ENGINE

ZERO-DEFECT MANUFACTURING FOR
GREEN TRANSITION IN EUROPE

HORIZON-CL4-2021-TWIN-TRANSITION-01-02:

Zero-defect manufacturing for green transition in Europe

Grant Agreement n° 101058179

D9.3: Data Management Plan

Lead Author: Anssi Laukkanen/VTT

Tero Frondelius/WAR

José J. de las Heras/ADV

Tommi Helander/NOME

with contributions from: Juho Könnö/UOULU

Andris Freimanis/VTT

Reijo Kouhia/TAU

Friedrich Halstenberg/GD

Reviewed by [Names and organisations]

Tom Andersson/VTT

Deliverable nature	R -- Document, report
Dissemination level	PU - Public
Contractual delivery date	M6
Actual delivery	M6
version	1.0

Co-funded by
the European Union

Version history

Version	Date	Editors	Description
0.1	2.11.2022	ANL	First version of the document
0.2	22.11.2022	Several	Updated version of the document
1.0	30.11.2022	ANL	Submission version

Abstract

The Data Management Plan (DMP) is the key element for appropriate data management i.e. producing, collecting or processing of research data. ENGINE – being a project highly centered around data and design workflows – will collect extensive volumes of several types of data. Therefore, we follow the Horizon Europe Guidelines on FAIR data management to ensure accessibility, interoperability, and reusability of data within ENGINE. The DMP describes the data management life cycle for the data collected, processed and/or generated in ENGINE, and implemented via the ENGINE exchange as an integral part of the ENGINE system.

The DMP will be updated over the course of the project whenever significant changes arise, for example, when new data types and formats are being introduced. The next update or evaluation of the DMP will be done latest with the 1st periodic evaluation/assessment of the project.

Degree of progresses

This document complies 100% with the expected degree of progress.

Dissemination level

The following Deliverable in ENGINE has a PUBLIC Dissemination level.

List of abbreviations

CC	Creative Commons
CERN	European Organization for Nuclear Research
CSL	Citation Style Language
DMP	Data Management Plan
DNI	Direct Normal Irradiance
DOI	Digital Object Identifier
EC	European Commission
EEA	European Economic Area (Iceland, Norway and Switzerland)
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
JSON	JavaScript Object Notation
SSL/TLS	Secure Sockets Layer/ Transport Layer Security

Co-funded by
the European Union

Table of contents

Introduction

Modern product development and manufacturing produces vast volumes of data from simulations, tests, monitoring, and diagnostics. Moreover, zero-defect manufacturing systems use a wide range of tools to prevent and reduce defects, which leads to a heterogeneous data pool. Importantly, all tools must interact seamlessly at all manufacturing locations, therefore, must be able to "talk" to each other, if we are to create a competitive solution that industry partners want to use after ENGINE. All these reasons make effective data management paramount to develop a sustainable solution.

ENGINE has three modules, which are integrated into ENGINE system. Each defines their own specifications, which include data topics. ENGINE system architecture ensures that modules work together seamlessly while remaining independent. Data specifications identify industrial systems that ENGINE exchange needs to interface with and ensure data-interoperability in ENGINE. Software specifications include a plan for the whole development process, and specify capabilities in ENGINE toolbox and software languages, standards, and external libraries used for development. Hardware specifications clarify sensors developed for ENGINE exchange, monitoring techniques investigated, and requirements for data fusion.

This report gives details about data collected and generated by ENGINE. Especially, the report provides the general framework for data management: (1) meta-data collection and publication, (2) a summary of what type of data will be produced, (3) naming conventions, (4) storage and security of data, (5) allocation of resources, (6) ethical aspects, as well as data access issues e.g. what data can be shared with the public, and what's reserved for the consortium members and the European Commission.

This deliverable is based on the template provided by the European Commission. As recommended, the DMP will be updated over the project and particular in time with the periodic evaluation/assessment of the project. The final version will summarize all the data during the whole duration of the project.

Data Summary

Purpose of data collection and generation

ENGINE collects data to enable operation of the ENGINE system, which combines the data, software, and hardware modules to enable zero-defect manufacturing. The primary objective is for all ENGINE components to function with maximum independence, while enabling collaboration on a supply-chain scale. The ENGINE system and its modules are outlined in Figure 1: from the base system and workflows to data management in ENGINE exchange, simulation solutions in ENGINE toolbox and NDE, monitoring and production control in ENGINE production.

