MaxH2DR

Maximise H2 Enrichment in Direct Reduction Shaft Furnaces

Data Management Plan

Deliverable 4.5

Thomas Piontek, Tobias Kempken, Thorsten Hauck
VDEh-Betriebsforschungsinstitut GmbH

Valentina Colla, Ismael Matino
Scuola Superiore Sant'Anna

November 2022

This project has received funding from the European Union under grant agreement NUMBER — 101058429 — MaxH2DR

Table of contents

1. Introduction

This document details the Data Management Plan (DMP) for the MaxH2DR project. The detailed Data Management Plan is set in this document and will be revised continuously during the project duration. This plan describes rules for production, collection and processing of data. It provides detailed information on the project data lifecycle, privacy, and the MaxH2DR policies for data collection, storage, access, sharing, protection, retention, and destruction. The process of data management is based on the EC's guidelines on FAIR Data Management. The FAIR Data Principles are explained in detail in the following section.

In this project different types of data are generated. Typical are sensor data, construction plans, experimental reports for the measurement parts or simulation results (including relevant boundary conditions and assumptions) and program codes for the modelling part. If industrial process data and reports are provided during the project they will be treated in agreement with the company.

To make data accessible, free online repositories like "Zenodo" or "ResearchGate" will be used to save data with Digital Object Identifier (DOI). Unique File names based on "date-project_name-content-creator" will be used for each data type. Common keywords like "direct reduction", "reduction kinetics" or "DR-Pellets" will be used. To make the data structure transparent, different levels of detail will be published and explained in the specific files. The rules for file names make the files unique and retraceable even if data will not get a DOI (e.g., private or internal data shared between partners).

The research data are divided into experimental data which will be fully published according to Open Science principles and simulation data. Some simulation codes cannot be made public because these were initially developed prior to MaxH2DR and are covered by according restrictions. These codes are adapted and/or used in MaxH2DR. To make data interoperable standard data types and SI Units are used. The file structure is fixed, and every partner uses the same wording and acronyms consistent to standard of publications. Common data files will be used to guarantee a wide range of accessibility. For the restricted data a contact and service person will be available. Most data and results will be shared with the scientific community through publications in scientific journals and presentations at conferences. If data cannot be made public, the data owners will briefly justify the reason (e.g. protecting IP, commercial agreements, personal data, etc). BFI takes responsibility for controlling the linkage and retractability between the partners during the project. At this stage, the licensing of datasets is not yet clearly defined. It depends on the results of the project, so that a decision will be made during the project duration.

Cost for storage and preservation of data are included in the project costs. Each partner is responsible for implementing the DMP and the WP leaders will manage the overall data flow and DMP implementation according to this document. Data access plans, right management and documentation for the main data repository will be implemented by BFI.

In the following text, the term metadata is used for "data providing information about other data". These may include descriptive, structural, administrative, reference, statistical or legal metadata. The term "(meta)data" is referencing to both metadata and data.

2. FAIR Data Principles

This Data Management Plan complies with "The FAIR Guiding Principles for scientific data management and stewardship", as laid out by Wilkinson et al. in 2016[1]. The FAIR Data Principles comprise of Findability, Accessibility, Interoperability and Reusability. This emphasizes machine-based utilization of data with none or minimal human intervention.

[1] M.D. Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship", Scientific Data 3, Article number 160018, 2016, Link

2.1 Findability

The first requirement in using data is finding it. Thus, metadata and data should be as easy to find as possible for humans and computers. In this context, machine-readable metadata are a basis for automatic discovery of datasets. Thus, this is an essential aspect regarding findability.

Wilkinson et al.[1] defined findability as follows:

- (Meta)Data are assigned a globally unique and persistent identifier
- Data are described with rich metadata (see Reusability)
- Metadata clearly and explicitly include the identifier of the data it describes
- (Meta)Data are registered or indexed in a searchable resource

2.2 Accessibility

If the first step of finding data is successful, the next step is accessing the data. This includes both authorization for access as well as authentication. Authorization refers to having the required access rights or privileges to resources. Authentication consists of proving the identity of a user, while authorization refers to proving that this specific user has specific access rights or privileges.

