UNEXMIN DELIVERABLE D9.3

DATA MANAGEMENT PLAN

Summary:

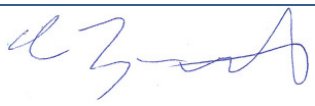
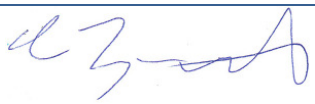
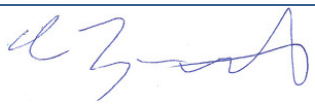
This is a review of data management of the UNEXMIN project. This document describes how the data will be exploited, stored, cured and preserved, what the project will generate. This document is the first version of the Data Management Plan which will evolve during the execution of the project to support the collection, curation, preservation and access of the data produced by UNEXMIN.

Authors:

Norbert Zajzon

This project has received funding from the European Union's Horizon 2020research and innovation programme under grant agreement No 690008.

Lead beneficiary:	**University of Miskolc (UNIM)**		
Other beneficiaries:			
Due date:	M6		
Nature:	Report		
Diffusion	Public		
Revision history	Author	Delivery date	Summary of changes
Version 1.0	Norbert Zajzon	13.07.2016.	
Version 2.0	Stephen Henley	13.07.2016.	English text improved
Version 3.0	Hilco van Moerkerk	14.07.2016.	English text improved
Version 4.0	Norbert Zajzon	15.07.2016.	Additional clarifications

Approval status			
Function	Name	Date	Signature
Deliverable responsible	Norbert Zajzon	18.07.2016.	
WP leader	Norbert Zajzon	18.07.2016.	
Reviewer 1	Stephen Henley	18.07.2016.	
Project leader	Norbert Zajzon	18.07.2016.	

Disclaimer: This report reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

Table of Contents

1) Executive summary 6
2) Objectives of the Data Management Plan 6
3) Types of data produced 6
4) Data and metadata standards 7
5) Policies of access and sharing 7
6) Plans for archiving and preservation 8

1. Executive summary

This Data Management Plan (DMP) describes the data management life cycle for all datasets to be collected, processed or generated during the UNEXMIN project. It covers:

- the handling of research data during and after the project,
- what data will be collected, processed or generated,
- what methodology and standards will be applied,
- whether data will be shared / made open access and how,
- how data will be curated / maintained and preserved.

Reviews, updating and adjustments of the DMP will take place during the execution of the project at meetings/workshops organized in the project (months 12-45). New version(s) of the DMP will be created if important changes occur and realized during the execution of the project. The DMP will be updated (at least) for the mid-term and final project reviews.

2. Objectives of the Data Management Plan

The purpose of this DMP is to support the management, by the project partners, of the data by the project partners which will be collected, generated and/or processed during the execution of the UNEXMIN project, and enhance the re-use, availability and survivability of these data.

3. Types of data produced

The UNEXMIN project will generate many different types of data in large quantities. These can be divided into four main groups: A) data relevant to the UX-1 robot design and performance, which range from stakeholder requirement evaluations, simple conceptual descriptions, general drawings to elaborated blueprints of all the parts and their test and performance reports from simulations to real hardware tests in relevant environments. B) Information produced about the test-sites, which, also, can widely vary from collected previously known information to high quality datasets and their visualization produced from the UX-1 robotic measurements. These cover many different data types, from xyz coordinates, distances, point clouds, pictures, temperature, conductivity, pH data, gamma-ray counts, spectral information of points / areas, mineral information of points / areas. It could also include databases produced by scientific instruments in a laboratory environment on reference samples to help the proper evaluation of the field data.
Type “A” data generally will be produced during the planning, production and validation of the UX-1 robots (WP1–5), whereas the type “B” data will be produced during the test dives and post-processing of the delivered data related to WP2, 6 and 7.
The “Flooded Mine Inventory” to be generated by WP5 in the project can be considered as type “C” data, which will contain an extensive list of the European flooded mines with their accessible information.
Other data will also be generated as deliverables (written reports) of the project, conference posters, brochures, talks (in ppt and/or pdf format) and publications (general and scientific). These type of data can be considered type “D”, and will be largely produced in WP8 and/or in cooperation with WP8.

4. Data and metadata standards

Data set reference and name

All data files produced in the project will include in their names the name of the project (UNEXMIN) and a short information on the type of data (e.g., "UNEXMIN_Ecton_3Dpointcloud") or, in cases where this may be impractical, will be held within folders that themselves follow this convention.

