Ref. Ares(2023)5334174 - 01/08/2023

MAXIMA

D1.3. Data Management Plan

Ángela Muñiz, FEUGA

Carmen Cotelo, FEUGA

Noelia Vilar, FEUGA

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No GA 101096097

Funded by the European Union

GRANT AGREEMENT NUMBER: 101096097
PROJECT ACRONYM: MAXIMA

EXECUTIVE SUMMARY

This report serves as the initial deliverable of Task 1.3 "Data Management Plan" and presents the initial Data Management Plan (DMP) for the MAXIMA project. The primary objective of the DMP is to provide a comprehensive overview of all datasets collected and generated within the project, while also defining the data management policy adopted by the MAXIMA consortium.

Deliverable Number	WP / T
D1.3	WP1 / T1.3
Lead Beneficiary	**Deliverable Author (S)**
FEUGA	Ángela Muñiz Carmen Cotelo Noelia Vilar
Beneficiaries	**Reviewers**
ENSAM	Stéphane Clenet, Asmaa Messaoudi
Planned Delivery Date	**Actual Delivery Date**
31.07.2023	01.08.2023

Type of deliverable	R	Document, report (excluding the periodic and final reports)	X
	DEM	Demonstrator, pilot, prototype, plan designs	
	DATA	Datasets, microdata, etc.	
	OTHER	Software, technical diagram, algorithms, models, etc.	

Dissemination level	PU	Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project's page)	X
	SEN	Sensitive, limited under the conditions of the Grant Agreement	

Version	Date	Author	Document history/approvals
V1	19/06/2023	Ángela Muñiz	1st consolidated version for partners feedback
V2	30/06/2023	Ángela Muñiz	Version sent for the internal review process
V3	01/07/2023	Stephane Clenet	Coordinator review
V4	10/07/2023	Ángela Muñiz	Integrated version after ENSAM revision
V5	01/08/2023	Asmaa Messaoudi	Final integrated version

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

INDEX

FIGURE INDEX

1. Introduction

The Deliverable 1.3 "Data management plan" is part of work package 1 "Project Management" and has been developed following the OpenAIRE (Open Access Infrastructure for Research in Europe) guidelines, a European initiative that aims to support and promote open access to research outputs. The Data Management Plan (DMP) is intended to be a living document where information will be regularly added and revised as the project implementation progresses.

This document serves as a comprehensive guide outlining our approach to effectively managing and handling data throughout the project's lifecycle. As a recipient of Horizon Europe funding, we recognize the critical importance of data management and the need to adhere to the guidelines set forth by the European Commission.

This deliverable was drafted to establish essential components that will enhance the possibility of reusing the data gathered and processed throughout the project, promoting its discoverability, accessibility, compatibility, and potential for future utilization accordingly with the FAIR (Findable, Accessible, Interoperable, and Reusable) principles.

The DMP aims also to ensure that data generated and/or collected during the project is handled, stored, and shared in a manner that maximizes its value, promotes accessibility, and guarantees the protection of sensitive information. By implementing robust data management practices, MAXIMA aims to foster scientific integrity, enable reproducibility, and facilitate knowledge exchange within and beyond the project consortium.

1.1. Data management plan objectives

The DMP for MAXIMA is designed to accomplish the following objectives:

- **Identify the datasets generated by MAXIMA:** The DMP will outline the different types of data generated throughout the project, including experimental data, simulation results, and any other relevant datasets.
- **Ensure FAIRness of the datasets**: The DMP will define strategies to ensure that the generated datasets are Findable, Accessible, Interoperable, and Reusable (FAIR). This may involve using standardized metadata, assigning persistent identifiers, and implementing data sharing and interoperability protocols.
- **Allocate resources for data management:** The DMP will determine the allocation of resources, both in terms of costs and responsibilities, for effective data management during and beyond the project.
- **Establish data security procedures:** The DMP will establish protocols for data security, including measures for data recovery, safe storage during the project's duration, and long-term preservation.

By addressing these objectives, the DMP for MAXIMA aims to ensure efficient and responsible management of data throughout the project's lifecycle, promoting data sharing, reproducibility, and long-term usability of the generated datasets.

2. Data summary

2.1. Data origin

The MAXIMA project utilizes a combination of generated, primary, and secondary data from various sources. It is important to note that all data used from partners follows the consortium agreement, and data extraction is conducted from publications (primary) and previous research work (secondary) as relevant to the project's objectives. Here is an overview of the data origin/provenance:

- Generated/Raw Data:

The project generates its own raw data through research activities and experiments conducted by the MAXIMA team. This includes data obtained from in-house tools (ENSAM). Additionally, results data obtained from simulations, experiments, and analyses conducted by the project team are considered an integral part of the project's data.

- Primary Data:

Data collected directly from main sources by researchers within the project are considered primary data. This includes data provided by the project partners and data extraction conducted from publications. For example, 4MP serves as the source of high-fidelity Computational Fluid Dynamics (CFD) data at the motor scale.