Wilkinson et al. [1] defined accessibility as follows:

- (Meta)Data are retrievable by their identifier using a standardized communications protocol
- The protocol is open, free, and universally implementable
- The protocol allows for an authentication and authorization procedure, where necessary
- Metadata are accessible, even when the data are no longer available

2.3 Interoperability

The term interoperability refers to the possibility of integrating data with other data. Additionally, the data needs to interoperate with applications of analysis, storage and processing workflows.

Wilkinson et al. [1] defined interoperability as follows:

- (Meta)Data use a formal, accessible, shared, and broadly applicable language for knowledge representation
- (Meta)Data use vocabularies that follow FAIR principles
- (Meta)Data include qualified references to other (meta)data

2.4 Reusability

Finally, the aim of FAIR principles is optimizing the reuse of existing data. For this purpose, (meta)data need to be well-described, enabling its replication and/or combination in different settings. The reusability of data includes the clarification of dedicated licenses.

Wilkinson et al. [1] defined reusability as follows:

- (Meta)Data are richly described with a plurality of accurate and relevant attributes
- (Meta)Data are released with a clear and accessible data usage license
- (Meta)Data are associated with detailed provenance
- (Meta)Data meet domain-relevant community standards

3. Data Summary

A summary of data generated or collected within MaxH2DR is given in the following Table 1. In this table, the content of each dataset is described and each responsible partner is stated. Additionally, the origin as well as the reuse and/or reusability of each dataset is given. The description of each dataset concludes with the definition of data types and formats used and an estimation of the number of stored values as well as of its storage size.

Table 1: Summary of data generated or collected in MaxH2DR

Experimental investigations of reduction kinetics in different scales

Responsible Partner(s): IMZ

Content & Description: To investigate the possible use of sinter as a feedstock in the DR process in an atmosphere with 80-100% H_2, special stand for testing sinter reduction in a hydrogen atmosphere will be used to obtain data such as CSV, i.e. "comma separated values". These values will be saved as Excel Files in the measurements. These will be analyzed by Excel and other statistical tools. The results will be merged with the measurement values in Excel files for post-processing and validation.

Origin & Reuse: own measurements (not reused, but available for reuse)

Types & Formats: Numerical values in CSV, i.e. "comma separated values" and Excel files

Expected Size: up to 10,000 numerical values, up to 10 GB storage space

Physical demonstration of linked solid and gas flow in DR shaft furnace

Responsible Partner(s): BFI

Content & Description: The test rig in demonstration scale at BFI will be utilized to obtain data such as gas velocities, particle mass flows and pressure drops. These values will be saved as Ascii-Files in the measurements. Additionally, the experiments will be filmed through the acrylic window, leading to MOV-Files. These will be analyzed automatically by a Matlab program, calculating e.g. local particle velocities. The results will be merged with the measurement values in Excel-files for post-processing and validation.

Origin & Reuse: own measurements (not reused, but available for reuse)

Types & Formats: Numerical values in Ascii- and XLSX-Files, film data in MOV-Files

Expected Size: Numerical values: 10,000 – 100,000 values, several MB storage space
Film data: 10 – 100 files, several GB storage space

Physical property investigation at pellet bulk scale

Responsible Partner(s): UNISA

Content & Description: Unisa will design and operate a test rig for the reduction and measurement of forces applied and acting on samples of iron pellets at high temperatures. Design data will be collected and generated. The collected data will regard the properties and the description of the material and of the parts to be purchased. The design process will generate calculation data mostly done on excel files and MatLab. Graphical design files in formats to be defined. The experiments

done on this set-up will produce time series of measured parameters, such as temperature, gas composition and forces acting on the elements of the set-up in CSV data files. The first elaboration of these files will be carried out on excel spreadsheets or with MatLab procedures.
Training data with model materials will be produced on a Schulze ring shear tester to pass to the modelling partners of RUB. These data also will consist of time series of the stresses, and the lid position of the shear tester generated in CSV data files and elaborated in Excel spreadsheets.

