Grant Agreement No.: 952697
Call: H2020-SU-ICT-2018-2020
Topic: SU-ICT-02-2020
Type of action: RIA

ASSURED

D8.2 DATA MANAGEMENT PLAN

VERSION 1.0

Work package	WP 8
Task	Task 8.3
Due date	28/02/2021
Submission date	01/03/2021
Deliverable lead	MARTEL
Version	1.0
Authors	Jean-Baptiste Milon (Martel)
Reviewers	Thanassis Giannetsos (UBITECH) Ellen Juel Nielsen (DTU)
Abstract	This deliverable provides the first version of the ASSURED Data Management Plan, identifying best practice and standards for the data collected and generated by and for the project, to assess suitability for sharing and reusing in accordance with official EC guidelines. The DMP ensures that all output is managed and maintained and that the data produced by the project is subject to appropriate levels of security, including the EU General Data Protection Regulation (GDPR).
Keywords	Data management plan, standards, metadata, policies, licensing

Document Revision History

Version	Date	Description of change	List of contributor(s)
V0.1	15/01/2021	1st full version of the deliverable for review	Jean-Baptiste Milon (Martel)
V 0.2	20/01/2021	Review comments	Athanassios Giannetsos (UBITECH)

V 0.3	26/02/2021	Addressed review comments	Jean-Baptiste Milon (Martel)
V 0.4	27/02/2021	Additional of technical inputs	Athanassios Giannetsos (UBITECH)
V 0.5	01/03/2021	Final Version	Jean-Baptiste Milon (Martel
V 0.6	01/03/2021	Final review and submission	Ellen Juel Nielsen (DTU)

DISCLAIMER

The information, documentation and figures available in this deliverable are written by the "Future Proofing of ICT Trust Chains: Sustainable Operational Assurance and Verification Remote Guards for Systems-of-Systems Security and Privacy" (ASSURED) project's consortium under EC grant agreement 952697 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

© 2020 - 2023 ASSURED Consortium

Project co-funded by the European Commission in the H2020 Programme		
Nature of the deliverable:	**to specify R, DEM, DEC, OTHER**	
Dissemination Level		
PU	Public, fully open, e.g. web	✔
CL	Classified, information as referred to in Commission Decision 2001/844/EC	
CO	Confidential to ASSURED project and Commission Services	

** R: Document, report (excluding the periodic and final reports)*

DEM: Demonstrator, pilot, prototype, plan designs

DEC: Websites, patents filing, press & media actions, videos, etc.

OTHER: Software, technical diagram, etc.

Funded by the Horizon 2020 Framework Programme of the European Union

EXECUTIVE SUMMARY

The Data Management Plan (DMP) is a living document that identifies which data is produced during the ASSURED project, who owns that data, how it will be documented, how it will be preserved, and with whom and under what form it will be shared. Most of the data produced during the project will correspond to documentation, source-code, mathematical proofs and experimental results. Most of the documentation that will be generated during the project has already been identified on the Description of Action (DoA).

The data generated during the project will be owned by the partners which have contributed to produce that data. The extent up to which this data will be made available and which restrictions will be imposed on its reuse will be decided on a case-by-case basis by the owners of the data. The partners will comply with the FAIR (findable, accessible, interoperable and reusable) guidelines of the H2020 programme, which state that data will be made as available as possible, so long that does not negatively affect the commercial advantage of the partners. The data will be shared among partners using internal repositories or through direct communication; and with the public through the project's website or public repositories. Finally, the data will be preserved from 3 and up to 30 years after the end of the project at the partners' repositories and cloud infrastructures, according to each partner internal policy.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020
Framework Programme of the European Union

TABLE OF CONTENTS

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

ABBREVIATIONS

DMP	Data Management Plan
DoA	Description of Action
TCG	Trusted Computing Group
ISO	International Organisation for Standardisation
IEC	International Electrotechnical Commission
FAIR	Findable, Accessible, Interoperable and Reusable
H2020	Horizon 2020
TSS	TPM Software Stack
ESAPI	Enhanced System API
SAPI	System API
TCTI	TPM Command Transmission Interface
API	Application Programming Interface
TLS	Transport Layer Security
SSL	Security Sockets Layer
QR	Quantum Resistant
VHDL	VHSIC Hardware Description Language
CSV	Comma Separated Value
JSON	Javascript Object Notation
PCI	Payment Card Industry
WP	Work Package
GNU	GNU is Not Unix
AGPL	Affero Generic Public Licence
CC	Creative Commons
FPGA	Field-Programmable Gate-Array
DOI	Digital Object Identifier
PQ	Post-Quantum
SME	Small or Medium-sized Enterprise
IMA	Integrity Measurement Architecture

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

1 INTRODUCTION

The current digital infrastructure prevents research and development from achieving their most productivity. While datasets resulting from research are often made publicly available, they: a) frequently go undocumented, b) are provided in non-standard formats, and c) are not uniquely identified, or are hardly accessible. This makes it hard to collect and compare results from several projects, in particular, using automated methods. It is thus important to take data management into account from the beginning of projects, so that these limitations are overcome. The main goal of this document is to identify the datasets that will be produced during the ASSURED project, research on which standards might be employed to potentiate their reusability and analyse up to which extent they might be disseminated.