- Secondary Data:

The MAXIMA project also utilizes secondary data, which refers to data that has already been collected through primary sources and made available for researchers to use. The *ecoinvent database*, used for certain analyses, provides secondary data on environmental impacts.

2.2. Data sources

Data reuse promotes collaboration, efficiency, and scientific progress by maximizing the value of existing data and enabling cumulative knowledge generation. In the MAXIMA project data reuse will be promoted to enhance the research and achieve the project goals.

The following datasets have been identified for re-use:

Physical Properties of ferromagnetic materials: We will re-use existing data related to physical properties of ferromagnetic material under different operating conditions. This data will help us understand the behavior of the material and its performance in various scenarios. By leveraging this information, we can assess the designed electric machine (EM) on the vehicle environment, evaluating its suitability and effectiveness. OCAS, Höganäs, and Mimplus provide valuable data on the physical properties of ferromagnetic materials. By reusing this data, the project can optimise the design and performance of our electric machines in real-world operating conditions. Their expertise enables us to make informed decisions for enhanced performance and reliability.

- **Effect of Manufacturing Processes**: Existing data on the effect of different manufacturing processes on material behavior will also be re-used. This information is valuable for assessing the impact of various manufacturing techniques on the overall performance and characteristics of the MAXIMA machine. By understanding how different manufacturing processes influence the machine's properties, we can optimize the manufacturing methods to enhance performance, durability, and cost-effectiveness.

- **Life cycle management**. The MAXIMA project aims to evaluate the environmental impacts of the designed electric machine throughout its life cycle. To achieve this, we will employ the Life Cycle Assessment (LCA) methodology, which considers the entire life cycle of the machine, from raw material extraction to end-of-life disposal. The team from VUB (Vrije Universiteit Brussel) will leverage the *ecoinvent database*[1], a comprehensive Life Cycle Inventory (LCI) database, to determine the environmental impacts associated with specific activities. This approach enables us to assess the sustainability and environmental footprint of the MAXIMA machine, considering factors such as energy consumption, emissions, and resource depletion.

- **Electrical machine design**. 4MP will actively reuse data from its own institution as well as data provided by project partners to support the design of electrical machines. The data reusability encompasses various aspects, including simulation methodologies, material numerical models, and good practices. The institution's existing data will be leveraged to enhance the design process. In addition, partners will contribute valuable data such as material characteristics, numerical results for integration into simulations, and experimental data obtained from prototypes.

The decision to re-use existing data has been made based on several reasons:

- **Availability and Accessibility**: The data provided by the partners of the consortium or the by the different institutions/companies is readily available, making it convenient for the project to access and utilize it in the project research.
- **Cost and Time Efficiency**: By re-using relevant data, time and resources that would otherwise be required to collect new data can be save. This allows the consortium to focus more on analyzing and interpreting the data for our specific needs.

However, it is important to note that not all existing data may be suitable or applicable for our project. If any datasets were considered but ultimately discarded, the reasons for discarding them could include irrelevance to our research objectives, inconsistency with the required data quality standards, or limitations in the data's scope or coverage.

Overall, the re-use of existing data, both internally and from our project partners, plays a significant role in enhancing the efficiency and effectiveness of our research efforts in designing the manufacturing process flow, LCA, and developing the digital twin for the MAXIMA machine.

2.3. Data purpose

The data generation and re-use activities within the MAXIMA Project serve several important purposes directly related to the project's objectives of accomplishing the design of the motor, recalibrating simulation models, generating low order models, and building digital twins. These purposes include:

- **LCA Data Re-use**: The re-use of data supports the LCA carried out in the project, helping to understand and analyse the environmental impacts associated with the MAXIMA electrical machine. This data re-use enables the project to assess the environmental sustainability of the machine throughout its life cycle, contributing to informed decision-making and eco-friendly design choices.
- **Understanding Impacts and Optimization**: The generated data allows for a deeper understanding of the impacts of the MAXIMA electrical machine and helps optimize the material properties for the developed motor topology. By analysing the data, the project can identify areas for improvement, refine design parameters, and enhance the overall performance and

[1] E*coinvent database*: https://ecoinvent.org/

efficiency of the machine. This optimization process aligns with the project's goals of developing advanced electric machine technology and efficiency.