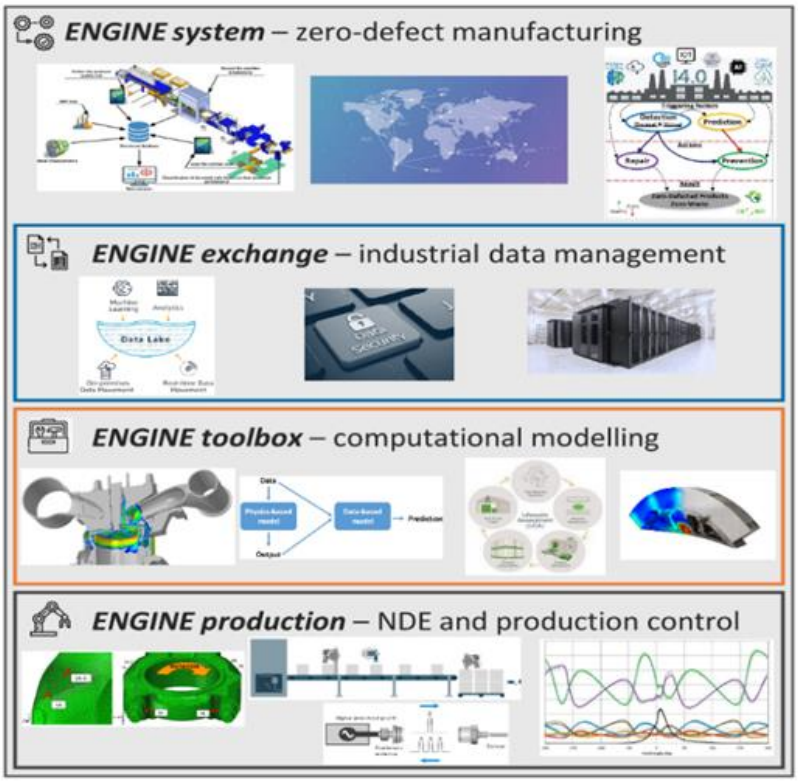

Figure 1. The ENGINE system and primary modules.

Data tools (ENGINE exchange) consist of cloud infrastructure, data ingestion layer, and data security layer. ENGINE exchange has four goals:

- Streamline exchange of open data within the supply-chain e.g. material specifications are not sent as excel files in an email, but available in ENGINE exchange for specific material batches.
- Enable using proprietary information within the supply-chain without actually sharing or revealing it (black-box approach) e.g. how would a change in steel composition affect expected lifetime without revealing the required change in processing parameters.
- Enable real-time decision-making during monitoring by combining simulation results with measurements during diagnostics e.g. monitoring system compares the obtained diagnostic results with simulations and recommends a decision (accept / further inspection / reject).
- Store the generated data for future use e.g. data from manufacturing for more refined LCA models, and computational model validation.

Types and formats of data collected

Regarding the main demonstrator in WP1 the following chapter specifies connecting rod main features, production process, and identified manufacturing / operation processes and data produced therein. The initial identified data sources and types are listed in Table 1, this data presented also in ENGINE deliverable D1.1 on specifications.

Table 1. Inputs and outputs of operation processes and sources of data.