More concretely, this plan describes the data management life cycle for all data sets that will be collected, processed or generated by the research project. It is a document outlining how research data will be handled both during the research project and after it is completed, describing what data will be collected, processed or generated and following what methodology and standards, whether and how this data will be shared and/or made open, and how it will be curated and preserved. It should be noted that the Data Management Plan (DMP) is not a fixed document: it evolves and gains more precision and substance during the lifespan of the project.

The remaining of the DMP is organised as follows. Chapter 2 includes global considerations about the type of datasets that will be generated during the project, which standards will be followed, how they will be documented, preserved and shared, among others. Chapter 3 will consider the generated datasets in more detail, providing more specific features for each, such as who owns each dataset, what it will be used for, and how it will be disseminated and stored. Chapter 4 covers Data Management specificities under the Horizon 2020 framework and chapter ASSURED's Participation to the ORDP. Finally, Chapter 6 concludes the Data Management Plan.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020
Framework Programme of the European Union

2 DATA DESCRIPTION

2.1 DATASETS

During the implementation and evaluation of the ASSURED, publicly available libraries will be used, in order to accelerate the development of proof-of-concept architectures and ensure its wide applicability.

The successful achievement of the project objectives relies on the production of documentation, proof-of-concept systems, performance measurements and scientific publications. The documentation that directly results from the project has been identified during the project planning and includes reports on the ASSURED use-case and system requirements, on the reference architecture, on the security and privacy risks in next-generation “Systems-of-Systems” deployments, on the novel lightweight cryptographic primitives, on the security models for the attestation and operational assurance protocols, on the threat modelling and risk assessment methodology, on the secure and trusted data sharing schemes by leveraging policy-compliant Blockchain structures, on the demonstrators implementation, on the validation results, performance evaluation and adoption guidelines, on the exploitation, standardization, dissemination and communication activities, on the project quality plan and on the risk assessment plan. These reports will be produced using standardized document types including Microsoft Word files and LateX descriptions.

The production of the other deliverables, such as demonstrators and libraries, will require the design of systems using source-code, possibly proving the security of systems with automated proof verifiers, and measuring/evaluating the performance of the implementations. Software systems will be described with languages that include *C*, *Java* and *Python*. Hardware systems will be described in *VHDL*. The code size is estimated to be in the thousands of lines of code. Security proofs will be done with functional programming languages, such as *Coq ProVerif, or TAMARIN*. Performance measurements, which might include timing and energy consumption, will be provided either in plaintext format or structured documents, like Microsoft Excel files or CSV. The data used by the demonstrators will include activity tracking data sets generated by dummy nodes and described in JSON.

The produced code for the attestation mechanisms will follow the TCG TPM 2.0 (ISO/IEC standard 11889-1:2015) standard as closely as possible. Changes to this standard will be introduced only when strictly required to ensure legacy requirements as well as alignment with other possible decentralized roots-of-trust to be considered such as the DICE architecture. In addition, UBITECH has been certified with ISO 27001. This certification provides confidence on the management of security of the data handled by UBITECH regarding the risk assessment process of the envisioned use cases.

The code will be documented through inline source code comments, and/or reference manuals. The documentation of the code may be done with standardized tools such as Doxygen.

Moreover, scientific publications, along with the accompanying presentation slides when applicable, will be generated to disseminate the results achieved by the project. Other dissemination documents include leaflets and posters. These documents will be produced with Microsoft Word/PowerPoint and LateX.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

2.2 INTELLECTUAL PROPERTY RIGHTS

The data generated throughout the project will be owned by the partners that have contributed to its creation. The licensing of the data will be agreed upon on a case-by-case manner by the partners involved in each WP, and this document will be updated accordingly.

Open-source libraries that will be used to support the implementations in this project will impose limitations on how the generated system may be distributed. For instance, the implementations of the attestation software stack impose the following restrictions:

"[…] Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. All advertising materials mentioning features or use of this software must display the following acknowledgement:

This product includes software developed by the <organization>.

4. Neither the name of the <organization> nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

Second, restrictions on usage might be put in place for code derived from privately owned libraries, such as SPH's CubeSat operating system.

2.3 ACCESS AND SHARING

The accessing and sharing of data is firstly ruled by two documents: the Consortium Agreement (CA), which stipulates under which conditions transmitted information between the project partners is deemed confidential and must not be further disseminated; and the Description of Action (DoA) which stipulates the dissemination level of each deliverable. Moreover, the project consortium will comply with the FAIR (findable, accessible, interoperable and reusable) (European Commission, 2016) guidelines of the H2020 programme.