Origin & Reuse: data by providers, and own-produced data

Types & Formats: PDF, xlsx, docx, m

Expected Size: 10^6-10^7 data values equivalent to 10-100 MB storage

Analysis output of searching innovators, inventors and business drivers related to H_2-enriched Direct Reduction in steelworks

Responsible Partner(s): PNO (CiaoTech Srl)

Content & Description: Output data from selecting funded projects, patents and commercial projects in relation to the H2-enriched Direct Reduction in steelworks. Aim of the analysis is the identification, classification and positioning of the most relevant actors (including innovators, investors and business drivers) within the MaxH2DR value chain. Multiple databases (such Lens.org or Patent Inspiration) and PNO's own search tools such as Wheesbee are used to perform this analysis.

Origin & Reuse: own researches and analyses (not reused but available for reuse)

Types & Formats: Structured Excel with multiple tabs in XLS

Expected Size: around 16 GB

Analysis output of technology market outlook

Responsible Partner(s): PNO (CiaoTech Srl)

Content & Description: Output data from market analysis. The approach for the Market value identification is based on secondary sources and data triangulation.

Origin & Reuse: own researches and analyses (not reused but available for reuse)

Types & Formats: Structured Excel with multiple tabs in XLS

Expected Size: around16 GB

Outcomes of the questionnaire/focus group led by Ciaotech as result of stakeholders' participation and for involvement in the exploitation activities

Responsible Partner(s): PNO (CiaoTech Srl)

Content & Description: Most relevant stakeholders identified during the Stakeholder Analysis will be invited to fill in an interactive questionnaire/survey or to participate in a focus group by PNO with the

purpose of assessing their perception towards main market drivers and barriers identified as well as MaxH2DR Key Exploitable Results. Participation of stakeholders will be entirely voluntary, under public consent. No vulnerable categories of individuals or persons unable to give consent will be involved. PNO will be responsible for questionnaires/ focus group preparation, data collection, analysis and storage in the Corporate intranet.

Origin & Reuse: own researches and analyses (not reused but available for reuse)

Types & Formats: PDF files

Expected Size: around 4 GB

Thermogravimetry experiments for kinetic measurements of pellet reduction under various gas composition

Responsible Partner(s): University of Lorraine (IJL-UL)

Content & Description: The experiments will mainly consist of reduction experiments using industrial ore pellets. A pellet will be placed in a thermobalance and its reduction by H2-CO reducing mixtures (of a composition chosen in agreement with the other partners) will be recorded. The temperature, gas composition, the type of pellet will be varied, and lump ore and sinter will also be tested. The data produced will be typically the measured mass, temperature and gas composition, as a function of time.

Origin & Reuse: own measurements (not reused, but available for reuse)

Types & Formats: Numerical values in Ascii, or Excel spreadsheet files.

Expected Size: around 50 files (one per experiment) of a few MB each – around 500 MB in total

Characterization of samples of reduced pellets

Responsible Partner(s): University of Lorraine (IJL-UL)

Content & Description: Some samples will be characterized before and after the tests, as well as intermediate stages, by SEM observation, X-ray diffraction, and Mössbauer spectrometry.

Origin & Reuse: own measurements (not reused, but available for reuse)

Types & Formats: TIFF, PNG, JPEG or PDF.

Expected Size: around 20 files (one per experiment) of a few MB each – around 200 MB in total

Shaft furnace simulations

Responsible Partner(s): University of Lorraine (IJL-UL)

Content & Description: Reductor is a 2D, 2-phase, 3-zone, 10-reactions, finite volume method based CFD model specifically developed to simulate the operation of DR shaft furnaces. It will be used to run simulations of the DR shaft furnace under various configuration. Some simulations will serve as benchmarks to be compared with other model or experimental results. The data produced

will be sets of text files (data and results) for every simulation. Only the most meaningful simulations will be stored.