Manufacturing / Operation Process (e.g. casting, usage, etc.)	Inputs	Physical outputs	Output data
Continuous casting	Scrap chemistry, steelmaking process parameters etc.	Solid steel	Data for LCA Material characteristics*
Continuous casting monitoring data			Data on steel chemistry and defects
Casting and rolling simulation	Process parameters, steel composition		Defect distribution, Reduction of cross-section
Rolling	Solid steel	Rolled bars	Data for LCA Material characteristics
Bar inspection	Rolled bars		Defect size
Die forging	Rolled bars	Raw forging	Data for LCA Material characteristics
Forging and machining simulation	Material defect distribution, forging parameters		Grain flow, defect distribution, material strength
Heat treatment	Raw forging	Heat treated forging	Data for LCA Material characteristics
Machining	Heat treated forging	Machined part	Data for LCA Surface roughness
UT inspection	Machined part		3D map of defect locations (point cloud)
UT data prefiltering	3D map of defect locations (point cloud)		Filtered data (Much smaller than raw. Threshold for inclusion size, and direction of defects. Throw the data away when inclusion below a certain size)
Decision making	Simulated defect maps + test data		Decision: keep/alert/throw away
Component simulation	Dynamic loads, component grain flow, strength data		Critical defect map
Assembly to engine	Component (data matrix QR code)	Assembled engine	Critical defect map
Engine operating conditions analysis			Component location (site data)
Full engine simulation	Engine loading		Engine loading data
Engine testing and usage	Dynamic moment on crankshaft &		Dynamic moment on crankshaft & combustion

	combustion force		force Online monitoring data (separately for testing and operation)
Engine maintenance	Defect map, operational history	Serviced/repaired engine	Data for LCA Updated defect map
Engine end of life	Defect map, operational history	refurbished/repaired/ scrapped engine	Data for LCA End of life defect and operational data
Fatigue simulation	Results of forging and machining simulation		Assessment of criticality of defect map, lifetime estimates

*"material characteristics" is used to describe the multitude of material features in general, such as microstructure, chemistry, residual stresses etc data and processed information stored over the workflow.

Data types

In addition to the data collected and generated by ENGINE, some of the data can be classified as personal data and therefore will be kept confidential. Therefore, the main categories of data collected and presented in Table 1 are twofold:

- Technology-specific data:
  - Data, for example, on materials, products, manufacturing conditions, characterization and measurement, production data and conditions, modeling and analysis, samples, and assemblies.
- Personal data:
  - Contact information (i.e. names, phone, e-mail addresses) of project partner representatives
  - Project external individuals and stakeholders (i.e.names, phone, e-mail addresses, affiliation).

Data formats

Different formats of datasets are collected and generated within ENGINE. The main data formats include:

- Audio files: .mp3, .wav, .wma, .ra
- Databases: .csv, various .sql like databases, hdf5 databases
- Specific modeling associated files: vtk, ivtk,
- Documents/Reports/Publications: .PDF/A, txt, doc/docx
- Pictures: .jpg, .png, .tif
- Spreadsheets: .xls/.xslx
- Video: .avi, mov, mp4, wmv

Data size

The size of different data formats differs greatly across various characterization, experimental and modeling formats. It is expected to have data ranges between 1 MB and 100 TB.

Existing data being re-used

At present this consists of public datasets from the Materials Project, NOMAD, OQMD, JARVIS repositories.

Origin of the data

ENGINE will involve both manually and automatically collected data:

- Manually collected data: literature review and open data (re-use of existing data)
- Automatically collected data:
  - data that will be generated through the testing and simulation of materials
  - data that will be generated during experimental and processing activities.
- Several research datasets will be re-used in performing ENGINE activities.
- Material codes and standards
- Material datasheets

Data utility: usefulness

The audience of DMP is mainly the ENGINE consortium and the European Commission. Concerning the project, the research data could be useful for several stakeholders:

- Project members: Internal use to perform the project activities.
- Other stakeholders, including sector associations, scientific community, service providers, policy makers and industry.

FAIR data

The FAIR data management (Findable, Accessible, Interoperable and Reusable data) principles will be ensured by ENGINE. While some closed data will be protected and thus not publicly available, the project will maximize access to and re-use of research data generated by the project.

Making data findable, including provisions for metadata

Quality control measures will be taken to maintain the accuracy of data during the project. Discipline compliant metadata elements will be used to describe the data, aid data discovery, and re-use. List of metadata elements and metadata standards used are provided in a separate spreadsheet. Metadata of opened data will be made available via FAIR compliant repository for research and re-use after project closure. Persistent identifiers provided by the repository will be used in linking to datasets.

ENGINE makes data accessible in three ways: public dissemination, internal data management solution, and sharing of datasets on the Zenodo repository. In addition to the project's website, ENGINE will use an internal repository as one of the main tools to internally collect, process, share, and make available/findable open research data. Information and data about materials, software development, LCA, monitoring and diagnostics process development produced during the project, their relations, and tests and characterizations performed will be collected in the internal repository. This will allow transparency of results for all partners and traceability of materials and processes developed in ENGINE.