The data necessary to successfully complete the project Work Packages (WPs) will be shared without any restrictions amongst the WP partners either via internal repositories or direct communication. Public data will be made available at the project's website, the DTU data repository (which provides the overall consortium shared space) or other repositories, as appropriate. Users will be made aware of this data primarily through research publications, patent applications, dissemination activities, invited talks, social networks, the project website (ASSURED Consortium, 2020), UBITECH's open-source repositories, SURREY and TUE website, the Research Portal from SURREY and the website of the TUDA research team (System Security Lab). Data will be made available to the project consortium as soon as it is available; to standardisation bodies when required; and to the public at the due date of the derivable, and, in case a research publication is based on that, as soon as the paper is submitted (if submission is anonymous, this will be postponed). If access to confidential data is necessary by the public, restrictive measures will be put in place.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

Finally, minting of Digital object identifiers (DOIs) is automatic for all deposits in the SURREY repository, providing unique identifiers for the documents made available therein.

2.4 ARCHIVING AND PRESERVATION

The project consortium will mostly keep the produced source-code, the accompanying documentation and the experimental results after the end of the project. These will be used for further research that push the state of the art in the core areas, targeted by ASSURED, of **remote attestation** (and underlying trusted computing technologies), **lightweight cryptography**, **dynamic real-time risk assessment**, and **enhanced and accountable knowledge sharing of operational threat intelligence data flows** (through the use of Blockchains).

The generated data will be stored by each partner for a minimum of 10 years beyond the end of the project. The costs associated with storing these data during the course of the project correspond to the server provisioning and maintenance and have been accounted for in the project overheads.

2.5 DOCUMENTATION & METADATA

Documentation will be provided in formats such as text files, source-code comments, man files, Word documents, Wiki and Markdown. It will be comprised of specification descriptions, code examples of API/library/executable invocations and testbeds.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

3 DATA COLLECTION

The following datasets have been identified as likely to being generated during the project.

3.1 REMOTE ATTESTATION ENABLERS

Data Owner

WP3 Leader (TUDA), WP3 Contributors

Dataset Description

New algorithms and systems for remote attestation will be developed and (partially) implemented as proof-of-concepts. These **runtime behavioural attestation services** will target both the software and hardware layers and will cover all phases of a device's execution; from the **trusted boot and integrity measurement of a CPS**, enabling the generation of static, boot-time or load-time evidence of the system's components correct configuration (Configuration Integrity Verification (CIV)), to the runtime behavioural attestation of those safety-critical components of a system providing strong guarantees on the correctness of the **control- and information-flow properties**, thus, enhancing the performance and scalability when composing secure systems from potentially insecure components.

End user

The remote attestation trust extensions will be made available to the scientific community.

Existence of similar data

There is a plethora of attestation protocols being present in the literature that, however, have a number of challenges and gaps when it comes to the efficiency and scalability of such solutions – especially when it comes to techniques that target the operational assurance of a deployed system. A detailed SoTa analysis will also be documented in the context of WP2.

Possibility of integration and reuse

These trust extensions may be used to support verifiable evidence towards the operational assurance of any type of embedded system.

Standards and metadata

Source-code of proof-of-concept implementations will be developed and will be disseminated as open-source. Furthermore, these extensions will also be presented to TCG as part of the project's standardization activities.

Data sharing

The attestation algorithms and primitives will be disseminated through scientific publications.

3.2 MULTI-LEVEL DETAILED RUNTIME TRACER

Data Owner

T3.4 Leader (MLNX), T3.4 Contributors

Dataset Description

New **tracing mechanisms** that will be used for extracting the **control- and data-flow graphs** as part of the ASURED Attestation Toolkit. This provides the **trusted anchor** with the compiled control- and information-flow graphs (CFGs & DFGs) that represent the **runtime state of a remote device, against the configuration and execution properties of safety-critical**

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

components to be verified. ASSURED advanced tracing techniques will be based on the novel use of: (i) lightweight eBPF execution hooks capable of providing near real-time low-level code inspection, thus, capturing the strict constraints of SoS-enabled ecosystems, and (ii) embedded OS introspection agents capable of traversing the entire physical memory of a CPS, via Direct Memory Access, for known execution signatures.

End user

The tracing extensions will be made available to the scientific community.

Existence of similar data

The goal is to target purely software-based runtime tracing capabilities, thus, existing libraries may be investigated in this direction including the “extended Berkeley Filters” (eBPFs) execution hooks and Intel PT tracing functionalities.

Possibility of integration and reuse

These tracing extensions may be used to support the monitoring and introspection of devices that are also outside the ones to be leveraged in the ASSURED use cases.

Standards and metadata

Source-code of proof-of-concept implementations will be developed and will be disseminated as open-source.

Data sharing

The tracing algorithms and primitives will be disseminated through scientific publications.

3.3 TSS2.0 EXTENSIONS FOR TRUST AND SECURITY BLOCK-CHAIN ENABLERS

Data owner

Task 4.4 Leader (SURREY), Task 4.4 contributors

Dataset description

The TSS is a software stack designed to isolate TPM application programmers from the low-level details of interfacing to the TPM. This stack will be extended to support ASSURED’s trusted on- and off-chain data management functionalities – i.e., stakeholder authentication, private ledger access control and ledger information verification.

End user

It is our goal that these extensions are proposed for standardisation.

Existence of similar data

Implementations of current versions of the API are available (e.g. IBM TSS2.0 implementation).

Possibility of integration and reuse

The extensions may be used to support future versions of the TPM.