Origin & Reuse: own calculations (not reused, but available for reuse)

Types & Formats: Numerical values in Ascii

Expected Size: around 20 files (one per simulation) of around 100 MB each - 2 GB in total

Experimental investigations of reduction kinetics at pellet bulk scale

Responsible Partner(s): Tata Steel (TS)

Content & Description: Reduction and softening tests will be conducted at the TS IJmuiden laboratories, using the combined HOSIM-BRASS test rig. In the experiments, mainly commercial type pellets will be used at specified size range. In the experiments, flat gas composition – temperature and varying gas composition – temperature vs. time profiles will be implemented. Conditions will be in alignment with the other partners. Experiment conditions and outputs (e.g., duration, inlet gas flow rate, temperature, inlet gas composition, mass sample, etc.) will be recorded.

Origin & Reuse: own calculations (not reused, but available for reuse)

Types & Formats: Numerical values in Ascii- and XLSX-Files

Expected Size: around 60 files of around 5 MB each - 300 MB in total

Characterization of the relevant physical properties of the oxides, intermediates and reduced pellets

Responsible Partner(s): Tata Steel (TS)

Content & Description: The reduced and partly reduced material will be investigated with respect to chemical and physical property changes during the H2-enriched DR experiments; e.g., degree of metallisation, mechanical strength, porosity, disintegration, softening, etc.

Origin & Reuse: own calculations (not reused, but available for reuse)

Types & Formats: Numerical values in Ascii- and XLSX-Files

Expected Size: around 60 files of around 2 MB each - 120 MB in total

Numerical DEM/CFD simulations of Direct Iron Reduction in Shaft Kilns

Responsible Partner(s): Ruhr-University Bochum (RUB-LEAT)

Content & Description: Description of the theoretical background, the implementation and results of a set of specific test-/application-cases of DEM/CFD targeting Direct Reduction of iron ore with hydrogen rich synthesis gas in shaft kilns

Origin & Reuse: Code, executable programs and applications specifically developed for MaxH2DR and the corresponding simulations (partial use/reuse/adaption of previously existing and published framework, available for reuse)

Types & Formats: Various data types like code repositories, HDF5 data files, AVI, MOV, JPEG, binary mesh files, PDF, plain ASCII and XML control files

Expected Size: Several TB, strongly dependent on the specific problems addressed and the strategy employed for raw data reduction, to be defined during the course of the project

4. Allocation of resources

Cost for storage and preservation of data are included in the project costs. The technical project coordinator BFI communicates this Data Management Plan to all partners. Each partner is responsible for implementing the DMP within their respective work packages and for the data it processes. BFI hosts a document exchange platform for all MaxH2DR partners, in which the communication and reporting documents will be aligned. SSSA is hosting the data exchange server for the process chain simulations in WP3. The partner responsible for the implementation of the Data Management Plan is BFI. As such, data access plans, right management and documentation for the main data repositories will be implemented by BFI.

For this purpose, a total of 4 person-months were dedicated in WP4. This includes the documentation of the overall workflow and process as well as the data check in each deliverable. These costs are included in the project costs. After conclusion of MaxH2DR, there will be service persons available at BFI and SSSA who can be contacted for further questions and work related to data management.

5. Data Security

The electronic documents are stored for at least one year after the completion of MaxH2DR on resident backup servers. Additionally, the documents are stored on a duplicated hard disk, to prevent data loss. Confidential data and information gathered in the context of this project will only be circulated among the staff of MaxH2DR members that are responsible for the specific data collection and analysis. To enhance security, such data and information will be stored in files encrypted with a password.

List of Acronyms and Abbreviations

Acronym	Full Name
BFI	VDEh-Betriebsforschungsinstitut GmbH
CIAOTECH	CIAOTECH s.r.l.
CSV	Comma Separated Values
DMP	Data Management Plan
DOI	Digital Object Identifier
DR	Direct Reduction
EC	European Commission
GB	Gigabyte
IMZ	Siec Badawcza Lukasiewicz – Instytut Metallurgii Zelaza IM. Stanislawastazica
MB	Megabyte
PDF	Portable Document Format
RUB	Ruhr-Universität Bochum / Ruhr-University Bochum
SSSA	Scuola Superiore Sant'Anna
TB	Terabyte
TS	Tata Steel
UNISA	Università degli Studi di Salerno / University of Salerno
UL	Université de Lorraine / University of Lorraine
WP	Work Package