All fully open research data can be found through the ENGINE community at Zenodo repository, the open data research repository of OpenAIRE and CERN.

Metadata

Furthermore, metadata features, in addition to the project name and Grant Agreement (GA) number, will be associated with all published data, including:

- Abstract/description
- Access and licensing information
- Associated project and community
- Associated publications and reports
- Bibliographic information
- Digital Object Identifiers
- Grant information
- Keywords
- Language
- Version numbers

Keywords

Specific keywords will be assigned to public datasets. In general, the consortium has defined several general keywords for all public datasets, publications and deliverables, etc.

Versioning

DOI versioning of all datasets hosted at Zenodo is provided by their respective native DOI service.

Making data accessible

Data accessibility

As outlined by the HEU Open Access guidelines, research data generated by HEU projects will be made accessible with minimum of restrictions related to protection of personal data or sensitive information that is governed by privacy concerns and/or commercial or security reasons.

Decisions concerning the sharing of (selected) datasets will be taken by project management group. Principal investigator in collaboration with project partners will take all the appropriate measures to make relevant data openly available and usable for third parties for study, teaching and research purposes.

If, after project closure, permission to re-use the data is required, all requests for further use of data will be considered carefully and whenever possible approved by the principal investigator or the person mandated with the task. Permission for data use will be granted providing there are no IPR or confidentiality issues involved or any direct overlap of research questions with the primary research. Permission will be provided by contacting Principal Investigator (project manager). Contact information and appropriate procedure will be provided in connection with other metadata.

Main focus in data sharing will be on the data underlying prospective scientific publications ensuring the validation of results presented in publications.

Published and FAIR-compatible data will be archived in a common and open data repository. Recommended generic and certified repository services, CERN's Zenodo in particular, will be used to enhance long-term accessibility and re-usability of the data.

Data accessibility platforms and tools

- ENGINE website public area: All publications (public deliverables, scientific articles in peer-reviewed journals and newsletter, etc.) will be available to third parties during the whole project's duration.
- VTT's Teams site for the project: It will be running/online during the entire project's duration. All beneficiaries will have access to this area to upload and download all restricted items to consortium.
- EC Participant's Portal: In addition to the deliverables (both public and confidential), project reports and contractual documents that are accessible to all project partners and the EC, the final summary report will be available through the EC Cordis (https://cordis.europa.eu/).
- ENGINE's internal data management platform, the ENGINE exchange: A data repository will be established to store production, test and characterization data related to produced materials and components and to store related meta data (production dates, author, materials, processes etc.). The platform will be accessible for all partners through web client to interact with the data or to create new entries and upload data. For advanced analysis of the data partners can access the repository with an API. The data model of the repository shall be standardized as much as possible. The repository will be kept confidential during the project. Parts or the full database will be released to a research data repository, e.g. Zenodo, at the end of the project.
- Zenodo: In order to comply with HEU Open Access and to increase the project's impact, all research open data will be put into a public repository like Zenodo. The main items to be provided in the Zenodo HEU ENGINE Community: public deliverables, all scientific publications in peer-reviewed journals, newsletters, and any open data. These materials will also be uploaded into the European Commission Funded Research (OpenAIRE) Community in Zenodo.

As the open research datasets will be included in the Zenodo repository: no methods or software tools are needed to access them and there are no restrictions on use.

Making data interoperable

Variables and value names will be constructed following general data processing conventions common to the research subject. List of value names and used vocabulary will be provided in a separate list. Examples of vocabulary information to be managed within the project will be e.g. number of variables / units of observation, list of variables with the name and label of each variable as well as its values and value labels, frequency distribution of each variable, information on the classifications used and meanings of abbreviations used. The standards and ontologies provided by related activities and respective networks will be utilized, such as the EMMC and EMCC, will be utilized as much as seen possible.

While knowing that not all the research datasets that is being collected and that will be generated during the project's duration are public, some confidential datasets will be kept confidential and therefore will not be available for access or inter-operability. Moreover, files such as simulation models cannot be used by other partners that do not have the software license.