Standards and metadata

When compatible with the themes of the INSPIRE Directive, the data and metadata added in the project will follow the format recommended by the INSPIRE Directive.
The quality of the data will be discussed between partners, who will decide which data are of sufficiently high quality to be used for publication of the research done in the project. Only these data will be kept in an open data repository.
Metadata will be created following the model of the INSPIRE Directive (they will offer the same types of information). As the provisions in the INSPIRE Directive are compulsory for EU members, this will ensure a high degree of compatibility with other data sets in the EU.

5. Policies for access and sharing

An important issue is to monitor the progress of the novel technology development, carefully distinguishing among confidential data (like trade secrets eg. confidential know-how) and Open Research Data in accordance with the H2020 directives, and if relevant seek patent protection to proprietary technology and methods, if such new discoveries evolve through the research in the project. The responsibility to initiate the patenting will be upon the respective WP leader participants, with ownership proportion agreed within the actively participating research team members.

Data sharing

In addition to the various scientific methodologies and concepts UNEXMIN will develop technologies that will be suitable for commercial exploitation in the future. This will require very careful management of Intellectual Property Rights (IPR). The principles governing IPR are outlined in the Consortium Agreement and further details will be developed and agreed in the Newco business plan.
As a general rule, the data representing syntheses of the knowledge from previously published research will be open data from the moment they are put in the repository. The unpublished geological data (produced in the project and older unpublished data provided by partners) will be confidential for a 3-year period for the purpose of ensuring the novelty of the data used in scientific articles published by the partners. The unpublished technological data will remain confidential over a period that is decided by their owners.

For components of the project that are sufficiently unique and innovative, the developing WP shall give early warning to the steering committee about the confidentiality relevant to this component to all consortium members, not to disclose details until protection is in place.
Consideration will be given to patent protection before publication of details of these components.

All the metadata and data of types "A", "C", and "D" will be stored in a repository, accessible from the UNEXMIN project website. The publicly available Open Research Data will be available by any simple web browser. The search, browsing, displaying and downloading the data will be available freely, without any registration.
Type "A" data that cannot be patented (or would be too expensive to patent) should by default being considered "trade secrets" unless agreed by all the owners of the data and/or the Steering Committee to be eligible to be placed in the public domain. It should automatically be passed to Newco at the end of the UNEXMIN project, and it will then be a decision for Newco to make, whether and when to release such data.
It may be impractical to make anticipated large volumes of data of type "B" available in this way. Such data will be held in offline storage with multi-site backup for security, and made available to users on request. Part of "B" type data will be put to the online accessible storage, to represent what kinds of data are accessible upon request. These data will be carefully selected by the consortium to have good quality, contain relevant information, but do not contain any confidential, or ethically questionable information.

Security of data will be addressed as an integral part of the task of developing online and offline data storage methods.

The confidential data will be available only for the authorized personnel.

6. Plans for archiving and preservation

Archiving and preservation (including storage and backup)

As the Open Research Data will be kept in multiple copies by Newco and relevant project participants, the archiving and preservation will follow the general procedures used by the project participants. The partners will assess the archiving and preservation procedures and decide whether special procedures are needed for the data produced in the project and also for how long the data are going to be kept in the repository.
After the end of the project, copies of all data will be maintained by Newco under conditions to be determined within the Newco business plan, with off-site secure storage of full backup copies. The online open research data will be available for the public for minimum of five years after the end of the project, which is a responsibility of Newco. For contingency plan, if the Newco would not continue for any case, all the data, and the responsibility with them will be transferred to a permanent organization in the UNEXMIN consortium, such as the Miskolc University, or to the Eurogeosurveys for safe keeping.

Dissemination of the data:

- data stored in the project's website,
- presentations in scientific conferences,
- open access peer reviewed publications.

The data produced in the project together with any metadata will be held in a central repository to be managed by Newco with copies also maintained by relevant project participants. Full copies will also be held in secure off-site storage.

Associated software used by or produced within the project will be held by Newco but will be made available to users only under commercial terms as agreed within the Consortium Agreement and the Newco business plan.

Where appropriate, data which may be of use or interest to the wider public and non-specialists will be made available in suitable digital or paper formats.