



REWIRE - Cybersecurity Skills Alliance A New Vision for Europe

R1.8. Open Data procedures and guidelines



Title	R1.8. Open Data procedures and guidelines
Document description	The Open Data procedures and guidelines document contains a description regarding the activities of the project involving personal and non-personal information, as well as a description of relevant principles, procedures and guidelines regarding the handling of open data.
Nature	Public
Task	Task 1.3. Open Data Management and Data Protection
Status	Final
WP	1
Lead Partner	APIROPLUS Solutions
Partners Involved	All partners
Date	11.11.2021

Revision history	Author	Delivery date	Summary of changes and comments
Version 01	Karras Apostolos	06/10/2021	Initial Version
Final Version	Karras Apostolos	11/11/2021	Final version

Disclaimer:

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

CONTENTS

1. Executive Summary	3
2. Introduction	4
3. Data summary	4
3.1. Introduction	4
3.1. Data and purpose of the collection / generation	5
3.2. Types and formats of data the project will generate/collect.	7
3.3. Origin of data	8
3.4. Data utility.....	9
4. Open data and open access.....	9
4.1. Research Data - Open access	11
4.2. FAIR data management.....	12
4.3. Open access to educational resources	13
5. List of Figures	14
6. List of Tables.....	15

1. EXECUTIVE SUMMARY

This deliverable presents the methodology that will be followed in the REWIRE project in relation to non-personal and Open Data Management, the data that will be collected, created or processed during the project and how these data will be stored or shared.

In order to correctly identify the relevant data, a questionnaire was distributed to all relevant partners and their comments were collected and analyzed, in relation to the data that will be managed during the project lifecycle.

The analysis of the data led to the creation of data maps (personal, non-personal and open data). These maps and the relevant information provided in this document, can be used by the project partners in their decision regarding the management of data.

2. INTRODUCTION

The Cybersecurity Skills Alliance – New Vision for Europe – REWIRE project develops a Blueprint for the Cybersecurity industry and a concrete European Cybersecurity Skills Strategy. It brings together 25 partners from academia and VET, cybersecurity industry, non-cyber industries, certification partners and umbrella organizations. Its work builds upon four pilot projects: CONCORDIA, SPARTA, ECHO, CyberSec4Europe implemented with the support of the European Union’s Horizon 2020 research and innovation programme.

The project deliverables are:

- A **EUROPEAN CYBERSECURITY BLUEPRINT**, addressing skills gaps in the cybersecurity sector.
- A **CYBERRANGE PLATFORM**, based on open-source components with specific elements and features.
- **CERTIFICATION SCHEMES**, implemented covering specific areas of cybersecurity and following international best practices.
- The **CyberABILITY** platform, A digital on-line publicly accessible European Cybersecurity Skills Digital Observatory which will provide up-to-date information regarding the job market, competences, training courses, certification schemes and a career roadmap.
- **ONLINE COURSES** and trainings on selected occupational profiles offered in the form of four (4) VOOCs.

3. DATA SUMMARY

3.1. Introduction

As described in R1.7 PRIVACY POLICY & PROCEDURES, a process was implemented almost from the start of the project, to identify the data that will be processed within the project’s lifetime.

This process included the collection of an initial estimation of data that will be processed by the project partners as part of their activities within the REWIRE project. The results of the collection were further processed, and specific procedures, policies and a file of processing activities in relation to personal data was created (please refer to R1.7 PRIVACY POLICY & PROCEDURES document).

The collection of data was facilitated through a specifically constructed questionnaire. This questionnaire introduces the concepts of data (personal and non-personal) and the answers to the following questions were requested:

- Type of Data to be collected or generated
- Description of Data
- Purpose of processing (Why we need this data)
- Task under which the data will be processed / generated / processed
- Source of the Data
- If the data or the output proprietary?

In relation to the non-personal data, this document contains the results of this initial estimation as well as the relevant guidelines and principles.

3.1. Data and purpose of the collection / generation

Within this section the data that are envisioned to be collected, generated or processed within the project are presented.

The data are correlated to the different aims and activities of the project and for completeness purposes, both personal and non-personal data are displayed.

In every case, the data are always correlated to the various objectives of project, providing an initial base for the processing.

