EXPLORATORY RESEARCH

sesar JOINT UNDERTAKING

D1.7 Data Management Plan

Deliverable ID:	**D1.7**
Dissemination Level:	**PU**
Project Acronym:	**SlotMachine**
Grant:	**890456**
Call:	**H2020-SESAR-2019-2**
Topic:	**SESAR-ER4-27-2019 Future ATM Architecture**
Consortium Coordinator:	**Frequentis AG**
Edition date:	**31. October 2022**
Edition:	**01.01.02**
Template Edition:	**02.00.05**

EUROPEAN PARTNERSHIP

Co-funded by the European Union

Authoring & Approval

Authors of the document

Name / Beneficiary	Position / Title	Date
Christoph Fabianek / Frequentis	Technical Manager	2022-10-03

Reviewers internal to the project

Name / Beneficiary	Position / Title	Date
Marie Carré / SWISS	WP 5 Leader	2022-10-19
Eduard Gringinger / Frequentis	Project Manager	2022-10-27

Reviewers external to the project

Name / Beneficiary	Position / Title	Date
none		

Approved for submission to the SJU By - Representatives of all beneficiaries involved in the project

Name / Beneficiary	Position / Title	Date
Christoph Fabianek / Frequentis	Technical Manager	2022-10-17
Eduard Gringinger / Frequentis	Project Manager	2022-10-27
Marie Carré / SWISS	WP 5 Leader	2022-10-19
Nadine Pilon / EUROCONTROL	WP2 Co-Leader	2022-10-31 (silent approval)
Christoph Schuetz / JKU	WP2, 3 Leader	2022-10-31 (silent approval)
Lorünser Thomas / AIT	WP4 Leader	2022-10-31 (silent approval)

Rejected By - Representatives of beneficiaries involved in the project

Name and/or Beneficiary	Position / Title	Date

Document History

Edition	Date	Status	Name / Beneficiary	Justification
01.00.00	2022-08-18	First draft	C. Fabianek	Initial draft
01.01.00	2022-10-03	Internal release	C. Fabianek	Ready for approval
01.01.01	2022-10-31	Final version	C. Fabianek	DMP release ready to deliver to SJU
01.01.01	2022-11-18	Final version	C. Fabianek	Update Copyright Statement

SlotMachine

A PRIVACY-PRESERVING MARKETPLACE FOR SLOT MANAGEMENT

This Data Management Plan is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 890456 under European Union's Horizon 2020 research and innovation programme.

Abstract

The document is the Data Management Plan (DMP) of the SlotMachine project. This document describes how the data collected or generated by the SlotMachine project will be handled during and after the project closure, describes which standards and methodology for data collection and generation will be followed, and whether and how data will be shared.

SlotMachine employs blockchain technology and secure multi-party computation to extend the existing User-Driven Prioritisation Process (UDPP) solution with the possibility to keep private the participating airlines' confidential information such as the cost structure of flights. Technology will allow for secure, auditable transactions without the need for a central broker, whereby stakeholders will be able to enter slot swapping transactions without disclosing information to other participants. By demonstrating the feasibility of a privacy-preserving platform for swapping ATFM slots, the foundation can be laid for the development of a product that will be an essential element in the aviation industry in the future. It contributes to better use of existing resources at airports, higher efficiency of airlines, lower emissions, and shorter delays for passengers.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

Table of Contents

Abstract 3
1 Introduction 5
1.1 Document Maintenance 6
1.2 Applicable Reference Material 6
1.3 Rules Reminder on Communication or Dissemination of Project Results 6
2 Data Summary 7
2.1 Project Objectives and related Data Collection 7
2.2 Data Description 7
2.3 Data Utility 9
3 FAIR Data 10
3.1 Making Data Findable, Including Provisions for Metadata 10
3.2 Making Data Openly Accessible 10
3.3 Making Data Interoperable 12
3.4 Data Re-use & Licensing Data 13
4 Allocation of Resources 14
4.1 Estimated Cost 14
4.2 Responsibilities 14
4.3 Long-Term Preservation 14
5 Data Security 15
6 Legal and Ethical Aspects 16
6.1 Data Protection 16
6.2 Measures to Ensure Ethical and Legal Standards 16
6.3 Privacy and Trust 16
7 References 17

List of Figures

Figure 1: Data Lifecycle 5
Figure 2: Basic classes used for documenting provenance 12

EUROPEAN PARTNERSHIP
Co-funded by the European Union

1 Introduction

Annex 1 of the Grant Agreement (GA - 890456) [1] provides the contractual baseline of the project by means of the Description of the Action. It contains a high-level description of what the project is aiming to achieve and how and which additional outputs will be produced in terms of Deliverables.