Nevertheless, some of the data that will be publicly available and put in open access repositories like Zenodo can be interoperable. JSON (JavaScript Object Notation) scheme is used by Zenodo for metadata. It is possible to export research data into other formats such as: BibTex1, CSL2, DataCite, Dublin Core, DCAT, JSON3, JSON-LD, GeoJASON, MARCXML4, and export to Mendeley, Citavi.

Increase data re-use

Ownership of datasets will belong to project consortium after the project completion. Creative Commons licence CC-BY-SA or CC-BY will be used for any opened datasets, unless there are compelling reasons to select more restricted type of CC-licence. Creative commons licences will by default include also a disclaimer of liability for the re-use of opened data. Data availability is summarized in Table 2.

Table 2. Data Availability

Data	Type	Availability
Underlying data published in scientific journals	Public anonymous	Will be made after publication
Underlying data published in scientific journals	Confidential anonymous	This will not be available due to privacy concerns, but the main results will be published.
Associated with public deliverables	Public anonymous	Once approved by the EC
Data not associated with neither papers nor deliverables	Public anonymous	Once approved by data controllers and to be available by end of the project at the latest
Other data	Confidential non-anonymous	Not available for reuse

No definite period or time limit is planned for access or re-use of the data. However, the opened data will be deposited in a repository that guarantees data integrity on the bit level. At this point no continuous data curation policy to guarantee full long-term digital preservation of datasets is planned.

Justification for possible case-specific embargo for published data will be decided by project consortium. Embargo will be sought primarily in connection with any potential patent application based on project results.

For all public open data, it will remain reusable via Zenodo for at least 20 years. As stated by the Zenodo Repository “Items will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least.” In case the repository must stop operations, continuity plans are also envisaged “best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories.”

Allocation of resources

Costs related to research data management and opening are eligible as part of the project grant. Cost allocation is based on the assumption that maximum of 5 % of total project costs will be needed to

make research data quality-controlled, FAIR-compatible and as open as possible. ENGINE has a full WP, WP2 Industrial Data Management, focusing on all aspects of data thoroughly. Additionally, the data manager, ethics board in collaboration with the coordinator dedicate resources for data management and continued development of the project data management as necessary.

During the project consortium partners will be responsible for managing and curating datasets at their possession. At the project ending, consortium steering group will mandate Principal Investigator or project data manager to take care of long-term preservation and sharing of datasets. Free of charge research data repository tools are used by ENGINE. The costs of data management are allocated in the project budget, as most partners have foreseen budget for Open Access fees, so no further costs are envisaged for such activity.

Data security

At the beginning of the research project, the research consortium will decide and agree on the tasks, roles, responsibilities and rights relating to data collection, dataset management and data use.

During the project research datasets will be available only to those project partners or project consortium members, who have been accredited by and their data usage has been approved by Principal Investigator or authorized project consortium member. Project partners will be responsible for curating, preserving, disseminating and deleting in appropriate manner the datasets in their possession. Retention time for curated datasets will be the same as for other project results at the project consortium partners.

Data collected or acquired within the project will be stored in a secure IT environment behind a firewall at Project consortium partners premises or in secure cloud environment provided by Project consortium partners selected IT service providers. Access to it will need registration and authentication. Principal Investigator will check applications for the use of data. Where access is granted to research data, this will be provided through secured telecommunications channels. ENGINE has a dedicated Data Protection Officer support the project's Data Management Officer and Coordinator in any related tasks.

Long-term and secure preservation of published research data will be ensured by using only certified and OpenAIRE guidelines compatible repositories.

Zenodo data security

As the ENGINE project will also use the public repository Zenodo, safety considerations are well ensured. Zenodo's data security settings as outline by Zenodo (https://about.zenodo.org/infrastructure/) are described below:

- "CERN Data Centre: Our data centre is located on CERN premises and all physical access is restricted to a limited number of staff with appropriate training and who have been granted access in line with their professional duties (e.g. Zendo staff do not have physical access to the CERN Data Centre) .
- Servers: Our servers are managed according to the CERN Security Baseline for Servers, meaning e.g. remote access to our servers are restricted to Zenodo staff with appropriate training, and the operating system and installed applications are kept updated with latest security patches via our automatic configuration management system Puppet.