The objectives of the REWIRE project are:

- design and deliver the European Cybersecurity Blueprint,
- develop the European Cybersecurity Skills Framework, update existing and create 4 new occupational profiles,
- deliver training programmes on highly innovative fields,
- develop a digital on-line publicly accessible Skills Observatory for cybersecurity skills merging the market needs, the profiles, the competencies and the available training courses.
- involve all stakeholders for exploiting the VET potential in cutting-edge subjects for creating growth and jobs in the Cybersecurity sector,
- enhance the use of cyberranges,
- promote the application of EQAVET and EQF/ECVET frameworks that ensure both quality and better transferability of the project's results.
- create a lasting partnership of all types of stakeholders that will monitor and adjust to changes in the sector's needs,
- facilitate transnational mobility between the sectors' stakeholders,
- provide transversal skills as well as career guidance, career management skills and access to the labour market, thus improving their long-term employability.

The following maps, depict the data that will be processed during the life-cycle of the project.

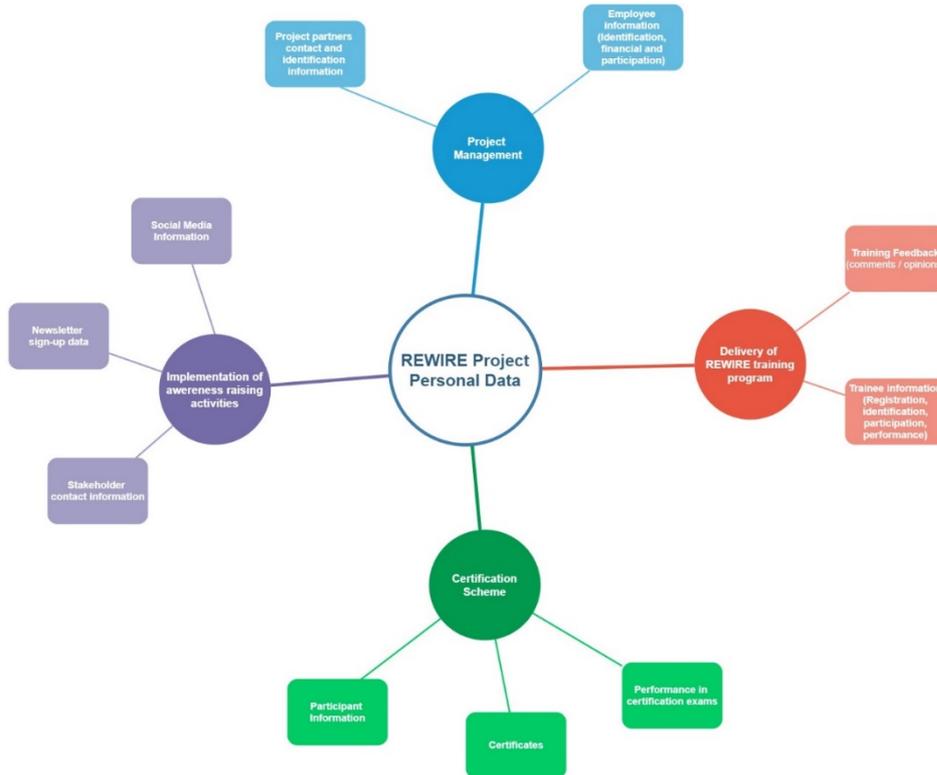


Figure 1. Map of Personal Data



Figure 2. Map of Non-Personal Data

3.2. Types and formats of data the project will generate/collect.

The following table contains the types and formats of the data that the project will generate, collect, process during its life cycle.

Type of Data	Format	Description of Data
Personal Data	Digital	Project partners contact and identification information
Personal Data	Digital	Employee information (Identification, financial and participation)
Personal Data	Digital	Training Feedback (comments / opinions / complaints etc)
Personal Data	Digital	Trainee information (Registration, identification, participation, performance)
Personal Data	Digital	Participant Information
Personal Data	Digital	Certificates
Personal Data	Digital	Performance in Certification Exams
Personal Data	Digital	Newsletter signup data
Personal Data	Digital	Stakeholder contact information
Personal Data	Digital	Social Media Information
Non-Personal Data	Digital	Web Site Information (Cookies)
Non-Personal Data	Digital	Policy Recommendations
Non-Personal Data	Digital	Key Stakeholders
Non-Personal Data	Digital	VLE training material
Non-Personal Data	Digital	Survey on skill needs
Non-Personal Data	Digital	Occupational Profiles
Non-Personal Data	Digital	Competencies
Non-Personal Data	Digital	Event Presentation
Non-Personal Data	Digital	Other Dissemination Activities
Open Data	Digital	Standardization Inputs
Open Data	Digital	Stakeholders Inputs
Open Data	Digital	PESTLE Analysis
Open Data	Digital	Scientific Papers
Open Data	Digital	Journal Publications

Open Data	Digital	Frameworks from other countries
Open Data	Digital	Knowledge, Abilities, Skills, Tasks

Table 1. Types and formats of data

3.3. Origin of data

The following table contains the description of the data and the source from where the data collected.