Figure 1 depicts the general data lifecycle foreseen in the SlotMachine project:

- Various stakeholders define dataset requirements (stakeholders and high-level requirements are described in Deliverable 2.3 Business Concepts)
- A data engineer generates the respective datasets provided by project partners (SWISS: real-word data, JKU: synthetic data) and makes those datasets available in data containers (see chapter 3)
- The SlotMachine prototype will be developed and iteratively updated using these datasets and feeds back insight to refine datasets
- Public access to generated data (synthetic data) is also provided through the above-mentioned data containers

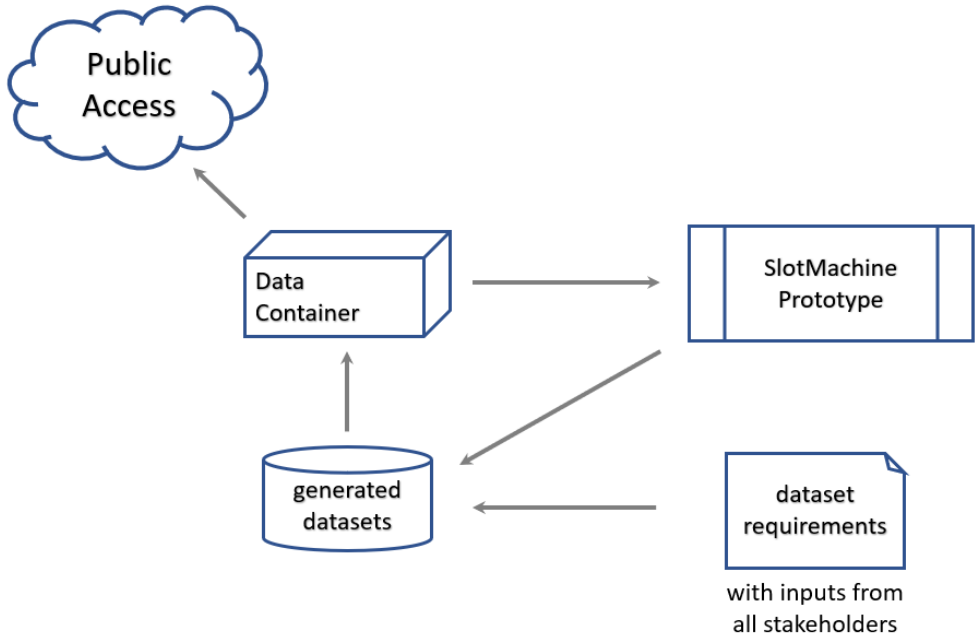

Figure 1: Data Lifecycle

Note: Public access to data is for datasets to demonstrate the functionality of software components in the SlotMachine project. In a production environment, data from Airspace Users about preferences to exchange slots is of course confidential and not publicly available. Indeed, it is one of the core objectives in SlotMachine to keep such data private in any circumstances.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

1.1 Document Maintenance

The Data Management Plan (DMP) has been produced at the beginning of the project as a contractual deliverable and was updated throughout the project ([3] & [4]). Its main purpose is to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur.

The Project manager (as defined in D1.1 Project Management Plan [2]) oversees the DMP updates. The DMP is now in its final version (at the end of the last reporting period) and describes all data sets that are used in testing and demonstrating the system.

1.2 Applicable Reference Material

Unless otherwise stated in this deliverable, the execution of the project will be fully compliant with the latest version of the S2020 Project Handbook [10].

1.3 Rules Reminder on Communication or Dissemination of Project Results

Any communication or dissemination of the project results (in any form, including electronic) will display the SJU logo and the EU emblem, and will include the following text:

"This project has received funding from the SESAR Joint Undertaking under grant agreement No 890456 under European Union's Horizon 2020 research and innovation programme".

When displayed together with another logo, the SJU logo and the EU emblem will have appropriate prominence.