Standards and metadata

It is our goal that the produced code follows the TCG TPM 2.0 (ISO/IEC standard 11889-1:2015) standard as closely as possible. Changes to this standard will be introduced only when strictly required in the context of decentralized trust anchors for DLT-based Identity Management.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

The code will be documented through inline source code comments, and/or reference manuals. The documentation of the code may be done with standardised tools such as Doxygen.

Data sharing

Should these extensions be accepted for standardisation, the specifications will be made available at the Trusted Computing Group's website.

Archiving and preservation

The source-code will be preserved for 3-30 years at the partners private and public repositories.

3.4 TPM2.0 EXTENSIONS

Data owner

Task 4.4 Leader (SURREY), Task 4.4 contributors

Dataset description

Extensions to the TPM2.0 library specifications will be provided, supporting more efficient and trustworthy verification of block data, mining validation, consensus management and agreement, membership authentication, and undeniable actions commitment.

End user

It is our goal that these extensions are proposed for standardisation.

Existence of similar data

The TPM2.0 standard is publicly available. Implementations of the current version of the architecture are available (e.g. IBM TPM2.0 SW implementation).

Possibility of integration and reuse

The extensions may be used to support future versions of the TPM.

Data sharing

Should these extensions be accepted for standardisation, the specifications will be made available at the Trusted Computing Group's website.

3.5 DICE SOFTWARE ABSTRACTION

Data owner

Task 4.4 Leader (SURREY), Task 4.4 contributors (MLNX)

Dataset description

New set of **designed and implemented software abstractions** – for the new generation of **DICE roots-of-trust** – that will provide the necessary interfaces to support a **256 bit seed** (used for the internal key generation process). Furthermore, abstractions will be developed, based on the integration of **symmetric (ARM ISA) and asymmetric engines and primitives**, for the provision of an additional software layer that will be able to support a more efficient variant of the **Direct Anonymous Attestation (DAA)** protocol with the DICE engine as the underlying trusted component.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

End user

The DICE updated models and abstractions may be described in a scientific publication.

Existence of similar data

Partially for other types of roots-of-trust including the trusted Platform Module (TPM).

Possibility of integration and reuse

The software abstractions may be reused in other security chips.

Standards and metadata

The code will be documented through inline source code comments, and/or reference manuals. The documentation of the code may be done with standardised tools such as Doxygen or similar.

Data sharing

The source-code will be kept private.

Archiving and preservation

The source-code will be preserved for 3-30 years.

3.6 CODE FOR THE RISK ASSESSMENT FUNCTIONALITIES OF THE ASSURED PLATFORM AND VULNERABILITY ANALYSIS & ATTACK VALIDATION COMPONENT

Risk Management is the service that offers monitoring, calculation, and evaluation of risks related to the security and privacy threats of each component (i.e., device), in the target "Systems-of-Systems" environment, to be protected by the ASSURED Attestation Toolkit (Section 3.1) and Cryptographic Primitives (Section 3.2).

For the risk evaluation, a meta-model will be defined that will include assets such as deployed devices (comprising the mixed-criticality services of the target supply chain) and datasets with their properties and their relationships and dependencies. Security and privacy calculation will be related to the ISO and GDPR definitions, respectively, and will also take into account the possible risks that are related with the device configuration and execution behavioural properties so as to identify the best set of mitigation strategies (i.e., security attestation policies) to be enforced.

Data Owner

WP2 Leader (UBITECH), WP2 Contributors

Dataset Description

A reactive run-time Risk Assessment (RA) and mitigation framework will be developed to ensure the security of the envisioned use cases in the face of emerging threats and vulnerabilities. This RA tool will also be enhanced with a collective threat intelligence engine for emulating and simulating a wide gamut of attack vectors so as to analyse their potential impact to the entire cartography of assets.

End user

The risk assessment and vulnerability analysis will be made available to users of the ASSURED platform.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

Existence of similar data

Partially for other types of hardware devices.

Possibility of integration and reuse

The framework will be developed to be as generic as possible to be easily adaptable to other devices.

Standards and metadata

The code will be documented through inline source code comments, and/or reference manuals. The documentation of the code may be done with standardised tools such as Doxygen.

Data sharing

It hasn't yet been agreed up to what extent and under which licence the code will be made available. However, the outputs of the risk assessment models to the envisioned use cases will be kept confidential.

Archiving and preservation

The outputs of the risk assessment tool will be preserved for 3-30 years at the partner's private repositories.

3.7 DIGITAL TWINS FOR SUPPLY CHAINS

Data Owner

T2.5 Leader (UNIS), T2.5 Participants (BIBA, UBITECH)

Dataset Description

Development of a scalable testbed (based on the concept of Digital Twins), using a **hybrid approach of emulation and simulation, to generate a digital representation of the deployed CPSoS, safety-critical components and services**, taking also into consideration their **real-time constraints**, towards emulating the overall SoS ecosystem and simulating the identified risk and threats for evaluating the compiled attestation policies in regards to CPS **attack resilience, fault tolerance and prediction of any unforeseen cyber threats.**

End user

The testbed will be made available to users of the ASSURED platform.

