC Job Role Forecasting:** IDC has its own employment benchmark tracker that, although discontinued, has been updated until 2023 and contains a forecasting up to 2032. The data provides quantitative inputs on full time employment for over 130 countries, including all EU Member States but is not split by gender. That said, data from the employment forecast tracker can be used as a more detailed benchmark to further refine other publicly available statistics.



Other sources, such as corporate sustainability reports, can also be considered for statistics on the gender composition of the workforce by role and/or seniority. The viability of additional sources such as professional social media platforms (e.g. LinkedIn), online job portals or open-source repositories will also be explored (additional information to be provided in section 6.4 Exploratory approaches).

### *Quality assurance*

To ensure that the Index relies only on robust and reliable secondary data sources, Women in Digital will perform a quality check on the accuracy, reliability, and validity of the secondary data used. This process will include:

- **Data triangulation:** Prioritising, when possible, data comparison from multiple sources to identify any discrepancies or inconsistencies.
- **Data Validation:** Cross-referencing data with other relevant sources or studies- such as GENDEX or SheFigures - to verify its accuracy where feasible, checking for any outliers or anomalies that may indicate errors in the data and, where possible, perform sense checks on the data.
- **Documentation and Transparency:** Maintaining detailed records of all data sources, methodologies, and validation procedures and clearly documenting any limitations or assumptions made during the data collection process.

### *Women in Digital Survey*

The Women in ICT career survey will be one of the key components of the Women in Digital methodology and is being designed to complement the secondary data collection. The survey results are expected to (i) be published in a separate deliverable (Women in ICT Career Survey report findings), (ii) contribute to the Index Development phase of the project as results will be cross-comparable in different countries and (iii) serve as an indication of certain local policies identified during the monitoring of initiatives phase

The survey will be rolled out by IDC CEMA and designed to capture detailed responses from women who are currently part of the digital workforce in Europe. WIDCON will make use of the following survey methodologies: CAWI (Computer-Assisted Web Interviewing) and CATI (Computer-Assisted Telephone Interviewing).

- **CAWI** provides high-quality, targeted responses to be obtained via online surveys. IDC has access through its online panel providers to a wide variety of respondents in many countries; the respondents are profiled for their professional background, their seniority and line of business, knowledge and their decision-making involvement.
- **CATI** is widely used in business (B2B) research when surveying ICT and non-ICT (business) decision makers and is best suited to structured questionnaires, where most of the possible answers are known from previous research and can be listed as pre-

coded responses. CATI allows for targeting high quality respondents and elicits better response rates for certain audiences (e.g., more senior respondents, niche roles, etc.)

While secondary data sources may provide a robust understanding of the current statistics of the gap for women in digital, the survey will help the consortium complement these metrics and deep dive into the more subjective barriers that women face throughout their education and career paths. While we will be focusing on women who are currently holding ICT/digital roles, questions will be added to capture their perception regarding educational paths prior to joining the workforce, as well as perceived enablers that allowed them to foster their interest in the area and further develop their careers.

Also, the survey will contain a set of screening questions, including (but not limited to) age, level of seniority, tenure with the company, industry area and level of education, which will enable WIDCON to create data cuts and drill down into specific barriers and enablers for different career stages. Furthermore, the sample will also allow us to correlate past educational choices to current career development to identify areas with the most potential.

### *Sample*

As the survey results will also feed into the WIDCON index development, the consortium will aim for a more granular country view and expect to have at least 20-50 respondents by Member State, provided IDC CEMA can cater for the whole country scope. That said, we estimate the total sample for the survey to be between 700-1400 respondents if CAWI is used as the main methodological approach. Country sampling will be conducted by IDC CEMA and will rely on economic and demographic factors such as population and size of industry to establish the correct number of respondents by country.