- **Design Accomplishment**: Data generation and re-use are crucial for accomplishing the design of the motor. Through simulation loops, the project can iteratively analyse the electromagnetic, thermal and mechanical aspects of the motor design. By generating data from these simulations, the project gains insights into the motor's performance, identifies areas for improvement, and refines the design parameters to enhance overall efficiency and effectiveness.
- **Recalibration and Validation**: The project aims to improve the accuracy and reliability of simulation models by recalibrating them based on experimental data. By comparing simulation results with real-world measurements, the project can refine and adjust the models, ensuring they accurately represent the motor's behavior. This recalibration process strengthens the confidence in the simulation results and helps validate the design choices made during the project.
- **Generation of Low Order Models**: The generation of low order models, electromagnetic, thermal and mechanical, is a valuable outcome of the data generation and re-use process. These simplified models serve as a bridge between detailed simulations and real-time applications. They provide a computationally efficient representation of the motor's behavior, which can be utilized for various purposes such as optimization, control, and real-time performance assessment.
- **Motor Testing and Data Collection**: Data results such as voltage, current, torque, speed, and temperature will be collected during motor testing. This data provides crucial insights into the motor's performance characteristics and enables the project to validate the design and functionality of the machine. The collected data aids in optimizing motor functionality and ensures that it meets the required specifications and performance criteria.
- **Calibration and Training:** High fidelity Computational Fluid Dynamics (CFD) simulations conducted at the motor scale will be used to train and calibrate the thermal module of the Digital Twin. This calibration process improves the accuracy and reliability of the digital twin's predictions regarding the motor's thermal behaviour. Accurate thermal modelling is essential for optimizing the cooling system design and ensuring the motor's reliability and longevity.
- **Building Digital Twins:** Data generation and re-use contribute to the development of digital twins, which replicate the behaviour of the motor. The data collected from simulations, recalibration, motor testing and low order models help create a digital twin that closely represents the real motor. This digital twin serves as a powerful tool for performance improvements, testing the inverter and control system, and advancing the overall development of the project.
- **Cooperation and Development**: The project involves collaboration among partners, and the generation of data, including the development of ferromagnetic material models and new source code, facilitates cooperation and the creation of a digital twin. This collaborative effort fosters innovation and knowledge exchange, enabling the project to leverage expertise from multiple domains. The digital twin, which replicates the motor's behaviour, serves as a valuable tool for performance improvements, testing the inverter and control system, and advancing the overall development of the project.

Additionally, the data re-use within the MAXIMA project helps update the electric machine specifications, assess the designed electric machine on the system and car level, and evaluate the impacts of the EM technology on the vehicles. These activities support the project's comprehensive approach in examining the performance, compatibility, and integration of the designed electric machine within the larger system and vehicle context.

Overall, the data generation and re-use activities within the MAXIMA project play a vital role in achieving the project's objectives and enhancing collaboration.

2.1. Data utility

The data generated or re-used within the MAXIMA project holds potential utility for various stakeholders beyond the project itself, particularly Original Equipment Manufacturers (OEMs) and automotive suppliers.

OEMs can benefit from the data by gaining insights for designing and developing advanced electric machines, optimizing performance, and assessing environmental impacts. Automotive suppliers can use the data for component development, manufacturing process optimization, and evaluating product performance.

Machine designers, Tier-1 traction motor providers, and electrical vehicle OEMs can also find value in the public data, as it can aid in the development and optimization of motor designs for improved sustainability.

Researchers in the field of motor manufacturing and eco-design can benefit from the data as it provides insights into the impacts associated with the manufacturing process of electric motors.

Additionally, the data can serve as a valuable for **Academia** and as an educational resource for students within the different institution involved, offering **support in teaching**, and enhancing their understanding of motor design and environmental considerations.

Moreover, the data generated by the MAXIMA Project holds significant value for all the **partners involved in the consortium**. As the partners collaborate and contribute to the project's objectives, the data becomes a valuable asset for their future development work. It is important to note that any confidential data will be safeguarded and not accessible for external use. The project recognizes the significance of intellectual property management in collaborative endeavours to ensure the successful implementation and exploitation of the project's results. For this purpose, it will be followed carefully the Article 16.1 Grant Agreement information where background is defined and Access Rights to data, know-how or information granting principles are stated. Besides, Access Rights to results and background specific restrictions and/or conditions for implementation will be also carefully followed as stated in Article 16.4 Grant Agreement and Annex 5 of the Consortium Agreement.

2.2. Data types and formats

The MAXIMA project will generate and reuse various types and formats of data to support its research activities. MAXIMA employs a diverse range of data formats to capture and analyse data, enabling effective research and analysis in various aspects of the project's objectives. These include:

- **Text documents:** The project generates and reuses text documents in formats such as .doc (Microsoft Word) and .pdf (Portable Document Format).
- **Images:** Various image formats such as .jpeg and .png are generated and utilized within the project.
- **Presentations:** Presentations are created and reused in formats like .odp (OpenDocument Presentation), .ppt (Microsoft PowerPoint), and .pdf.
- **Spreadsheets:** The project generates and reuses spreadsheets in formats such as .ods (OpenDocument Spreadsheet) and .xlsx (Microsoft Excel). The MAXIMA Project may utilize calculation spreadsheets from software tools like Mitcalc and MathCAD. Mitcalc is a mechanical engineering calculation software that provides a wide range of calculations for various

mechanical components and systems. MathCAD, on the other hand, is a mathematical and engineering calculation software that enables users to perform complex mathematical calculations, analyze data, and create visual representations of results.