Existence of similar data

Partially for other types of environments.

Possibility of integration and reuse

The framework will be developed to be as generic as possible to be easily adaptable to other application domains.

Standards and metadata

The code will be documented through inline source code comments, and/or reference manuals. The documentation of the code may be done with standardised tools such as Doxygen.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

Data sharing

It hasn't yet been agreed up to what extent and under which licence the code will be made available. However, the outputs of the attack validation models to the envisioned use cases will be kept confidential.

Archiving and preservation

The outputs of the risk assessment tool will be preserved for 3-30 years at the partner's private repositories.

3.8 CONSTRAINT LEARNING AND POLICY ENFORCEMENT

Data Owner

UTRCI as T2.5 Participant

Dataset Description

Development of a **model-based design workflow for safety and security assurance in mixed-criticality applications executed on multicore platforms.** This will enable the **safe configuration and implementation** of those software components with a high level of criticality to be developed on the same CPS with shared software and hardware resources. More specifically, this model-based approach will leverage existing tools (i.e., Simulink) for solving optimization problems regarding the **best resource and time partitioning** at the available multicore resources. This will also enable for the more efficient process of the ASSURED remote attestation mechanism.

End user

This constraint problem solver will be made available to users of the ASSURED platform so that they can adopt it for identifying the best configuration of their internal devices.

Existence of similar data

Partially for other types safety properties but not tailored to security and operational assurance as is the vision of ASSURED.

Possibility of integration and reuse

The framework will be developed to be as generic as possible to be easily adaptable to other devices.

Standards and metadata

The code will be documented through inline source code comments, and/or reference manuals. The documentation of the code may be done with standardised tools such as Doxygen.

Data sharing

All models will be disseminated through scientific publications.

Archiving and preservation

The outputs of the risk assessment tool will be preserved for 3-30 years at the partner's private repositories.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

3.9 EXPERIMENTAL MEASUREMENTS

Data owner

WP5 Leader (INTRA), WP5 contributors

Dataset description

The performance of demonstrators will be evaluated by measurements that might include timing and power consumption.

End user

Experimental results pertaining to timing and energy consumption may be included in scientific publications.

Existence of similar data

Performance metrics are typically included in scientific publications wherein attestation and/or cryptography is addressed.

Possibility of integration and reuse

The results may be used in other publications for performance comparisons.

Standards and metadata

Experimental results will be published under the form of an Excel or a CSV file.

Archiving and preservation

The experimental results will be preserved for 3-30 years at the partners private and public repositories

3.10 HRC DATASET

Data owner

BIBA

Dataset description

The Human-Robot-Collaboration (HRC) Data consist of lines of activities/events which are used in the HRC Monitoring infrastructure, and these, if combined can produce analytics. These data include elements such as “Type”, “Location”, “Event”, etc.

End user

The data-set will be used for the HRC demonstrator and could be used by SMEs and Research Organisations for their own experiments. Scientific Publications based on the demonstrator may be pursued.

Possibility of integration and reuse

The same data-set could be reused for experiments with a similar structure.

Standards and metadata

The format to be used is JSON as it is used in the employed storage facility (MongoDB).

Data sharing

The data will be licensed under GNU AGPLv3.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

Archiving and preservation

The dataset will be preserved for 3-30 years at the partners public repositories

3.11 (SYSTEM) REQUIREMENTS, USE CASES (SYSTEM) ARCHITECTURE

Data owner

WP1 Leader (DTU), WP1 contributors

Dataset description

These documents describe the requirements of the ASSURED framework and the use-case system architecture and will establish concrete goals for the remainder of the project.

End user

The ASSURED documentation will be made available to the public so that other groups may contribute to the project or take advantage of it.

Possibility of integration and reuse

The system requirements and architecture may be used as a basis for other projects on security, privacy and operational assurance services in various application domains.

Standards and metadata

These documents will be produced with Microsoft Word or LateX.

Data sharing

The documentation will be made publicly available.

Archiving and preservation

The documentation will be preserved for 3-30 years at the partners public repositories

3.12 EFFICIENT CRYPTOGRAPHIC ALGORITHMS AND PRIMITIVES

Data owner

WP4 Leader (TUE), WP4 contributors

Dataset description

Novel algorithms and primitives will be proposed to achieve the performance requirements of the security services (i.e., key establishment, secure communication, signature of collective attestation reports, Attribute-based Encryption, Searchable Encryption, Proxy Re-encryption, etc.) while targeting the resource-constrained devices comprising the supply chains environments.

End user

The proposed cryptographic primitives should be reviewed by the scientific community and disseminated through entities that might benefit from exploiting it.

Existence of similar data

There's no similar available data for the considered security models. No official standards exist for such data. We will follow the de-facto standard of submitting to eBACS for benchmarking.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

Possibility of integration and reuse

The proposed algorithms may be integrated in other security protocols.

Standards and metadata

Proofs of security of new cryptographic protocols might be provided in computer verifiable languages.