*Table 8 WiD survey's sampling approach*

<b>Project name:</b>	WIDCON ICT Career Survey
<b>Total sample size:</b>	N=500-1200 respondents
<b>Data Collection Method</b>	CAWI
<b>Quant. / Qual.</b>	Quantitative
<b>Company size:</b>	100+ employees
<b>Industry segments:</b>	all verticals
<b>Target respondent:</b>	Senior management or IT management/specialists (Examples: CIO, CTO, VP, director IS/IT-related), IS/IT middle management, IS/IT operations, IS/IT staff, senior business management (e.g. VP, director of technology related areas), non-technical ICT related roles such as IT Project Manager, Product Owner, Scrum Masters
<b>Screening Criteria:</b>	Women working in ICT/digital roles
<b>Weighted:</b>	No

## *Sections/segments*

The questionnaire was structured to gather comprehensive insights into various aspects of women's involvement in the ICT sector. It begins by inquiring about the respondent's current role and educational background, including any ICT-specific education they may have pursued. The survey also explores motivations for entering the ICT field and barriers faced during education, such as lack of role models or financial constraints. Additionally, it assesses the respondent's experience with ICT education and the types of support they believe would have been beneficial. This section provides a foundation for understanding the foundational elements that influence women's participation in ICT.

Furthermore, the survey looks to identify employment status, career longevity, and professional development of women within the ICT sector, as well as investigating challenges faced by women in their current respective roles, such as work-life balance, skill development, and whether they perceive to have equal opportunities for advancement compared to male colleagues. The survey also touches on gender discrimination experiences and perceptions of job satisfaction and work arrangements. Moreover, it explores career transitions, retention efforts, and organizational initiatives supporting women in ICT. Finally, it seeks opinions on barriers to entry for women and resources needed for career advancement, providing a holistic view of women's experiences and perspectives in the ICT workforce

## *Member State initiatives*

### *Identified sources*

The main data sources are identified in this stage, While applying the same methodology, the list could be extended in the future. The initial list is the following:

- Recovery and Resilience Facility projects have a clear assignment to gender equality. RRF Scoreboard ([https://dashboard.tech.ec.europa.eu/qs\\_digit\\_dashboard\\_mt/public/hub/stream/1d5f8b8b-6b97-4642-914a-4de9d040a245](https://dashboard.tech.ec.europa.eu/qs_digit_dashboard_mt/public/hub/stream/1d5f8b8b-6b97-4642-914a-4de9d040a245)) already have all data about the projects. The data could be structured and extracted for further analysis and is the most trusted source.
- Kohesio system (<https://kohesio.ec.europa.eu/en>) has all information about projects supported by EU Cohesion policy. The provided data and description of the projects will be used to identify women in digital related projects.
- EU Funding & Tender Portal (<https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/projects-results>) publish main information about projects & results. All the information, after relation to women in digital topic matching will be analysed.

- Global internet (via search engines and social networks) will be used to find some local or non-EU funded initiatives. These sources are not structures, therefore the information, presented in the webpage or post will be analysed to extract its relevance to women in digital topic.

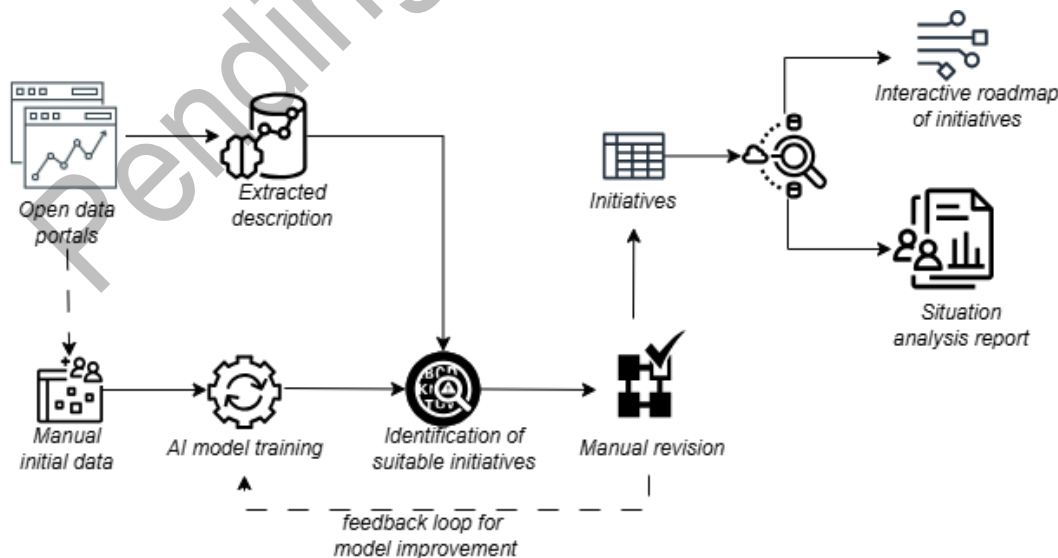
This is the initial list of data sources to identify women in digital initiatives. Taking into account most of it is not structured and linked to the gender diversity and ICT, digital area, certain level of automation is needed.