Co-funded by the European Union

2 Data Summary

2.1 Project Objectives and related Data Collection

SlotMachine has the following detailed objectives:

- **Objective #1**: Development of a novel slot swapping platform to improve utilization of available resources at airports and reduce cost for airlines
- **Objective #2**: Research and development for a secure and trustworthy system for slot swapping which employs an evolutionary algorithm in conjunction with the PrivacyEngine in a privacy-preserving way
- **Objective #3**: Realization of a proof-of-concept implementation of the SlotMachine Platform Demonstrator, which offers privacy-preserving slot management

All objectives involve data collection and in the following section, these objectives will be referenced when describing the relevant data.

Note: While it is the objective of the SlotMachine project to develop a privacy-preserving marketplace, we provide in this Data Management Plan mainly datasets to demonstrate the functionality using synthetic data. This data is openly shared so that others can more easily work with the developed open-source components (Heuristic Optimizer, Privacy Engine). Real-world data provided by project partner SWISS is not shared but has the same data structure as the provided synthetic data.

2.2 Data Description

For Objective #1 data is identified as follows:

- Resource description of airports and airlines
  data consumed by project – dataset "NM Simulation" (publicly available)
  format: XML, size: 16.760 bytes
  description: this dataset is a subset of data generated during an UDPP Trial exercise performed by SWISS in October 2021 as part of Solution 39 (unrelated to SlotMachine) and is based on data provided by the NM B2B Interface (Flight List By Aerodrome messages); SWISS selected a relevant dataset that could be made public; this dataset is an example for input data necessary from NM

- Example slot configurations and demonstration of successful reordering, i.e., show cases of improved utilization through slot swapping
  data generated by project – dataset "AU Preferences" (publicly available)
  format: JSON, size: 5.938 bytes
  description: the dataset with synthetic data was generated to demonstrate functionalities of the SlotMachine Heuristic Optimizer and is based on feedback about airspace user preferences from SWISS; it presents the full life-cycle of an optimization from an initial flight-sequence → preferences provided by airspace users → optimization run with multiple solutions → and

EUROPEAN PARTNERSHIP
Co-funded by the European Union

finally a selected successful reordering of the initial flight sequences submitted to NM (this dataset demonstrates the data format for submitting solutions to NM as specified in D2.1 Requirements Specification [5])

For Objective #2 data is identified as follows:

- Edge cases to demonstrate usefulness of genetic algorithms:
  data generated by project – dataset "Optimizer Evaluation" (publicly available)
  format: JSON, size: 19.116 files
  description: this dataset was generated during development of the Heuristic Optimizer and includes input data, configuration, and results

- Data to demonstrate performance capabilities of the PrivacyEngine:
  data generated by project – dataset "PrivacyEngine Evaluation" (publicly available)
  format: results described in D3.3 Privacy Engine Software Component [7]
  description: data generated during development of the Privacy Engine to evaluate run-time behaviour of the system

For Objective #3 data is identified as follows:

- Real-world data:
  data consumed by project – dataset "Swiss22" (private data not publicly shared)
  format: XML & JSON, size: 274 files
  structure:
    - b2b.xml – response from NM B2B IF (FlightListByAerodrome) query
      same format as in "NM Simulation" dataset from Objective #1
    - AU-<3-letter ICAO code of airline>-preferences.json – preference (margin+prioriy as specified in D2.2 System Design Document [6]) of an airline for each flight;
      same format as in "AU Preferences" dataset from Objective #1

  description: this dataset was generated by SWISS in the time between August and October 2022 and includes data provided by the NM B2B Interface (FlightListByAerodrome messages) and Airspace User preferences for those flights from SWISS

- Results and performance analysis records:
  data generated by project – dataset "SlotMachine Evaluation" (private data, details not publicly shared and overall results documented in D5.1 Platform Demonstrator [9])
  format: JSON, size: 124 files
  structure:
    - solution.json – file with all solutions as calculated by the Heuristic Optimizer
    - solution_3-letter ICAO code of airline>.json – individual solution of flight reordering for each airline together with information about credit exchange
    - *.log – log output files

  description: this dataset comprises of the evaluation output generated D5.1 Platform Demonstrator [9] and was used for the Advisory Board #3 to demonstrate the SlotMachine Prototype capabilities

EUROPEAN PARTNERSHIP

Co-funded by the European Union

2.3 Data Utility

The data was used, during the project execution, as main input for each WP to trigger software development and solution prototyping and validation. There are software scripts and programs (e.g., Java code or python script) that were developed to allow data transformation and model generation and assessment.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