- **Interviews:** Data from interviews will be collected and stored as text documents and other formats such as .mp3, and .mp4, .avi, .mov., .raw
- **Calculation models:** The MAXIMA Project will utilize various calculation models from software tools such as JMag, ANSYS, MatLab, Abaqus, Code Carmel, Python library and KissSoft. These software packages offer powerful capabilities for simulating and analyzing different aspects of the design and performance of electrical machines.
    - JMag is a simulation software specifically designed for electrical devices, enabling researchers to model and analyze electromagnetic phenomena, such as magnetic fields and electrical currents, within the machines.
    - ANSYS is a widely-used engineering simulation software that offers a broad range of capabilities, including structural analysis, thermal analysis, and electromagnetic simulation. It allows researchers to study and optimize the mechanical and thermal behavior of electrical machines.
    - MatLab is a high-level programming and numerical computing environment that provides a versatile platform for algorithm development, data analysis, and simulation. It is commonly used in engineering and scientific research for modeling and analyzing various systems, including electrical machines.
    - Abaqus is a finite element analysis (FEA) software that specializes in solving complex structural and thermal problems. It can be utilized to simulate and analyze the mechanical behavior and thermal performance of electrical machines.
    - KissSoft is a software tool specifically designed for the design and optimization of machine elements, including gears, shafts, and bearings. It aids in the calculation and analysis of mechanical components within electrical machines.
- **Source code:** The project will generate source code using programming languages like Matlab, Python, C++, and in-house software such as Code_Carmel. The code_Carmel software is a calculation code developed by the Modelling team of the L2EP supported partially by U. Lille and ENSAM and the electromagnetism modelling team of EDF R&D. It is primarily used for solving quasi-static electromagnetic problems in the modeling of electrical machines. The code_Carmel software is implemented in Fortran 90 and utilizes input and output files in .unv or .med format, which are provided by various CADsoftware or the Salome platform, respectively. The source code files may have specific formats associated with each programming language, including Matlab (.m), Python (.py), and C++ (.cpp, .h).
- **Web code**: Web-based languages such as HTML, CSS, and JavaScript will be used for web development tasks. Additionally, JavaScript Object Notation (.json) may be used as a data interchange format within the project.
- **Finite element model:** If required, the project may generate a finite element model using JMag software. JMAG interfaces with third-party software and integrates with major CAD and CAE systems, enabling seamless analysis.
- **Digital Twin:** The specific format for the source code of the Digital Twin implementation is yet to be defined.
- Life cycle inventory database(s): LCI databases compatible with brightway2 framework .bw2package will be generated.

2.3. Data size

The expected size of the data that we intend to generate, or re-use varies depending on the specific aspects of the MAXIMA project. Here is an overview:

- VUB: The data generated for the LCA is expected to be compiled in a document of at least 30 pages. This document will contain comprehensive information and analysis related to the environmental impacts of the MAXIMA project. The LCI database(s) is expected to range between a hundred KB to less than MB.
- OCAS: The size of the Finite Element Models will depend on the required accuracy and complexity of the simulations. Excluding the Finite Element Models, the data size can range from 1 to 3 GB.
- ENSAM & Emotors: Each of these partners anticipates generating or re-using data in the range of 10-100 GB.
- STLA: The data generated is expected to be gathered in a pdf document of at least 10 pages.
- UPC Research Groups: For electrical topics, the data size may range from a few kilobytes (KB) to megabytes (MB). However, for thermal topics, the data size can be in the order of gigabytes (GB).
- 4MP: Given the complexity and computational requirements, the data size 4MP can range from 50 to 100 terabytes (TB).
- FEUGA: The project may generate tens of gigabytes (GB) of data over the course of a year. This would include website analytics data, user engagement data, form submissions, content data, and metrics, as well as videos and pictures.
- MPT: It is expected to generate 10 GB of data per year.

It's important to note that the sizes provided are approximate and can vary based on the specific circumstances, requirements, and request of each of the partners within the MAXIMA project.

3. FAIR data

3.1. Making data findable, including provisions for metadata

The data generated within the MAXIMA project will follow provisions for making it findable and accessible. Persistent identifiers, such as DOIs (Digital Object Identifiers), will be assigned to the outputs of the LCA and other relevant data. These outputs, which include the inventory, will be stored in platforms such as Zenodo or GitHub, ensuring their long-term accessibility and traceability.

Furthermore, the outputs resulting from the building and development of the manufacturing process flow, as well as the creation and utilization of the digital twin, have the potential to serve as inputs for publications in scientific journals or for patent applications. In such cases, the publications will be assigned DOIs, and the researchers involved will be identified through their respective ORCID (Open Researcher and Contributor ID).