### *Automated collection approach*

The need of automation is caused by two main reasons: reduction of resource on text data analysis; consistency in data analysis and possible reuse for future data collection. Taking into account the available data is not properly structures for computerised analysis (just few a sources have well-structured data, needed for the analysis), the text interpretation is needed. Considering the possibilities of existing artificial intelligence (AI), machine learning (ML), natural language processing (NLP) solutions, these technologies will be utilised to automate the data extraction and analysis.

The initial schema for the gathering of initiatives of women in digital is presented in figure below. To train the ML model, the initial list of activities is gathered from the defined data sources manually. It provides the URL to the activity and structured, labelled data. This initial data is used to train ML models for estimation whether the data is about women in digital related initiative or not, what is the target audience and geographical scope. The identified initiatives are presented in a list for manual human revision. This stem is needed to increase the accuracy of the data processing as the revised results are used both for the list of initiatives, as well as ML model tuning to achieve a higher accuracy and possible concept drift over time.

*Figure 6 Proposed model for automated initiative data extraction*



Source: Connecting Women in Digital

The gathered data automatically will be processed to visualise for further situation analysis and recommendation system to search for relevant initiative. The possibility to integrate external party initiative proposal for analysis could be done in the future.

### *Categorisation and labelling*

As stated in 5.3.1 section, all data sources will be analysed to identify it is relevant to the women in digital topic. For those which are, the matching to each of the four (STEM, ICT, Digital, Leadership) principal phases and respective stages of the Women in Digital Pipeline. One initiative could cover more than one stage, therefore classifiers for each of them will be developed.

The geographical coverage of the initiative will be estimated by identifying listed countries or geographical areas, to associate the initiatives to the appropriate country.

The funding source mostly will be estimated automatically by the data sources. Meanwhile for the initiatives, found on the internet, the funding source as well will be identified based on the initiative description.

### *Approximations and limitations*

In some cases, the initiative description might not include any data, which could indicate the target group, geographical scope or funding source. Therefore, each of the classifiers will be adapted to provide “No data” option, rather than replaying on inaccurate assigning of any category.