Data sharing

The algorithms and primitives will be disseminated through scientific publications. ASSURED may produce or reuse existing source code for implementing a PoC (Proof of concept) and/or production level deliverables as well as benchmark certain properties (e.g., timing, size) of such code. Textual format (for the source code) and tabular format (for the benchmarks) are the most natural way to store such information and organic methods to store, backup, and access such information are widely known and deployed.

Archiving and preservation

The proofs will be preserved for 3-30 years at the partners public repositories

3.13 CODE FOR THE PROVABLE SECURITY MODELLING AND ANALYSIS OF THE ASSURED FRAMEWORK

Data owner

T3.3 Leader (DTU), T3.3 Contributors

Dataset description

The consortium aims at proving the security of the ASSURED framework in a form as wide as possible.

End user

The proofs will be made available to the scientific community.

Existence of similar data

There's no similar available data for various decentralized roots-of-trust that will be considered in the context of ASSURED Attestation Toolkit (e.g., DICE). There are a number of works focusing on the cryptographic primitives leveraged in such crypto accelerators; however, the focus in ASSURED would be to focus on "idealized functionalities" that go beyond the crypto aspects that in our case are considered secure.

Possibility of integration and reuse

Other researchers might build upon the proofs.

Standards and metadata

Proofs of security of new cryptographic protocols might be provided in computer verifiable languages.

Data sharing

The proofs will be made publicly available.

Archiving and preservation

The proofs will be preserved for 3-30 years at the partners public repositories

Partially for other types of hardware devices.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

3.14 SECURITY EVALUATION AND ASSESSMENT

Data owner

Task 6.6 Leader (S5), Task 6.6 Contributors

Dataset description

This document will provide guidelines for the implementations of the ASSURED taking into consideration the new attestation primitives and how they relate to low-level software attacks including buffer overflows and race conditions, as well as hardware attacks such as a fault injections and cold boot attacks.

End user

This document will provide greater confidence on the security and privacy by the ASSURED framework both to researcher partners (including universities and research organisations) and commercial partners; and provide important guidelines for the further adoption of the ASSURED architecture.

Existence of similar data

Companies such as Riscure and River Loop Security perform similar security analysis.

Possibility of integration and reuse

The security analysis may be integrated on projects that make use of trusted computing technologies to improve their security.

Standards and metadata

The document might be accompanied with experimental results that practically demonstrate how the proposed guidelines mitigate low-level attacks, such as power-traces. These data might be provided in standard data formats that can be readily imported to frameworks dealing with these types of attacks, such as Jlsca (https://github.com/Riscure/Jlsca).

Data sharing

This document will be made publicly available.

Archiving and preservation

The document will be preserved for 3-30 years at the partners public repositories

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020
Framework Programme of the European Union

4 DATA MANAGEMENT IN HORIZON 2020

According to the European Commission (EC) all project proposals submitted to "Research and Innovation actions", "Innovation actions" and “Coordination support actions” have to include a section on research data management which is evaluated under the criterion 'Impact'. Projects participating in the pilot action on open access to research data have to develop a DMP) to specify what data will be open.[1]

The DMP is defined as:

“*Data Management Plans (DMPs) are a key element of good data management. A DMP describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project. The use of a Data Management Plan is required for projects participating in the Open Research Data Pilot. Other projects are invited to submit a Data Management Plan if relevant for their planned research*.”

The purpose of a DMP is to provide a discussion of the main elements of the data management policy that will be used by the applicants with regard to all the datasets that will be generated by the project.

TABLE : 1 CLARIFICATIONS OF TERMS

Research data	Research data is the evidence that underpins all research conclusions (except those which are purely theoretical) and includes data that have been collected, observed, generated, created or obtained from commercial, government or other sources, for subsequent analysis and synthesis to produce original research results. These results are then used to produce research papers and submitted for publication.
Open research data	Openly accessible research data can typically be accessed, mined, exploited, reproduced and disseminated, free of charge for the user.
Secondary data	Secondary data are data that already exist, regardless of the research to be conducted.
Open access	Open access is understood as the principle that research data should be accessible to relevant users, on equal terms, and at the lowest possible cost. Access should be easy, user-friendly and, if possible, Internet-based.
Metadata	Metadata is data used to describe other data. It summarizes basic information about data, which can make finding and working with instances of data easier.

[1]http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

Research data repositories	Research data repositories are online archives for research data. They can be subject based/thematic, institutional or centralized.

Overall, having taken into account all relevant principles regarding lawful processing of personal data, scientific research data should be easily discoverable, accessible, assessable and intelligible, useable beyond the original purpose for which it was collected and interoperable to specific quality standards.

The ASSURED Data Management also follows the Guidelines on FAIR Data Management in Horizon 2020, released by the European Commission Directorate – General for Research & Innovation. This Horizon 2020 FAIR DMP template[2] has been designed to be applicable to any Horizon 2020 project that produces, collects or processes research data. According to these guidelines the management and organization of data should be based on four basic principles, which determine how research outputs should be processed so that they can be more easily accessed, understood, exchanged and reused. This means that data must be findable, accessible, interoperable and re-useable, for example by researchers interested in using the data in further research in the field.