Description of Data	Source of Data
Project partners contact and identification information	Project partners
Employee information (Identification, financial and participation)	Data subject
Training Feedback (comments / opinions / complaints etc)	Data subject
Trainee information (Registration, identification, participation, performance)	Data subject
Participant Information	Data subject
Certificates	Data subject
Performance in Certification Exams	Data subject
Newsletter signup data	Data subject
Stakeholder contact information	Project partners
Social Media Information	Data subject
Web Site Information (Cookies)	Cookies consent box
Policy Recommendations	Project partners
Key Stakeholders	Project partners
VLE training material	Project partners
Survey on skill needs	Data subject
Occupational Profiles	Project partners
Competencies	Project partners
Event Presentation	Project partners
Other Dissemination Activities	Project partners
Standardization Inputs	Project partners

Stakeholders Inputs	Stakeholders
PESTLE Analysis	Project partners
Scientific Papers	Project partners
Journal Publications	Project partners
Frameworks from other countries	Public domain
Knowledge, Abilities, Skills, Tasks	Project partners

Table 2. Origin of data

3.4. Data utility

The project Cybersecurity Skills Alliance – A New Vision for Europe, in short REWIRE, is co-funded by the ERASMUS+ programme of the European Union.

REWIRE aims to build a Blueprint for the Cybersecurity industry and a concrete European Cybersecurity Skills Strategy. Its work focuses on delivering concrete recommendations and sustainable solutions that will lead to the reduction of skill gaps between industry requirements and sectoral training provision and contribute to the growth, innovation and competitiveness of the Cybersecurity sector.

It should be noted that the data of this project as mentioned above, are proprietary and belong to the various partners as identified internally.

No information may be re-used without the written authorization of the appropriate project partner. Information regarding open data and open access of data of the project are provided in the following sections.

4. OPEN DATA AND OPEN ACCESS

Open Data is based on the idea that public information should be freely available for use and re-use.

The following diagram depicts the spectrum where data exists as depicted by the Open Data Institute (<https://theodi.org/about-the-odi/the-data-spectrum/>). The spectrum ranges from closed data to open data.

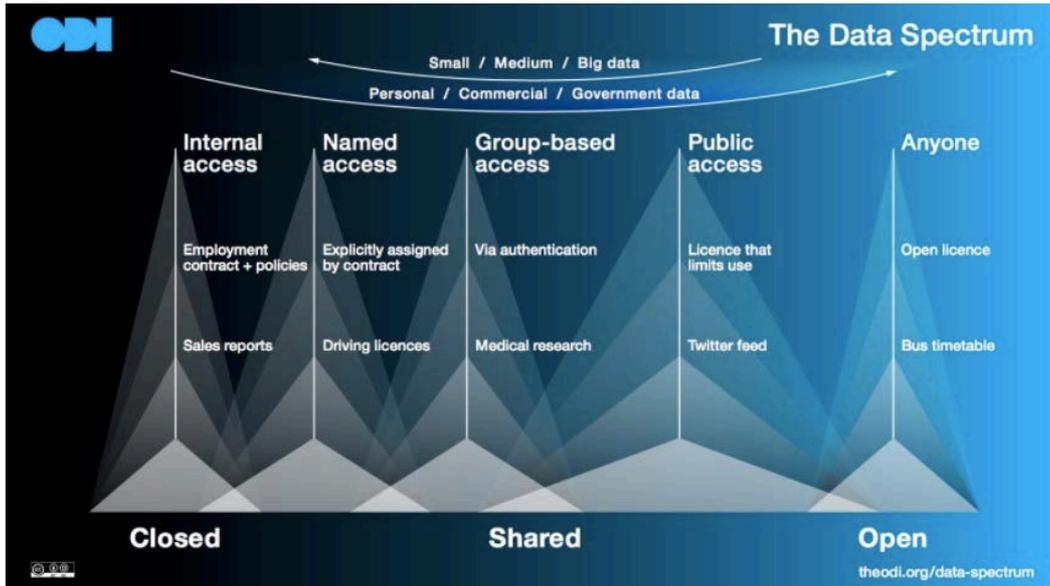


Figure 3. The Data Spectrum

As mentioned before, questionnaires were sent to the project partners in order to, through their answers, identify the data generated, collected and any way processed and identify their classification based on the above spectrum.