- Network: CERN Security Team runs both host and network based intrusion detection systems and monitors the traffic flow, pattern and contents into and out of CERN networks in order to detect attacks. All access to zenodo.org happens over HTTPS, except for static documentation pages which are hosted on GitHub Pages.
- Data: Zenodo stores user passwords using strong cryptographic password hashing algorithms (currently PBKDF2+SHA512). Users' access tokens to GitHub and ORCID are stored encrypted and can only be decrypted with the application's secret key.
- Application: We are employing a suite of techniques to protect your session from being stolen by an attacker when you are logged in and run vulnerability scans against the application.
- Staff: CERN staff with access to user data operate under CERN Operational Circular no. 5, meaning among other things that
    - staff should not exchange among themselves information acquired unless it is expressly required for the execution of their duties.
    - access to user data must always be consistent with the professional duties and only permitted for resolution of problems, detection of security issues, monitoring of resources and similar.

staff are liable for damage resulting from any infringement and can have access withdrawn and/or be subject to disciplinary or legal proceedings depending on seriousness of the infringement."

Ethics

Legal issues, personal data and privacy aspects

Privacy of data subjects will be secured by following closely the General Data Protection Regulation (Regulation (EU) 2016/679 of the European Parliament and of the Council). The project consortium has appropriate technical and organisational measures in place to carry out data protection during the project.

Processes that handle personal data have been designed and built with the GDPR principles taken into account. Processes provide safeguards to protect research data (e.g. using pseudonymization or full anonymization where appropriate), and use the highest-possible privacy settings by default. No person or organisation involved will unintentionally be identifiable directly or indirectly in the datasets. Any indirect reference to sensitive personal information or e.g. lines of businesses, branches or industries will be removed and destroyed after the anonymised dataset has been checked and validated.

After curation no person-related data is available publicly without explicit, informed consent, of the data subject and – if no anonymization is required – publicly available data cannot in any circumstances be used to identify a subject without additional information stored securely in a separate place. Principal investigator and processors of research data will always retain an unambiguous and individualized affirmation of consent from the data subject and the subject will always have the right to revoke her/his consent at any time.

During and after closure of the project the project coordinator will clearly disclose any datasets, which have been collected during the project and declare the lawful basis and purpose for their

processing. In addition, project coordinator will state how long the data will be retained and also unambiguously declare, if it is being shared with any third parties or outside of the EEA. Data subjects of the project will have the right to request a portable copy of the data collected in a common format, and the right to have their data erased under specified circumstances. VTT employs a data protection (privacy) officer (DPO), who is responsible for managing compliance with the GDPR.

Ethical issues

Research integrity and ethical principles related to data collection and use are covered in detail in the ethics self-assessment section of the grant application and monitored and developed by the ENGINE Ethics Board.

References

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

Co-funded by the European Union

Appendices: preliminary data summary

WP	Name of dataset	Description	Format	Responsible / Data controller	Origin	Dissemination level	Comments
1, 2	Casting - Material certificates	Certificates, material characteristics and properties		ABS	ABS	CO	
1, 2	Casting - Production process and product results	Certificates, material characteristics and properties		ABS	ABS	CO	
1, 2	PPAP (Part Approval Process) documentation	FMEA, Control Plan and Flow Chart of the process		ABS	ABS	CO	
1, 2	Die forging, heat treatment and product results	Certificates, material characteristics		SIFR	SIFR	CO	
1,2	Machining production process and product results	Certificates, respective product measurements		WAR	WAR, WIT	CO	
1	Engine simulation	Full engine simulation dataset containing loading data		WAR	WAR	CO	
1,2	Engine operating conditions analysis	Analysis dataset including site specific data		WAR	WAR	CO	
1	Component simulation, simulated defect maps, measured defect maps	demonstrator associated datasets for decision-making		WAR	WAR, WIT, ABS, SIFR	CO	