These principles precede implementation choices and do not necessarily suggest any specific technology, standard, or implementation-solution. EC provides a Template with the FAIR principle. This template is not intended as a strict technical implementation of the FAIR principles, it is rather inspired by FAIR as a general concept. The template represents the set of questions that someone should answer with a level of detail appropriate to the project.

It is possible to develop a single DMP for any project to cover overall approach. However, where there are specific issues for individual datasets (e.g. regarding openness), someone should clearly spell this out.

The template proposes the following issues to be addressed:

- Data Summary
- FAIR data
- Allocation of resources
- Data security
- Ethical aspects

[2] H2020 Programme Guidelines on FAIR Data Management in Horizon 2020 Version 3.0, 26 July 2016, available at:

http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

5 ASSURED ORDP PARTICIPATION

The Open Research Data Pilot (ORDP) of the European Commission enables open access and reuse of research data generated by Horizon 2020 projects. There are two main pillars to the Pilot: a) developing a Data Management Plan (DMP) and b) providing open access to research data.

A project that opts-in ORDP have to adhere to the following conditions:

- Develop (and keep up-to-date) DMP.
- Deposit the data in a research data repository.
- Ensure third parties can freely access, mine, exploit, reproduce and disseminate this data.
- Provide related information and identify (or provide) the tools needed to use the raw data to validate the research.

The ORDP applies to:

- The data (and metadata) needed to validate results in scientific publications.
- Other curated and/or raw data (and metadata) that are specified in the DMP.

From the current consensus within the consortium some of the ASSURED Artefacts will not be publicly available.

5.1 PUBLISHING INFRASTRUCTURE FOR OPEN ACCESS

The ASSURED publication infrastructure consists of a process and several web-based publication platforms that together provide long-term open access to all publishable, generated or collected results of the project. The implementation of the project will be done in accordance with the applicable regulations in national and EU level and, especially, with the General Data Protection Regulation (GDPR)[3] protection of personal data.

More specifically, there are not cases where personal data information or sensitive information of internet users is collected (IP addresses, email addresses or other personal information) or further processed.

In the potential future case where the ASSURED consortium will collect and/or further process personal data, this will be done in accordance with GPDR. Overall, it is aimed that ASSURED only collects and/or further processes personal data are necessary for the attainment of the project objectives.

Both the process and the used web-based platforms are described in the following subsections.

[3] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

5.1.1 Publishing Process

The ASSURED partners defined a simple, deterministic process that decides if a result in ASSURED must be published or not. The term result is used for all kind of artefacts generated during ASSURED like white papers, scientific publications, and anonymous usage data. By following this process, each result is either classified public or non-public. Public means that the result must be published under the open access policy. Non-public means that it must not be published.

For each result generated or collected during ASSURED runtime, the following questions must be answered to classify it:

Does a result provide significant value to others or is it necessary to understand a scientific conclusion?
If this question is answered with yes, then the result is classified as public. If this question is answered with no, the result is classified as non-public. Such a result could be code that is very specific to the ASSURED platform (e.g., a database initialization) which is usually of no scientific interest to anyone, nor does it add any significant contribution.

Does a result include personal information that is not the author's name?
If this question is answered with yes, the result is classified as non-public. Personal information beyond the name must be removed if it should be published. This also bears witness on the repetitive nature of the publishing process, where results which are deemed in the beginning as non-publishable can become publishable once privacy-related information is removed from them.

Does a result allow the identification of individuals even without the name?
If this question is answered with yes, the result is classified as non-public. Sometimes data inference can be used to superimpose different user data and reveal indirectly a single user's identity. As such, in order to make a result publishable, the included information must be reduced to a level where single individuals cannot be identified. This can be performed by using established anonymization techniques to conceal a single user's identity, e.g., abstraction, dummy users, or non-intersecting features.

Does a result include business or trade secrets of one or more partners of ASSURED?
If this question is answered with yes, the result is classified as non-public, except if the opposite is explicitly stated by the involved partners. Business or trade secrets need to be removed in accordance to all partners' requirements before it can be published.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

Does a result name technology that is part of an ongoing, project-related patent application?
If this question is answered with yes, then the result is classified as non-public. Of course, results can be published after patent has been filed.

Can a result be abused for a purpose that is undesired by society in general or contradict with societal norms and ASSURED's ethics?
If this question is answered with yes, the result is classified as non-public.

Does a result break national security interests for any project partner?
If this question is answered with yes, the result is classified as non-public.

5.1.2 Publishing Platforms

In ASSURED, we use several platforms to publish our results openly. The following list presents the platforms used during the project and describes their concepts for publishing, storage, and backup.

The project Website

The partners in the project consortium decided early to setup a project-related website. This website describes the mission and the general approach of ASSURED and its development status. A blog informs about news on a regular basis. Later in the project the developed ASSURED platform will be announced. A dedicated area for downloads is used to publish reports and white papers as well as scientific publications (in pre-camera ready form, or through links to the publisher's websites in case these are not open access). All documents are published using the portable document format (PDF)[4]. All downloads are enriched by using simple metadata information, such as the title and the type of the document. The website is hosted by partner MARTEL. All webpage-related data is backed up on a regular basis. All information on the project website can be accessed without creating an account. The website is backed up once per month.