The following map contains data that could possibly be identified as open data. The data identified in the map constitute inputs or outputs of the various processes of the process and are provided here for reference purposes only (as examples).

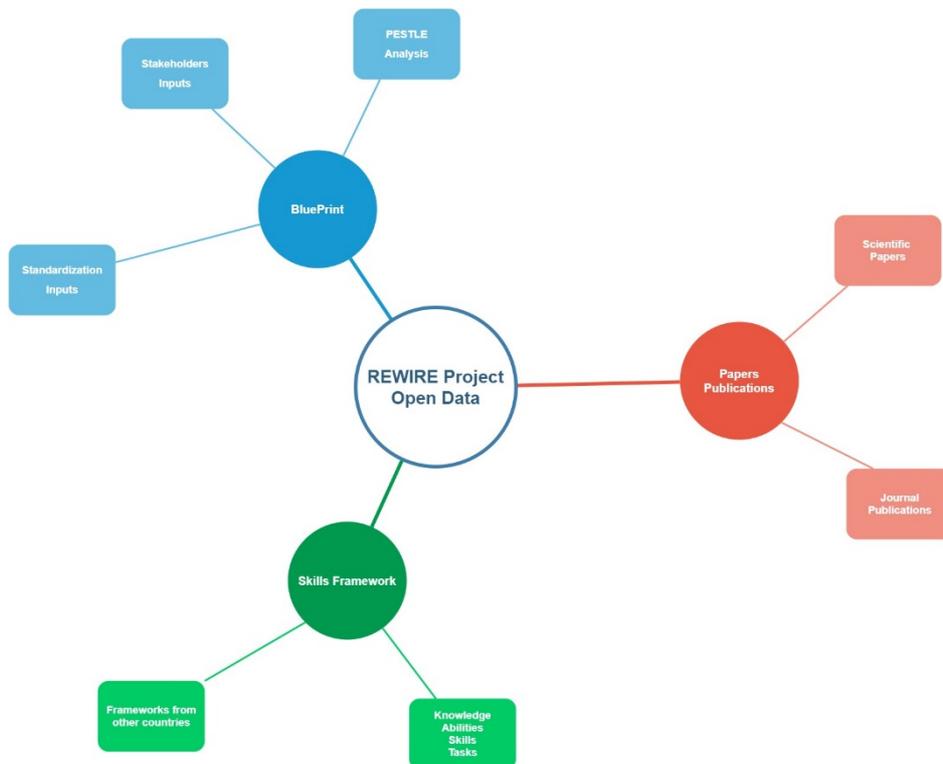


Figure 4. Map of Open Data

To further promote understanding the following examples are provided:

In relation to the development of the Skills Framework, the project partners may use as reference material Frameworks and sets of Knowledge, Abilities, Skills and Tasks from other countries or organizations. The project partners should make sure, before re-using such information, that they are allowed to reuse the information and or modify it, as well as if there are any specific conditions for such reuse (attribution, notification etc). At the same time, the project partners will work on the compilation of Role Profiles and the determination of the knowledge, skills, abilities and tasks that these Role Profiles should encompass. The Role Profiles and any adaptation of the knowledge, skills, abilities and tasks would then be provided openly by the project partners to the relevant systems (e.g. the Cyberability platform, the Skills Panorama, ESCO etc).

Since, the project is still in progress, the identification process will continue and when other data is discovered, this document will be updated and directions will be provided to the project partners.

In the following subsection, some more information is provided on the topic of open access to scientific information and about the FAIR principles that should be followed for such information. This information is provided as a guidance to the project partners to assist them during the identification of data and the decisions regarding the management of such data.

4.1. Research Data - Open access

The following information refers to Research Data. In this context Research data, refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation.

In a research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The focus is on research data that is available in digital form.

Open access (OA)¹ refers to the practice of providing online access to scientific information that is free of charge to the end-user and reusable. 'Scientific' refers to all academic disciplines. In the context of research and innovation, 'scientific information' can mean:

- peer-reviewed scientific research articles (published in scholarly journals) or
- research data (data underlying publications, curated data and/or raw data).

In the context of research funding, open access requirements do not imply an obligation to publish results. The decision to publish is entirely up to the project partners. Open access becomes an issue only if publication is chosen as a means of dissemination.