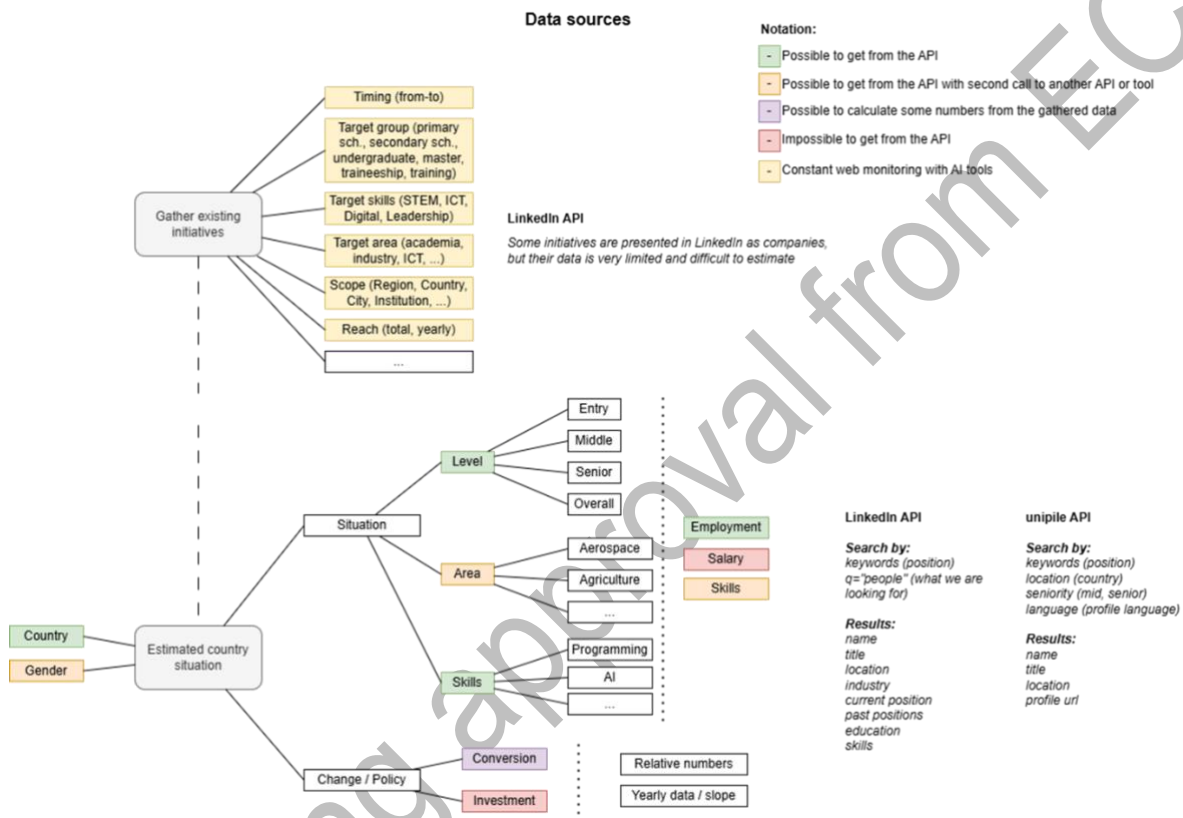
The text processing methods are advanced and allow text analysis automation. Generative artificial intelligence (GenAI) solutions will not be used for the data processing, only machine learning classification and NLP models, for more accurate text analysis. However, it will be oriented on English language mainly. In case of missing data for some countries, the option to integrate multilanguage models or automated text translation to be suitable for English-based text analysis, will be tested.

To build very accurate text analysis model, a large number of training data is needed. To cope with this issue, the manual data revision is included. Taking into account the model accuracy and training data, the initial threshold values for the classifiers will be adjusted to be oriented on having higher false positive rate, rather than false negative. Integrating the manually revised data into the model tuning, retraining process, the threshold values later could be readjusted to reduce the need for manual revision.

## Exploratory approaches

Beyond the research approaches explained above, WIDCON will also explore alternative methods to collect publicly available data, such as (but not limited to) the usage of LinkedIn and Unipile APIs to collect employment, salary information and job role trends across multiple member states, as the example below:

Figure 7 Plan for LinkedIn skills mapping



Source: Connecting Women in Digital

While this approach can be relevant to complement the WIDCON methodology, there are current technical barriers to perform such tasks; For example, third party data sources may be limited and be proprietary to the platforms (eg., LinkedIn, although a public platform, provides a limited availability to source data), issues on embedding gender intersectionality principles when genderising names during data collection of roles and salaries, and having oversight on the data collection sample. That said, exploring these processes (albeit the barriers) will be a part of the data collection phase of WIDCON, and any collected data using exploratory approaches will be evidenced during the index building, presentation of results and recommendations.