GitLab

GitLab is a well-established online repository which supports distributed source code development, management, and revision control. It is primarily used for source code data. It enables world-wide collaboration between developers and provides also some facilities to work on documentation and to track issues. GitLab provides paid and free service plans. Free service plans can have any number of public, open-access repositories with unlimited collaborators. Private, non-public repositories require a paid service plan. Many open-source

[4] Note that the site will not host spreadsheets. It exclusively host PDFs

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

projects use GitLab to share their results for free. The platform uses metadata like contributors' nicknames, keywords, time, and data file types to structure the projects and their results. The terms of service state that no intellectual property rights are claimed by GitLab over provided material. For textual metadata items, English is preferred.

All source-code components that are implemented during this project and decided to be public will be uploaded to an open access GitLab repository.

5.1.3 Access and Sharing

The accessing and sharing of data is firstly ruled by two documents: The Consortium Agreement (CA), which stipulates under which conditions transmitted information between the project partners is deemed confidential and must not be further disseminated; and the Description of Action (DoA) which stipulates the dissemination level of each deliverable. Moreover, the project consortium will comply with the FAIR (findable, accessible, interoperable and reusable) (European Commission, 2016) guidelines of the H2020 programme.

The data necessary to successfully complete the project Work Packages (WPs) will be shared without any restrictions amongst the WP partners either via internal repositories or direct communication. Public data will be made available at the project's website or other repositories, as appropriate. Users will be made aware of this data primarily through research publications, patent applications, dissemination activities, invited talks, social networks and the project website. Data will be made available to the project consortium as soon as it is available; to standardization bodies when required; and to the public at the due date of the derivable, and, in case a research publication is based on that, as soon as the paper is submitted (if submission is anonymous, this will be postponed). If access to confidential data is necessary by the public, restrictive measures will be put in place.

Each of the previously defined (data and/or research) artefacts has its own set of questions that has to be addressed. The proposed template states that it is not required to provide detailed answers to all the questions of the DMP that needs to be submitted by month 6 of the project, subject -also- to potential future updates. Rather, the DMP is intended to be a living document -to the extent necessary- 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.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union

6 CONCLUSIONS

This document reflects the vision of the ASSURED consortium for the generation and handling of data as of February 1st, 2021. It is foreseen that most of the generated data will be related to the source-code and the data necessary to support the demonstrators, the security proofs associated with the novel cryptographic protocols, the documentation, and the experimental results. Efforts will be put in place, such as the adherence of the developed code as much as possible with the TCG standard and the respective working groups (depending on the adopted Root-of-Trust; TPM 2.0 standard or the DICE specification), so that the ASSURED security and assurance mechanisms can be generic enough to be able to execute on a wide gamut of heterogeneous devices comprising mixed-criticality service graph chains. In addition, other generated datasets, like experimental results, will follow standardised formats, and will be documented, so that they can be easily used as reference in other research projects. The datasets will be disseminated through research publications, patent applications, invited talks, among others, and will be preserved at the partners repositories for at least 3 years.

It should be noted that since it is very early in the project, this document only presents preliminary proposals in terms of sharing, volume and archiving. The DMP will be updated periodically during the project to reflect changes in the properties of the data that may be made available by the project, and to add more concrete information about the datasets.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020
Framework Programme of the European Union

REFERENCES

[1] Bos, J., Costello, C., Ducas, L., Mironov, I., Naehrig, M., Nikolaenko, V., … Stebila, D. (2016). Frodo: Take off the ring! Practical, Quantum-Secure Key Exchange from LWE.

[2] Chisnall, D. (2007). *The Definitive Guide to the Xen Hypervisor* (First). Upper Saddle River, NJ, USA: Prentice Hall Press.

[3] Costello, C., Jao, D., Longa, P., Naehrig, M., Renes, J., & Urbanik, D. (2016). Efficient compression of SIDH public keys.

[4] European Commission. (2016). *Guidelines on FAIR Data Management in Horizon 2020*. Retrieved from http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[5] ASSURED Consortium. (2018). ASSURED. Retrieved May 22, 2018, from https://ASSURED.eu/

[6] Goldman, K. (n.d.-a). IBM's Software TPM 2.0 download | SourceForge.net. Retrieved May 22, 2018, from https://sourceforge.net/projects/ibmswtpm2/

[7] Goldman, K. (n.d.-b). IBM's TPM 2.0 TSS download | SourceForge.net. Retrieved May 22, 2018, from https://sourceforge.net/projects/ibmtpm20tss/

[8] Love, R. (2010). *Linux Kernel Development* (3rd ed.). Addison-Wesley Professional.

[9] Shaw, A., Jacquin, L., Lioy, A., Pitscheider, C., Basile, C., Risso, F., … Bosco, F. (2015). *Specification of the SECURED architecture (beta version)*. Retrieved from https://www.secured-fp7.eu/files/secured_d232_arch_beta_v0101.pdf

[10] The OpenSSL Project. (2015). {OpenSSL}: Cryptography and {SSL/TLS} Toolkit.

© 2020-2023 ASSURED Consortium

Funded by the Horizon 2020 Framework Programme of the European Union