This systematic approach ensures that the data generated within the MAXIMA project can be easily identified, referenced, and linked to related publications and researchers, promoting transparency, and facilitating further research and innovation in the field.

- Metadata

Metadata plays a crucial role in ensuring the discoverability and understanding of research data. While the data itself may have a specific availability period, metadata related to the publication and patent can be preserved separately and made available through electronic libraries or other appropriate platforms.

This allows researchers and users to access information about the data, even if the data itself is no longer available.

Metadata will be created for the journal publications resulting from the MAXIMA project. This metadata will include information such as the title of the publication, authors' names and affiliations, publication date, abstract, keywords, and potentially additional descriptive information depending on the specific requirements of the target journal. Additionally, some program files within the MAXIMA Project may include metadata.

The MAXIMA Project website will also have metadata associated with it. Website metadata provides information about the website's content, structure, and other relevant details. This metadata helps search engines, browsers, and other tools understand and categorize the website's content accurately.

The metadata will be carefully curated to ensure that the publications are findable and can be easily discovered and accessed by researchers. The metadata will be associated with the publications in databases such as Web of Science or Scopus, which are widely used by the research community to search for scholarly articles and track citation metrics.

By including comprehensive metadata, the MAXIMA project aims to enhance the visibility and discoverability of its publications, allowing other researchers and stakeholders to locate and access the valuable research outputs generated within the project.

- Standards followed

The MAXIMA project will adhere to various disciplinary and general standards to ensure the quality, integrity, and proper documentation of its research outputs. The following standards will be followed:

- **ORCID for authors:** The Open Researcher and Contributor ID (ORCID) will be used to uniquely identify and distinguish authors. ORCID provides a persistent digital identifier that helps to establish authorship and ensure proper attribution.
- **Abstract:** Each publication will include an abstract, which provides a concise summary of the research objectives, methods, and key findings. The abstract helps readers to quickly assess the relevance and significance of the publication.
- **License:** The publications and associated data from the MAXIMA project will be assigned appropriate licenses to specify the terms of use, distribution, and attribution. This ensures that the intellectual property rights and usage rights are clearly defined.
- **Publication date:** The publication date will be included in the metadata of each publication. This information is crucial for tracking the chronology of research findings and establishing precedence in the scientific literature.
- **Funding:** The MAXIMA project will acknowledge and provide information about the funding sources that have supported the research. This demonstrates transparency and helps to attribute credit to the funding agencies.
- **Intellectual property rights references:** In cases where the outputs of the project are eligible for patenting, relevant patent references will be included to indicate any associated intellectual property rights.

By adhering to these disciplinary and general standards, the MAXIMA project aims to uphold best practices in scholarly communication, enhance research reproducibility, and facilitate the proper dissemination and recognition of its research outputs.

3.2. Making data accessible

The Maxima project recognizes the importance of data accessibility and has taken steps to ensure that the project's data will be deposited in a trusted repository. A Zenodo community has been created specifically for the Maxima project, where all papers and datasets will be uploaded. Zenodo provides a reliable and user-friendly platform for sharing research outputs, ensuring long-term preservation and accessibility. Zenodo is a general-purpose repository that allows researchers to openly share various research outputs, including datasets, publications, software, and more.

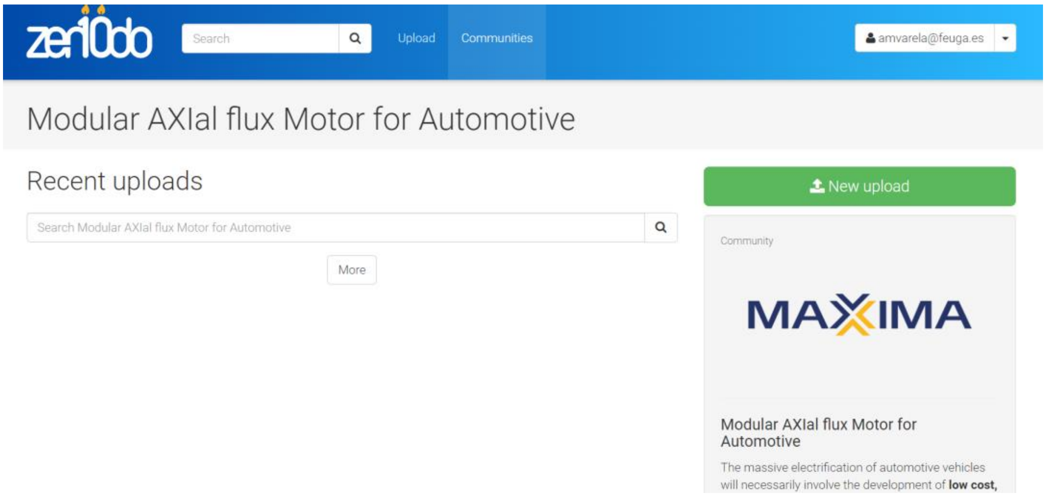

Figure 1 Maxima Zenodo profile

Zenodo is integrated with OpenAIRE, which means that when researchers deposit their data or publications on Zenodo, they have the option to link their content to the corresponding project on OpenAIRE. This integration helps enhance the visibility and discoverability of research outputs and ensures compliance with open access requirements.