## Quality control mechanisms

*Table 9 Data quality control mechanisms*

Mechanism	Data type	Actions	Control points
Cross-source validation	Secondary data that has been gathered from desk research	Compare key metrics across multiple sources (e.g., Eurostat vs OECD) to ensure consistency and flag discrepancies.	Any time where it is possible to cross-reference data points, it should be done.
Data gap flagging	Secondary data that has been gathered from desk research	Identify and document missing or outdated data for specific countries or indicators to inform exploratory approaches (e.g., LinkedIn, job portals)	Prior to reporting
Survey design validation	CAWI and CATI questions proposed by IDC	Consortium / partner validation, followed by a further third-party validating the questionnaire	Prior to the roll out of the survey
Mixed survey approaches	Survey data	Conducting both CAWI's (Computer assisted web interviews) and CATI's (Computer assisted telephone interviews) gives Women in Digital the ability to gather profoundly informative data	During the primary field research / surveys.
Manual data revision	Primary initiatives and survey data that have been gathered	The initiatives and the data must be reviewed by humans when the data is collected	After data collection and cleaning

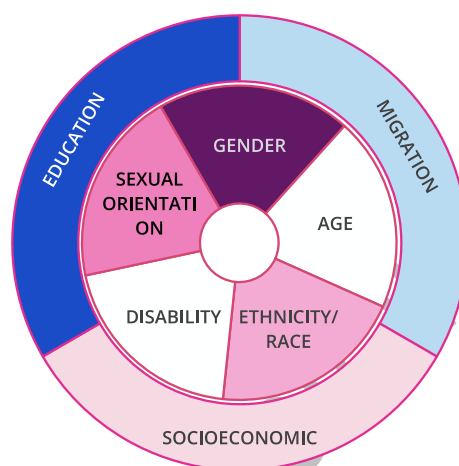


## ANNEX 1 – Addressing diversity as a whole

### Holistic view on diversity

Taking into account the previously identified frameworks and the particular scope of WIDCON, the following model has been elaborated as a starting position. It follows the Loden model for internal and external aspects while tailoring for relevant aspects related to founders and investors.

*Figure 8. Proposed model for definition of diversity (adapted from GENDEX study)*



*Source: Adapted from GENDEX Study*

It is composed of two key layers:

- Internal components – immutable characteristics that remain fixed over time<sup>22</sup> or cannot be readily changed these include:
  - Gender.
  - Ethnicity and Race.
  - Age.
  - Sexual orientation.
  - Disability.
- External components – those characteristics that are related to the past or current experiences and environments of the individual which includes:
  - Educational background.

---

<sup>22</sup> On the use of the term immutable – it is understood by the WIDCON consortium that aspects like gender, ethnicity and race are social constructs and may be externally presented in one form or another and can be fluid in practice but are stable core aspects of an individual's identity. Age is also central to an identity that cannot be changed readily by the individual.



- Migration status.
- Socioeconomic status.

Below the additional individual components are described in detail, providing existing nuance. In the review of the parameters, we cannot as a consortium, make a conclusive statement on all fields and aspects of diversity but attempt to provide an overview. In the end, diversity is someone who doesn't look or talk like you.

## Ethnicity & Race

Important to discussion of ethnicity or race is the scope of the EU-27 to be applied in this index. With that, the concept of ethnicity or race is both highly contextual and linked to national and regional cultures and respective histories. It is also controlled by legislation within various Member States in respect to legality and capacity for collection and categorisation, and further to this is defined alternatively or not at all within what is legally permissible.<sup>23</sup>

The position of the EU and the European Commission is the rejection of definition of ethnicity or race on the basis of race and/or ethnic origin which seek to determine the existences of separate human races.<sup>24</sup> It is recognised and accepted that both ethnicity and race are social constructs, often the labels are decided for by society where one may self-identify or in other cases cannot avoid the assignation.