	and test data						
1, 4	Engine testing and usage dataset	Dynamic moment on crankshaft & combustion force. Online monitoring data (separately for testing and operation)		WAR	WAR	CO	
1, 6	Engine maintenance dataset	Includes defect maps and operational histories, dataset for LCA		WAR	WAR, GD	CO	
1, 6	Engine end of life dataset	Includes defect maps and operational histories, dataset for LCA, data on refurbishing/repair/scrapping		WAR	WAR, GD	CO	
3	Continuous casting and rolling simulation data	Simulation outputs on defects, reduction of cross section etc.	various databases (e.g., hdf5, vtk)	ABS	ABS, VTT simulations	Partly restricted open publication	Simulation methods and results also in part published
3	Forging and machining simulation data	Simulation outputs on defects, microstructures etc.	various databases (e.g., hdf5, vtk)	VTT	VTT, WAR simulations	Partly restricted open publication	Simulation methods and results also in part published
3	Heat treatment simulation dataset	Simulation outputs on heat treatment simulation, effect on microstructures	vtk	VTT	VTT, ABS, WAR simulations	Partly restricted open publication	Simulation methods and results also in part published
3	UT inspection simulator dataset	Physics-driven UT simulator outputs	ivtk, vtk	VTT	VTT simulations	Partly restricted open publication	Simulation methods and results also in part published
3	MBSE simulation dataset	Multibody and system scale simulation outputs	various databases	UOULU	UOULU, WAR	Partly restricted open publication	Simulation methods and results also in part published

3	Fatigue and fretting simulation	Material property simulation outcomes, assessment of lifetime and defect map criticality	various databases (e.g., hdf5, vtk)	VTT	VTT, WAR, TAU, BOKU	Partly restricted open publication	Simulation methods and results also in part published
3	AI training dataset for modeling defect generation processes	Hybrid dataset (experimental and synthetic) for training AI for predicting defect generation processes	various databases	VTT	VTT, ABS, SIFR	Partly restricted open publication	Simulation methods and results also in part published
3	AI training dataset for defect identification	Hybrid dataset (experimental, synthetic, full scale measurements) for training AI for defect identification in NDE tests	various databases	AEONX	AEONX, WAR, WIT, VTT	Partly restricted open publication	Simulation methods and results also in part published
3	AI training dataset for lifetime prediction	Hybrid dataset (experimental and synthetic) for training AI for predicting lifetime	vtk	VTT	VTT, AEONX, WAR, BOKU, UPD, GBW, TAU	Partly restricted open publication	Simulation methods and results also in part published
3	Multiscale modeling AI training dataset	Synthetic data for training AI for performing scale transitions between product and material scale models	vtk	TAU	TAU, VTT	PU	
4	Ultrasonic inspection	Material ultrasonic inspection		WAR	WAR production site measurements	CO	Testing already running
4	Immersion UT	Submerged UT material inspection	point cloud, vtk	ABS	ABS test site measurements	CO	Investment decision made
4	Surface roughness	Material surface roughness test	CSV	UOulu	UOulu laboratory	Partly restricted open publication	
4	Residual stress measurements	Material residual stress measurements	CSV	GBW	GBW laboratory	Partly restricted open publication	

5	Fatigue tests, notch sensitivity tests, multiaxiality tests	Fatigue test results	CSV	UPD	UPD, BOKU, UOULU, GBW	Partly restricted open publication
5	Fatigue test	Ultrasonic fatigue test	CSV	GBW	GBW, BOKU	Partly restricted open publication
5	Fretting test	Fretting test results	CSV	TAU	TAU	Partly restricted open publication
6	Scrap chemistry, steelmaking process parameters etc.	Dataset for LCA	various	GD	ABS	Partly restricted open publication
6	Rolling process description and parameters	Dataset for LCA	various	GD	ABS	Partly restricted open publication
6	Die forging process description and parameters	Dataset for LCA	various	GD	SIFR	Partly restricted open publication
6	Heat treatment process description and parameters	Dataset for LCA	various	GD	SIFR	Partly restricted open publication
6	Machining process description and parameters	Dataset for LCA	various	GD	WAR	Partly restricted open publication
7	Training material collector dataset	Dataset for drafting the training materials	various	VALTEH	ABS, SIFR, WIT, WAR	CO
7	Training materials of	Disseminated training materials for manufacturing plant workers and	video, pdf,	VALTEH	VALTEH	PU

	WP7	design engineers	html