By linking Zenodo and OpenAIRE, researchers can benefit from broader exposure for their work, increased citation potential, and alignment with European Commission policies on open access and research data management.

Additionally, partners involved in the Maxima project will also make use of trusted repositories for specific data types. For journal publications, it is likely that the data will be deposited in repositories such as Scopus, which is a widely recognized database for scientific literature. Furthermore, the project's inventory data will be stored on repositories like Zenodo or GitHub, which are popular platforms for hosting and sharing code, datasets, and other research materials.

- **Storage in PURE:** The publication outputs from VUB will be stored in PURE (https://researchportal.vub.be/), which is a research information management system. PURE provides a centralized platform to manage and store research publications, ensuring their long-term preservation and accessibility within the institution.
- **Storage in SAM:** Another platform where data will be archived is the open Archive SAM (https://sam.ensam.eu/), the Arts et Métiers Open Access Repository. SAM serves as the open access repository for Arts et Métiers, providing a comprehensive collection of research papers, book chapters, patents, and conference materials from researchers affiliated with Arts et Métiers laboratories. This platform ensures that the research outputs from Arts et Métiers are easily

accessible to the wider academic community and the public, promoting the dissemination and impact of their scholarly work.

- **Storage in HAL**: ENSAM will be also using Hyper Article en Ligne (HAL - https://hal.science/?lang=en) as storage platform. HAL is an open science platform with an international scope. It provides researchers with a platform to easily share and access their publications, which are well referenced by search engines and connected to other services such as ORCID and preprint servers. HAL is considered a common good for research, as it is supported by the largest research organizations and most French universities. It is a public, sustainable, and responsible infrastructure that aims to promote open access to scholarly work.

- Restricted access conditions

In the MAXIMA project, the openness of data depends on the nature, origin and purpose of the action that they will be subjected (According to Article 9 within the Consortium Agreement, i.e. implementation, exploitation) and therefore, some of the project data access will remain restricted to consortium members. When the disclosure of data does not pose a risk to the partners background or to the possibility of the intellectual property (IP) protection of the consortium, there is potential for greater openness. In such cases, data can be made available to a wider audience, ensuring transparency, increasing impact, and enabling broader collaboration and knowledge exchange.

It is important to carefully evaluate the IP implications of data disclosure and consider the contractual agreements and legal obligations of the project. If the consortium determines that sharing certain data does not compromise IP protection or conflict with any contractual obligations, it can be made openly available, facilitating access for other researchers and stakeholders.

Identified restrictions to share are:

- Data shared or generated by Stellantis, a beneficiary of the project, is to be used exclusively for the MAXIMA project, as stated in Annex 5 of the Consortium Agreement. It is not permitted to utilize this data for other projects or purposes. This restriction is in place to protect Stellantis' legitimate interests and ensure the confidentiality of their proprietary information. Furthermore, the data shared or generated during the project must be used solely for the purposes of the MAXIMA project. This restriction ensures that the data remains within the intended scope of the MAXIMA project and maintains its confidentiality and integrity.
- LCA results can generally be made openly available for sharing. However, some data inputs required for the LCA, may not be suitable for open access. The availability of these data inputs depends on the partners who provide the data, that may restrict their sharing.
- Regarding information related to materials, manufacturing processes, and design, it is considered confidential by default. This is because this information comes from industrial partners and may be subject to IPRs considerations. Therefore, it will not be openly available.
- The use of the code_Carmel software will be limited to ENSAM and his third party, the University of Lille.

However, synthetic data obtained from the project and used in publications or patents can be made openly available. This ensures that the project's findings and innovations can be shared with the wider research community.

- Internal repository

During the MAXIMA Project, access to the data will be provided through a Microsoft SharePoint platform, with collaboration and communication facilitated through Microsoft Teams. However, it is important to note that access to the data will be restricted to individuals who are part of the consortium and directly involved in the project. This ensures that only authorized personnel can access and utilize the data, maintaining confidentiality.

Microsoft SharePoint and Teams offer secure and controlled environments for data management, sharing, and collaboration. The platforms provide features such as user authentication, access controls, and permission settings to ensure that only approved individuals can access the project's data. This helps protect the sensitive and confidential information generated within the project and restricts its availability to authorized stakeholders.

To ascertain the identity of the person accessing the data, it will be necessary to verify their affiliation and involvement in the project. Only individuals who are part of the project and can be authenticated as authorized participants will be granted access to the data. This helps maintain data security and ensures that only eligible individuals can retrieve and utilize the shared information.