Addressing race and ethnicity (and national origin), is required for proving where discrimination has taken place and, in the case, here, in the identification where gaps in equality occur for the purpose of effecting change.<sup>25</sup>

The definition of both race and ethnicity is, however, unclear. While race can be taken to refer to geographical origin, colour, or descent it can encompass religion, nationality, ethnicity and other terms. Ethnicity tends to refer to more cultural aspects like religion, shared language and traditions. There is no clear boundary between race and ethnicity and the lines can be drawn arbitrarily and moved along time and are context specific. For example, being Catholic can be considered ethnicity linked to national origin in one case (i.e. Northern Ireland) or solely a religion in another (e.g. Germany), what is considered in one Member State is not the same in the other. Added to this is the historical contexts of migration and global populations, where certain countries like the UK have a track record in applying categories to identify populations that is well accepted, while others have less of a tradition relying on migration or place of birth or do not collect at all.

---

<sup>23</sup> European Commission, Directorate-General for Justice and Consumers, Farkas, L., The meaning of racial or ethnic origin in EU law – Between stereotypes and identities, Publications Office, 2017.

<sup>24</sup> Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin.

<sup>25</sup> European Commission, Directorate-General for Justice and Consumers, Guidance note on the collection and use of equality data based on racial or ethnic origin, 2021

For this reason, given the complexity in definition of the categories and labels across the EU, within WIDCON, where we aim to collect such data, with a reasonable purpose, it should recognise the subjective nature and allow for free self-declaration and multiple selections and following the guidance, be permitted to not answer or declare.

### Sexual Orientation

Addressing the sexual orientation of founders is particularly difficult due to the ‘masked’ nature of this attribute and the potential risks, either real or perceived, for discrimination.<sup>26</sup> This is similar for socioeconomic status, ethnicity or neurodivergence.

Over past decades, the vocabulary for addressing sexuality has developed significantly. While we can define well sexual minorities along accepted terms like gay, lesbian, bisexual, asexual, queer<sup>27</sup> we do not have reliable and statistical sources of data for members of these communities within investment and innovation systems. Existing studies have attempted to extrapolate key assumptions applied to general population, but fundamentally we do not have reliable statistical reference points, save the exception of few mainly western European Member States.

Further to this, is the existing overtly hostile and prejudiced societies within which many operate, and the risk posed to their work and personal outcomes by ‘outing’ themselves for the purpose of a study. If we were to address this within an index, the resulting data would not be considered reliable and will be skewed based on the individual location and/or experiences of a founder or investor.

### Education background

The educational background of both founders and investors is of interest as an indicator of capacity to deliver, subject matter expertise or networks and reach. It is of interest to explore key influences, trends and existing pathways from successful founder and investment teams to observe for convergence in the areas of:

- Educational attainment - EQF levels 1-8<sup>28</sup>
- Subject matter and study – e.g. STEM vs non-STEM
- Attendance to group or network of establishments – e.g. TES University Rankings or elite or top tier business schools

---

<sup>26</sup> Masking refers to the hiding of a specific attribute of one’s identity to avoid negative social consequences projecting a different personality or identity externally or in distinct social groups.

<sup>27</sup> Stonewall - Global Workplace Equality Index

<sup>28</sup> European Qualifications Framework rev. 2017

## Socioeconomic status

In the definition of diversity, while socioeconomic status is important to consider and ensure that there is greater equality of opportunity, it is also another example of a potentially 'masked' attribute.

European Commission defines socioeconomic status is the social standing or class of an individual or group. It is often measured as a combination of education, income and occupation.<sup>29</sup> Individuals' socio-economic status is broadly defined as one's access to resources and is usually measured by income, wealth, education and occupation, all bearing significant implications on nutrition, health and living conditions.<sup>30</sup>

There is a layer of added complexity in that there is no set definition across all EU Member States and set thresholds can vary significantly between one region or another. It is also highly linked and intertwined with all other aspects described and can be dependent or correlated with one or more of them.

## Migration status

With migration status, we refer simply to whether an individual is a national of their country of residence and what was their prior country of residence without prejudice towards the status of the individual. This includes migration from both second and third country individuals, i.e. intra-EU and extra-EU.

The migration status of a founder is anecdotally linked to higher outcomes and propensity to become a founder in the first instance. It is believed that founders from other nations or born to those who have come from other nations have a greater appetite for risk and are more likely to be internationally focused on their start-up journey.