3 FAIR Data

3.1 Making Data Findable, Including Provisions for Metadata

Data compiled for the purpose of testing and demonstration of the SlotMachine prototype will be created in an iterative way and made available publicly. The technical framework of Docker containers is used to package data together with metadata: Semantic Container[1] provides a platform for sharing commercial and non-commercial data. To make data discoverable, metadata is provided through API endpoints[2]. A description of the metadata attributes can be found in the current version of the Semantic Container Design Document[3]. Initial work on Semantic Container was performed by project partners Frequentis and Johannes Keppler Universität in the course of BEST (Achieving the Benefits of SWIM by making smart use of Semantic Technologies) – a SESAR Exploratory Research project (#699298) finished in 2018.

3.2 Making Data Openly Accessible

Semantic Container can store and provide any kind of data and it is up to the data providing container operator (data controller) to decide if and how data is made accessible to defined recipients.

In case data is shared between 2 parties the following mechanisms are used to provide access control:

1. OAuth 2.0: OAuth is an open standard for access delegation and Semantic Container use the Authorization Framework based on Bearer Token Usage as described in RFC 6750[4].
2. Usage Policy matching: two actors can exchange data if and only if the receiving container (data controller) has a Usage Policy that is equal or a subset of the providing container (data subject). A semantic reasoner is used to evaluate compliance between the Usage Policies. Note: Usage Policies are based on the policy language as defined by the SPECIAL project[5] in a specific deliverable [11].
3. Digital Watermarking: a unique digital fingerprint can be applied to the provided data, i.e., any data request results in a dataset with insignificant errors that uniquely identifies the recipient of the data set; in case such a dataset is leaked and appears in an unintended location, the actor who originally requested and leaked the dataset can be identified

The datasets listed in chapter 2 are publicly available at the following locations:

[1] https://www.ownyourdata.eu/semcon/

[2] https://api-docs.ownyourdata.eu/semcon/

[3] https://www.ownyourdata.eu/semcon/design

[4] https://datatracker.ietf.org/doc/html/rfc6750

[5] https://www.specialprivacy.eu

EUROPEAN PARTNERSHIP

Co-funded by the European Union

- **Dataset "NM Simulation"**
  - description: this dataset holds information to simulate the Network Manager B2B Interface
  - content: 1 regulation (=scope for an optimization) with 10 flights
  - API endpoints documentation: https://slotmachine-nm-simulation.data-container.net/api-docs/index.html
  - Docker Image for long term storage: https://hub.docker.com/r/slma/sc-nmf_220310
  - retrieve data at the following endpoint: https://slotmachine-nm-simulation.data-container.net/api/FlightListByAerodrome
- **Dataset "AU Preferences"**
  - description: this dataset defines slot preferences of airspace users for a given regulation
  - content: AU preferences for 10 flights (mapped to NM Simulation dataset)
  - API endpoints documentation: https://slotmachine-au-simulation.data-container.net/api-docs/index.html
  - Docker Image for long term storage: https://hub.docker.com/r/slma/sc-ausim_220310
  - Retrieve data at the following endpoint: https://slotmachine-au-simulation.data-container.net/api/data
- **Dataset "PrivacyEngine Evaluation"**
  - description: the Privacy Engine was evaluated in the course of writing Deliverable 3.3 PrivacyEngine Software component
  - content: see section 5.1.2 Performance and Scalability for details
  - the document is available here: https://www.frequentis.com/sites/default/files/support/2022-01/D3.3%20PrivacyEngine%20Software%20component.pdf
- **Dataset "Optimizer Evaluation"**
  - description: generated data during development of the Heuristic Optimizer; it includes input data, configuration, and results
  - content: datasets produced for Deliverable 4.1 [8]
    - 9504 input data records (id: "testsNrAAA-*")
      example: https://slotmachine-optimizer.data-container.net/api/data/testsNr021-13-3--c099ed24-350d-40fb-800b-0963905dcd84?p=dri
    - 108 configuration records (id:"testsNrAAA")
      example: https://slotmachine-optimizer.data-container.net/api/data/testsNr021?p=dri
    - 9504 result data records (id: "result--testNrAAA-*")
      example: https://slotmachine-optimizer.data-container.net/api/data/result--testsNr021-13-3--c099ed24-350d-40fb-800b-0963905dcd84?p=dri