After the project concludes, access to confidential data that were inputs for the LCA will not be granted. However, the data that were shared within the project and are not subject to restrictions will remain accessible on the SharePoint platform. Access to this data will continue to be limited to individuals who participated in the project and were involved in the specific task or activity related to the data.

- Software documentation and references

Documentation or references about the software used to access or read the data will be included to facilitate data understanding and utilization. This documentation can provide instructions, guidelines, or explanations on how to use the software effectively. Including documentation and making relevant software available enhances the reproducibility and transparency of the data analysis process, enabling others to replicate and build upon the research findings.

Furthermore, if the relevant software, such as the activity-browser Python-based software, is open-source, it is possible to include the software itself along with the data. This allows users to directly access and use the software for processing or analysing the data. To ensure proper referencing and version control, it is important to include information about the specific version of the software used. This information helps users replicate the analyses or experiments conducted within the project, ensuring transparency and reproducibility.

In the case of the ENSAM in-house software, where a website is available, serves as a platform to provide documentation, instructions, and access to the software. The above-mentioned website provides further details on how to access, install, and use the code_Carmel software effectively. When relevant, MAXIMA will refer to the website for information on how to access and utilize the software for working with the associated data.

3.3. Interoperable data

To ensure interoperability and facilitate data exchange and re-use within and across disciplines, the MAXIMA project will follow relevant data and metadata vocabularies, standards, formats, and methodologies. Some of the practices and standards that may be considered include:

- **CAD (Computer-Aided Design):** CAD data for the design of the electrical machine will be prepared in a standard format that allows interoperability with other partners' systems and tools. Common formats such as STEP (Standard for the Exchange of Product Data) or IGES (Initial Graphics Exchange Specification) are commonly used for interoperability in CAD design.
- **Simulation Results:** Simulation results generated within the project will be shared using standard format documents such as Microsoft PowerPoint (.ppt), Excel (.xlsx), and PDF (.pdf). These formats

are widely recognized and can be easily accessed and interpreted by partners, facilitating data exchange and analysis.

- **Material and Manufacturing Process Information:** Data related to materials and manufacturing processes will be encoded in standard formats to enable interoperability with partner systems. This could include formats such as STEP (Standard for the Exchange of Product Data) for 3D models or other relevant industry standards. In addition to these standard formats, the project uses an Excel file accessible on the SharePoint platform, titled "Process Flow Analysis.xlsx." This Excel file contains relevant information and data pertaining to the manufacturing process flow analysis. It serves as a resource for understanding and documenting the sequential steps involved in the manufacturing processes.
- **Digital Twins:** The implementation of digital twin codes/modules will follow standard formats or languages to ensure interoperability. Community-endorsed best practices for digital twin development and integration may be followed to enable seamless collaboration and data exchange. Commonly used languages such as Python, MATLAB, or C++ may be employed to ensure compatibility and interoperability.
- **Results Data:** The results data obtained from simulations, experiments, and analyses will be presented in tabular formats, which are widely recognized and easily interoperable across disciplines. This could include formats such as CSV (Comma-Separated Values) or Excel files. Results from IP searches, including tables, information, and findings, will be presented in formats such as CSV, PDF, and PPT.

The specific vocabularies, standards, and methodologies selected will depend on the nature of the data and the requirements of each task. The project team will assess community-endorsed best practices and standards relevant to each specific data type to ensure effective interoperability and maximize data exchange and re-use within and across disciplines.

Moreover, the MAXIMA project data will include qualified references to other data sources. This includes referencing other data from within the project itself, as well as datasets from previous research that are relevant to the project's objectives. By including qualified references, the project ensures proper attribution and enables traceability of data sources, allowing other researchers to access and utilize the referenced data for their own investigations or as a basis for further research.

3.4. Increase data re-use

To provide documentation needed to validate data analysis and facilitate data re-use, the MAXIMA project will employ several approaches.

- **Journal Publication:** The research findings and data analysis will be documented in a journal publication. The publication will undergo a review process, ensuring that the methodology, data cleaning, analyses, variable definitions, units of measurement, and other relevant information are properly documented and validated.
- **Standardized Data Format:** For data collection, a specific format will be imposed to standardize the data across different sources. If a suitable format does not exist, the project will define a format that meets the requirements of the research and facilitates data analysis and re-use. This standardized format will ensure consistency and compatibility among the collected data, making it easier for researchers to understand and utilize the data.

By incorporating these practices, the MAXIMA project aims to provide comprehensive documentation that enables the validation of data analysis and promotes data re-use. Researchers will have access to essential information such as methodology, codebooks, data cleaning procedures, analytical techniques, variable definitions, and units of measurement, supporting transparency, replicability, and the potential for further exploration and collaboration.