Statistically, however this is not borne out by existing studies. Recently an assessment of the EU context, with the GEM survey, found that while non-natives are more likely to embark on a venture, there is not a positive correlation with outcomes.<sup>31</sup> There is, though, certain evidence that returning migrants enjoy higher rates of venture and success.

For the application within WIDCON, it is more of interest to measure flows of people from within the EU and outside, tracking concentrations of talent or founders, with reference to whether they are a founder operating in their native country or in another country.

---

<sup>29</sup> European Commission Socioeconomic status. Knowledge for Policy, 2021

<sup>30</sup> Parliamentary Assembly of the Council of Europe. Socio-economic inequalities in Europe: time to restore social trust by strengthening social rights., 2021

<sup>31</sup> Rillio, Cesare Fabio Antonio and Peroni, Chiara (2022): Immigration and entrepreneurship in Europe: cross-country evidence

## ANNEX 2 - IDENTIFIED SECONDARY DATA SOURCES

Source:	Data set:
Eurostat	educ_uoe_ent02; educ_uoe_grad10, educ_uoe_grad02; isoc_sks_itsps; isoc_ski_itsex; isoc_sks_itsps; earn_gr_gpgr2; eq_dskl07; lfsi_emp_a; lfst_hheredy;
OECD	OECD's Programme for International Student Assessment (PISA)
IDC	IDC Job Role Forecasting, Role & importance of women as cheque-writers & start-up founders, Social Sustainability — Breakdown by Geography, ESG Training & Upskilling as ESG Maturity Grows, Diversity as a Pivotal Point in Organisational Change, Diversity and Inclusion for the Attraction & Retention of IT Talent?, PitchBook (IDC primary external data provider), European trends in Diversity and Inclusion, IDC Europe Future of Work Employee Experience Survey, DIE Data Sustainability Indices, 2020-2023
GENDEX	D2.2 Quantitative Survey Results, GENDEX Index
PitchBook	<a href="https://pitchbook.com/">https://pitchbook.com/</a>
TIMMS International Database (IEA) – Grade 4	<a href="https://timss2023.org/data/">https://timss2023.org/data/</a>
Atomico	The State of European Tech
EIGE	Gender Equality Index, EIGE's Women and men in decision-making
European Commission	Cohesion Open Data Portal, RRF Country Files
Hal Op	What Can Statistics Tell About the Gender Gap in ICT? Tracing Men and Women's Participation in the ICT Sector
REWIRE	Cyber Equality Now: Tackling Gender Disparities in the European Cybersecurity
Studia Europejskie	Women in ICT Sector in EU States: Facing Gender Inequalities
EPRS	Women in the digital sector
EISMEA	European Startup Scoreboard
JRC	European Innovation Scoreboard: Gender perspective in innovation
e-WOMEN IN ICT	ICT & entrepreneurship access conditions for young women
EQUALS-EU	Women in tech: Breaking down barriers to female leadership
Honeypot	The state of women's representation in Europe
EWOB	EWOB' Gender Diversity Index
Bloomberg	Bloomberg's Gender Equality Index
GEM	Women's Entrepreneurship Report
WEF	Unleashing the power of Europe's women entrepreneurs
StartOut	The State of LGBT Entrepreneurship in the U.S
Statista-FT	Diversity Leaders ranking

GEN	L'observatoire numérique national (emploi/formation)
IP Owners	Diversity in the European Innovation Industry & IP Profession
UN	Minimum Set of Gender Indicators
GOVPH	Women and ICT Development Index (WIDI) Survey
PSADA	Women and ICT Development Index Survey
EITD	Deep Tech Talent Initiative, EIC Community awardees and seal of excellence beneficiaries, EIT Digital Venture & Open Innovation Factory applicants, EIT Digital Alumni, Girls go Circular Cross-KIC activity
Economics	Digital gender divide and convergence in the EU countries