  Comments regarding nomenclature of the datasets: AAA-*
  - AAA is the 3-digit internal test number (from 001-108)
  - * is either B-C--E or B-C-D--E (only for certain configuration numbers)
  - B is the experiment number, details in D4.1, Appendix A.1, Table 3
  - C is the configuration number, details in D4.1 Appendix A.2, Table 4 and 5
  - D is the repetition number of the experiment with the given configuration
  - E is the optimization identification of the record

EUROPEAN PARTNERSHIP

Co-funded by the European Union

- API endpoints documentation: https://slotmachine-optimizer.data-container.net/api-docs/index.html
- Docker Image for long term storage: https://hub.docker.com/r/slma/sc-optimizer_220310
- Retrieve data at the following endpoint: https://slotmachine-optimizer.data-container.net/api/data

Note: hosting the container at https://data-container.net is provided during the project duration and about a year after project completion. To be independent of this hosting, Docker Images with all the data, meta-attributes, and necessary API logic for easy data access are provided on the free Docker Hub platform in the public SlotMachine repository available here: https://hub.docker.com/orgs/slma/repositories

A Docker image can be started to make the data available at the specified endpoints by starting the instance with the following command (example)

```
docker run -d -p 3000:3000 slma/sc-nmf_220310
```

and is then accessible via http://localhost:3000.

3.3 Making Data Interoperable

To make data interoperable the Semantic Container platform provides a standardized API described above (bullet point "API endpoints documentation" for each dataset) and inter-container data operations generate a well-defined provenance trail based on the W3C Prov-Ontology[6]. This Provenance Ontology is used as-is (without extending any classes) and uses the basic framework as depicted in Figure 2.

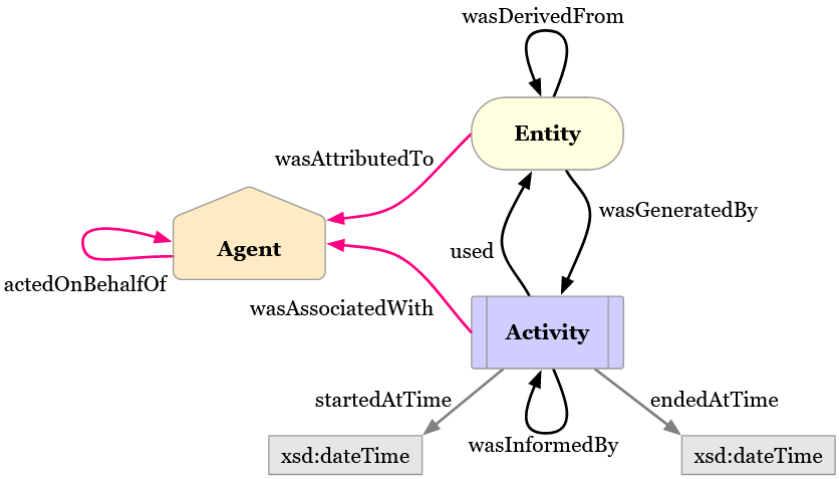

Figure 2: Basic classes used for documenting provenance

[6] https://www.w3.org/TR/prov-o/

Co-funded by the European Union

3.4 Data Re-use & Licensing Data

All data, data operations, and the Docker container itself are cryptographically hashed and the information is written into the public Ethereum blockchain for auditability using the OwnYourData Notary service[7].

Note: not the data itself but only the fingerprint of the data (the cryptographic hash value) will be stored on the blockchain

Synthetic data compiled in the course of the SlotMachine project is licensed under the CC-BY-SA 4.0 licence[8]. Any other data – especially data provided by SWISS – has exactly the same format as the synthetic data but usage rights are only with the respective provider of the data. For the course of the project these restricted datasets were made available at no cost.

[7] https://notary.ownyourdata.eu

[8] https://creativecommons.org/licenses/by-sa/4.0/deed.en

Co-funded by the European Union

4 Allocation of Resources

4.1 Estimated Cost

At the end of the project, a final price tag for the data management task during the project runtime is available: Containers storing data for the SlotMachine project are hosted on the Frequentis network and access to the data is covered by the project budget. Data access is guaranteed for the entire duration of the project and beyond at least 1 year after the end of the project.