The following figure, depicts the decision tree that each of the project partners could follow in order to identify the options regarding access to research data.²

¹ [Open access | European Commission \(europa.eu\)](https://ec.europa.eu/euro-iss/faq/faq-open-access)

² Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, Version 3.2, 21 March 2017

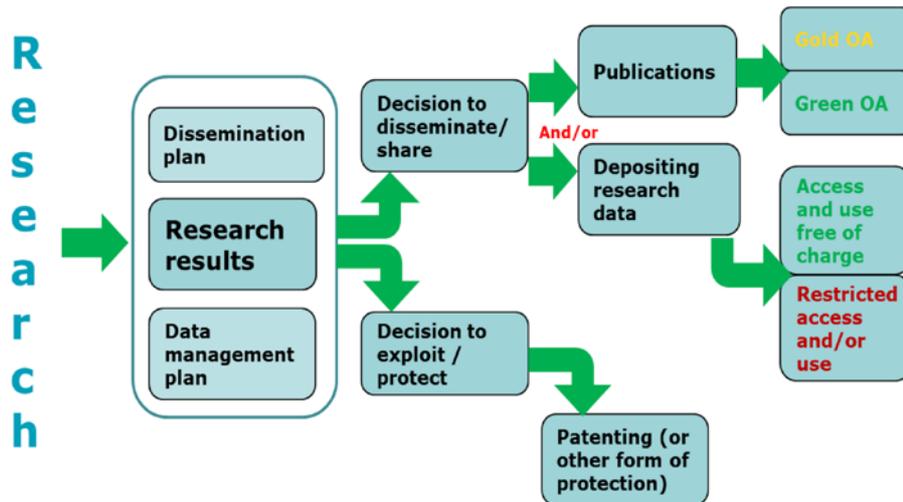


Figure 5. Open Access to Scientific Data Decision Tree

4.2. FAIR data management

In general terms, the research data should be 'FAIR', that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard, or implementation solution.

In 2016, the 'FAIR Guiding Principles for scientific data management and stewardship'³ were published in Scientific Data. The authors intended to provide guidelines to improve the Findability, Accessibility, Interoperability, and Reuse of digital assets. The principles emphasise machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention) because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data.⁴

In the following paragraphs, some more information is provided on these 4 principles:

Findable

The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the FAIRification process.

Accessible

Once the user finds the required data, she/he needs to know how can they be accessed, possibly including authentication and authorisation.

Interoperable

The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.

Reusable

The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

³ <http://www.nature.com/articles/sdata201618>

⁴ Source: FAIR Principles - GO FAIR (go-fair.org)

The principles refer to three types of entities: data (or any digital object), metadata (information about that digital object), and infrastructure. For instance, principle F defines that both metadata and data are registered or indexed in a searchable resource (the infrastructure component).

4.3. Open access to educational resources

Open Educational Resources (OER) are teaching, learning and research materials in any medium – digital or otherwise – that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation and redistribution by others with no or limited restrictions.⁵

The REWIRE project is committed to provide and adopt Open Educational Resources in English.

A key project priority is the production of Open Educational Resources in English, with the possibility to extend to more European languages after the end of the project.

The project's educational resources will be offered freely and openly for educators, students and self-learners for use, reuse, adaptation and sharing. Constraints imposed will be regarding commercial reuse of the aforementioned material as well as giving appropriate credit and license when reusing the material. Thus, the resources will be released under the Creative Commons license: Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). These resources include lesson plans, courses material, learning objects, digital teacher and trainee guides as well as e-assessment tools tailored to the project's learning objectives. In order to access these resources the user will be required to register to the project portal. All resources will be made available in downloadable formats so that the user can store them locally and access them when offline (such as text documents, presentations and videos). Learning objects will be developed under the SCORM specification (sharable content object reference model) so that they can be easily loaded into a Learning Management System (LMS) such as edX.

The quality of these open educational resources will be assured based on OER commons recommendations, as well as the EU Open Education 2030 vision on lifelong learning.

This approach aims to foster peer collaboration and to provide easily accessible material to all interested parties. It also aims to promote reuse of the developed resources, a key element for the sustainability of the project's results.

⁵ UNESCO, <https://en.unesco.org/themes/building-knowledge-societies/oer>

5. LIST OF FIGURES

Figure 1. Map of Personal Data	6
Figure 2. Map of Non-Personal Data	6
Figure 3. The Data Spectrum	10
Figure 4. Map of Open Data.....	10
Figure 5. Open Access to Scientific Data Decision Tree.....	12

6. LIST OF TABLES

Table 1. Types and formats of data	8
Table 2. Origin of data.....	9