## ANNEX 2 – WOMEN IN DIGITAL SURVEY - SAMPLE QUESTIONS

No	Questions
1	What is your current role in ICT?
2	What is your highest level of education?
3a	Did you pursue any ICT-specific education?
3b	If yes, what type of ICT-specific education did you pursue?
4	What motivated you to pursue a career in ICT?
5	Did you study STEM subjects in your basic education?
6	What barriers did you face in pursuing ICT education?
7	How would you rate your experience in ICT education?
8	What type of support would have helped you in your ICT education?
9	What is your current employment status?
10	How long have you been working in ICT?
11a	Have you participated in any upskilling or reskilling programs in ICT?
11b	If yes, what type of program did you participate in?
12	What challenges do you face in your current role?
13	Do you feel you have equal opportunities for advancement compared to your male colleagues?
14	Have you ever experienced gender discrimination in your ICT career?
15	If yes, in what form did you experience gender discrimination?
16	What is your current job level?
17	What is the % leadership roles (Director level and above) in your organization that are women?
18	What type of contract do you have?
19	How satisfied are you with your current work arrangement?
20	Have you ever taken a career break?
21	If yes, how easy was it to return to your ICT career after the break?
22	How many times have you changed roles within the ICT sector?
23	What is the average time you've spent in each ICT role?
24a	Have you ever transitioned from a technical to a non-technical role in ICT?
24b	If yes, what was the primary reason for the transition?
25	How would you rate the retention efforts for women in your organization?
26	What initiatives does your organization have to support women in ICT?
27	How satisfied are you with the gender diversity in your workplace?
28	Do you feel your ideas and contributions are valued in your workplace?
29a	Have you ever considered leaving the ICT sector?
29b	If yes, what was the primary reason?
30	How likely are you to recommend a career in ICT to other women?
31	What do you think is the biggest barrier for women entering the ICT sector?
32	What resources or support would help you advance in your ICT career?
33	How do you stay updated with the latest developments in ICT?
34	Do you actively mentor other women in ICT?
35	What do you think organizations can do to attract and retain more women in ICT roles?
36	How optimistic are you about the future of women in ICT?

## ANNEX 3 – IDENTIFIED DIGITAL / ICT ROLES

Eurostat and OECD's definitions of digital roles			
I. ICT MANAGERS, PROFESSIONALS AND ASSOCIATE PROFESSIONALS			
	133	ICT Service managers	
25	Information and communications technology professionals		
	251	Software and applications developers and analysts	
		2511	Systems analysts
		2512	Software developers
		2513	Web and multimedia developers
		2514	Application programmers
		2519	Software and applications developers and analysts not elsewhere classified
	252	Database and network professionals	
		2521	Database designers and administrators
		2522	Systems administrators
		2523	Computer network professionals
		2529	Database and network professionals not elsewhere classified
35	Information and communications technicians		
	351	ICT operations and user support technicians	
		3511	ICT operations technicians
		3512	ICT user support technicians
		3513	Computer network and systems technicians
		3514	Web technicians
	352	Telecommunications and broadcasting technicians	
		3521	Broadcasting and audio-visual technicians
3522		Telecommunications engineering technicians	
II. OTHER UNIT GROUPS THAT PRIMARILY INVOLVE THE PRODUCTION OF ICT GOODS AND SERVICES			
	2152	Electronic engineers	
	2153	Telecommunications engineers	
	2166	Graphic and multimedia designers	
	2356	Information technology trainers	
	2434	ICT sales professionals	
	3114	Electronics engineering technicians	
	742	Electronics and telecommunications installers and repairers	
		7421	Electronics mechanics and servicers
		7422	ICT installers and servicers

## ANNEX 4 EDUCATION LEVELS CLASSIFICATION (POST 2011)

ISCED 2011 Education classification	
Level 0	Less than primary education
Level 1	Primary education
Level 2	Lower secondary education
Level 3	Upper secondary education
Level 4	Post-secondary non-tertiary education
Level 5	Short-cycle tertiary education
Level 6	Bachelor's or equivalent level
Level 7	Master's or equivalent level
Level 8	Doctoral or equivalent level

Pending approval from EC