Other services for Semantic Containers like notary and validation are provided by OwnYourData (https://www.ownyourdata.eu) at no charge. The notary service groups together write operations on a blockchain (a Merkle Tree summarizes all transactions of a day). As blockchain Ethereum is used (as of October 2022: Mainnet after the merge; also known as Ethereum 2.0) and fluctuations in gas prices are not relevant because of the grouping of operations.

4.2 Responsibilities

The consortium partner FREQUENTIS is responsible for implementing the DMP and ensures that it is reviewed and revised during the project runtime.

Name and contact details of the person responsible during the project runtime:

Christoph Fabianek
Innovationsstraße 1
1100 Vienna, Austria
christoph.fabianek@frequentis.com
+43 664 60850 2348

The project will only be responsible for storing, preserving, and backing up the datasets mentioned in chapter 2 of this document. Any other data created and managed by project partners in the course of development and testing SlotMachine will be the responsibility of the respective partner (i.e., data controller).

4.3 Long-Term Preservation

For long-term preservation of a dataset, it is recommended to “commit” a container and store the resulting Docker image either on a public repository (e.g., https://dockerhub.com) or in a private repository. See the Semantic Container Design document for a description on how to commit, backup, and restart a container.

For the SlotMachine project as long-term preservation all public data are made available as Docker images here: https://hub.docker.com/orgs/slma/repositories. The concrete images for each dataset are listed in Section 3.2.

EUROPEAN PARTNERSHIP
Co-funded by the European Union

5 Data Security

Storing data in a container provides only a very basic level of security through Authorization. It is up to the data controller to perform regular backup and setup a recovery strategy. For real world data – provided in confidence by project partners – this is taken care for at Frequentis through standard IT procedures. Please note that this data is not shared and only used in a closed environment within the SlotMachine project setup using Semantic Container.

Risks related to data security for datasets as described in Chapter 2 are not relevant for the Docker images made available on Dockerhub since all (synthetic) data is made publicly available as described in section 3.4 Licensing Data.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

6 Legal and Ethical Aspects

6.1 Data Protection

Whenever personal data is processed, the compliance with the principles of data protection are to be proven by the controller. These principles encompass, for instance, data minimisation, meaning to only process the data necessary for the pursued purpose. Privacy by design indicates to create data processing technically already in favour of strong protection of personal data.

Data generated and used for the SlotMachine project does not include any personal data.

Datasets generated from inviting experts to the Advisory Board are handled separately and are covered by the FREQUENTIS GDPR-compliant data management principles.

6.2 Measures to Ensure Ethical and Legal Standards

All measures are taken to ensure the use of legal and ethical unquestionable datasets in the course of the project as described in the initial project proposal.

6.3 Privacy and Trust

Issues of privacy and trust amongst data trading participants need to be identified by the data controller. The following list of indicators leading to negative levels of privacy and trust can be used as guideline:

- an inconsistent level of protection for natural persons and private data,
- divergences in the handling and storage of data hampering the free movement of personal data,
- a lack of knowledge regarding data sharing,
- difficulties in determining the trustworthiness of data suppliers,
- lack of knowledge of the law leading to potential violations,

and inconsistent levels of protection for members across participating organisations.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

7 References

[1] SlotMachine Grant Agreement Description of Action - GA-890456-SlotMachine

[2] SlotMachine - D1.1 Project Management Plan (PMP), v01.02.00

[3] SlotMachine - D1.3 Data Management Plan (DMP), v01.01.01

[4] SlotMachine - D1.6 Data Management Plan (DMP), v01.01.01

[5] SlotMachine – D2.1 Requirement Specification, v01.02.02

[6] SlotMachine – D2.2 System Design Document, v01.00.01

[7] SlotMachine – D3.3 Privacy Engine Software Component, v01.00.00

[8] SlotMachine – D4.1 Report on State-of-the-Art of Relevant Concepts, v01.01.01

[9] SlotMachine – D5.1 Platform Demonstrator, v01.01.01

[10] Project Handbook (Programme Execution Guidance) for ER projects - v03.00.00, dated 14th March 2018

[11] SPECIAL - D2.1 Policy Language V1, SPECIAL_D21_M12_V10.pdf (specialprivacy.eu)

EUROPEAN PARTNERSHIP

Co-funded by the European Union