- Data validation

The MAXIMA project implements several data quality assurance processes to ensure the reliability and accuracy of the collected and generated data. Some of the relevant processes include:

- **Data Quality Matrix:** For LCA inputs, a data quality matrix will be used to assess the quality of the data provided by partners. This matrix evaluates various aspects of data quality, such as data source, completeness, representativeness, reliability, and validity. It helps in identifying and addressing any potential issues or uncertainties associated with the data, ensuring that only high-quality data is used for the analysis.
- **Data Validation:** The project will conduct rigorous data validation procedures to verify the accuracy and consistency of the collected and generated data. This process involves cross-checking data against established standards, comparing data with independent sources, and performing data verification tests. Data that does not meet the required quality standards will be flagged for further investigation and potential correction.
- **Documentation and Metadata:** Comprehensive documentation and metadata will be provided for each dataset to ensure transparency and enable data quality assessment. This includes information on data collection methods, data sources, data processing and cleaning procedures, variable definitions, units of measurement, and any associated limitations or uncertainties. Clear and well-documented metadata allows users to understand the context and limitations of the data, facilitating informed decision-making and appropriate data usage.
- **Peer Review:** The project will subject its research findings and data analysis to a peer review process. Independent experts in the field will evaluate the methodology, data quality, analysis techniques, and interpretation of the results. Peer review helps to identify any potential flaws or biases in the data and analysis, enhancing the overall quality and reliability of the research outcomes.

These data quality assurance processes contribute to ensuring that the data used in the MAXIMA project is of high quality, reliable, and suitable for its intended purposes. They promote transparency, validity, and confidence in the data, enabling sound decision-making and facilitating data-driven research and innovation.

4. Allocation of resources

In the MAXIMA project, there are several costs associated with making data and other research outputs FAIR (Findable, Accessible, Interoperable, and Reusable). These costs may include:

- **Cost of publishing in open access:** Making research outputs openly accessible to the public may incur publication fees in open access journals or conferences. These fees cover the costs of peer review, editing, production, and hosting the research outputs in freely accessible platforms. The project has allocated a budget for open access publication to ensure the availability of research outputs to a wide audience.
- **Cost of data storage and archiving:** Storing and archiving research data requires appropriate infrastructure and resources. This may involve costs for cloud storage, data repositories, or dedicated servers. The project may allocate funds for data storage and archiving to ensure the long-term preservation and accessibility of the data beyond the project's duration.
- **Cost of data security and privacy:** Ensuring data security and protecting sensitive information may involve additional costs. This may include measures such as data encryption, access controls, anonymization techniques, and compliance with data protection regulations. The project may

allocate resources to implement appropriate security measures to safeguard the data and protect the privacy of individuals involved.

The project management team will assess and allocate the necessary resources to fulfil the FAIR principles and promote the accessibility and usability of research outputs while considering the available budget and funding constraints.

The decision on what data will be kept and for how long will be based on several factors. These may include legal requirements, data sharing agreements, ethical considerations, and the potential value and relevance of the data for future research or reference. The project management team, in consultation with relevant stakeholders, will determine the specific data that will be preserved and establish guidelines or policies for long-term preservation.

In addition, each partner has designated a data management representative to oversee the appropriate handling and governance of data within their respective organizations. The roster can be referred to in Annex 1, a confidential section exclusively accessible to internal consultation.

5. Data security

In the MAXIMA project, provisions are in place to ensure data security throughout the data lifecycle. Here are the key considerations:

- **Secure storage and archiving:** During the project, the data will be stored in a SharePoint platform, which provides secure storage capabilities. SharePoint offers access control mechanisms, data encryption, and backup features to protect the data from unauthorized access and ensure its availability.
- **Internal drive storage:** The project data will also be stored on internal drives, which are likely to have appropriate security measures in place to protect sensitive information. These internal drives may be hosted on secure servers or network storage systems, which can include access controls, encryption, and regular data backups to facilitate data recovery.
- **Data transfer:** When transferring sensitive data within the project, secure methods will be employed. This may include encryption of data during transfer, using secure file transfer protocols (SFTP) or virtual private networks (VPNs) to establish secure connections and prevent unauthorized interception or access**.**
- **Data recovery:** To ensure data recovery in case of any unforeseen incidents, regular backups will be performed. This includes maintaining backup copies of the data at different time intervals, preferably in separate physical locations or cloud storage services. Backup procedures and schedules will be defined to minimize the risk of data loss and facilitate recovery if needed.

The project team will follow best practices for data security and adhere to any relevant regulations or guidelines to safeguard the confidentiality, integrity, and availability of the data throughout the project duration and its conservation afterwards.

Annex 1. Data management responsible

PARTNER	PERSON IN CHARGE
4MP	Ioan-Claudiu Cotovanu
ENSAM	Stéphane Clénet
EMOTORS	Ioan Deac
UPC	Daniel Montesinos Miracle
HÖGANÄS	Anna Ahlqvist
OCAS	Ahmed Abouelyazied
FEUGA	Ángela Muñiz
VUB	Lea D’amore
STELA	Belhadi M'Hamed
UTCN	Claudia Martis
MIMPLUS	Johannes Maurath