RADIATE

Project: **RADIATE**

Research and Development with Ion Beams – Advancing Technology in Europe

Project No.: **H2020, # 824096**

Project duration: **01.01.2019 – 30.06.2023**

Deliverable D6.2: First update of the Data Management Plan

Lead author(s): Stefan Facsko (HZDR)

Final reviewer(s): Lino da Costa Pereira (IST), Wolfhard Möller (HZDR)

Work package:	WP6: Data Management
Task:	T6.2: Data Management, Archiving and Open Access
Deliverable type:	Open Research Data Pilot
Dissemination level:	Public
Delivery date:	M32 (August 2021)
Total number of pages:	8
Approved by coordinator:	20.12.2021 J.F. (HZDR)
Submission date:	20.12.2021
Keywords:	Data Management Plan, Open Access

Acknowledgment:

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 824096.

1 Introduction

This document is the first update of the Data Management Plan (DMP) of the RADIATE project, which has been made public on 28.06.2019 as D6.1. It provides the updated strategic framework for the data curation in RADIATE and sets out how research data will be managed throughout its lifecycle to ensure its authenticity, accessibility and interoperability for reuse and open access in order to comply with legislation, best practice and EU requirements in the RADIATE project. Care will be taken to balance data openness and protection of scientific data due to commercialisation, IPR, privacy and security. This document will be further updated regularly to adapt for the requirements and needs of the users, as well as for the fast-moving developments in the field of open access for Research Data.

RADIATE is an INFRAIA (Integrating and opening research infrastructures of European interest) project in H2020 with the main objective to provide user access to Europe's key ion beam facilities. Totally more than 16.000 h trans-national access are offered to users at 14 ion beam facilities in Europe during the project duration of 4 years. The portfolio of the ion beam facilities ranges from specialized low-energy ion beam systems to large accelerators far beyond laboratory scale. They are operated as user facilities offering access to very specialized analysis and materials modification techniques for users from academia and industry.

A major task in RADIATE is to develop, implement and operate the **Ion Beam Web Portal** www.ionbeamcenters.eu, collecting all information and services of the European ion beam centres and providing it to the ion beam research community. This portal will act as the central hub for exchange of information and data not only among the European ion beam centres, but also for sharing information and software with the user community and for access to ion beam-related databases, software, and research data.

Furthermore, the Ion Beam Web Portal will supply building blocks for

- a single-entry point for proposal submission, evaluation and beam time scheduling at the RADIATE infrastructures,

- protocols, guidelines and QA manuals for quality control and best practice in the ion beam centres,

- newsletters and reports

- a database of all scientific publications and published data obtained in the framework of RADIATE,

- database of software for simulation of ion solid interaction and data evaluation for ion beam analysis.

The RADIATE Consortium approves the principle of open access to research data and provides the necessary infrastructure for data management and regulates the access to research data within these terms and conditions. Open access to research data should be ensured wherever possible following the path of the citable data publication. Where in place, local repositories at the facilities or freely available repositories, like Zenodo, should be used. Furthermore, all data generated in the joint research activities of RADIATE can be requested from the open access responsible persons at the facilities as listed in the Appendix.

2 Data Summary

Access to the RADIATE facilities is provided in two categories: ion beam analysis (IBA) on the one hand and materials modification on the other hand (Table 1).

TNA topic distribution		
Category	**Topic**	**Method**
Analysis	Elemental Analysis & Depth Profiling	RBS
		ERDA
		NRA
		PIXE / PIGE / PES
		MEIS
	Ultrahigh Sensitivity	Cosmogenic dating AMS
		High-energy AMS
		Environmental tracer AMS
	Lateral Imaging	External microbeam IBA
		(ambient) MeV SIMS
		PIXE (μ-beam, camera)
		He-Ne microscopy
		H microscopy
		IBIC
	Defect Analysis	RBS/PIXE-C
		In-situ TEM
	Real-time in-situ Analysis	Dynamic / high-T IBA
Implantation & Irradiation	Broad Beam	Implantation / Doping
		Multi-beam
		Clean environment
	Local	Non-Ga FIB, He/Ne microscope
		Single ion implantation
		Cell irradiation
	Deep	MeV ions
		Swift heavy ions
	Shallow	Low-energy ions
		Highly charged ions

Table 1: ***Topics for trans-national access to ion beam facilities in RADIATE***

IBA comprises elemental analysis and depth profiling, which is mostly realised by standard techniques, i.e. RBS[1], ERD[2], NRA[3], PIXE/PIGE[4] and MEIS[5], ultrahigh sensitivity with AMS[6] for cosmogenic dating, for environmental tracing and with high energies for special nuclides (^{53}Mn, ^{60}Fe, ^{63}Ni), lateral imaging by μ-beam, defect analysis by RBS and PIXE in channeling conditions, and real-time in-situ analysis for dynamic changes in stoichiometry during high temperature or other extreme experimental conditions.

Materials modification includes ion implantation with low ion fluences and ion irradiation with high ion fluences, being performed homogeneously on a large area with a broad beam or locally with a focused beam and with single ion impacts. The depth of the modification or implantation can be adjusted between several hundreds of micrometres by using swift heavy ions with energies between few MeV and GeV and a few nanometres by using low-energy or highly charged ions. In particular, for surface modifications and implantation in 2D materials, ultralow ion energies of a few tens of eV will be provided.

[1] Rutherford Backscattering Spectrometry
[2] Elastic Recoil Detection
[3] Nuclear Reaction Analysis
[4] Proton Induced X-Ray Emission / Gamma - Ray Emission
[5] Medium Energy Ion Scattering
[6] Accelerator Mass Spectrometry

A trans-national user campaign at one of the RADIATE ion beam facilities proceeds along the process chain as sketched in Fig. 1.

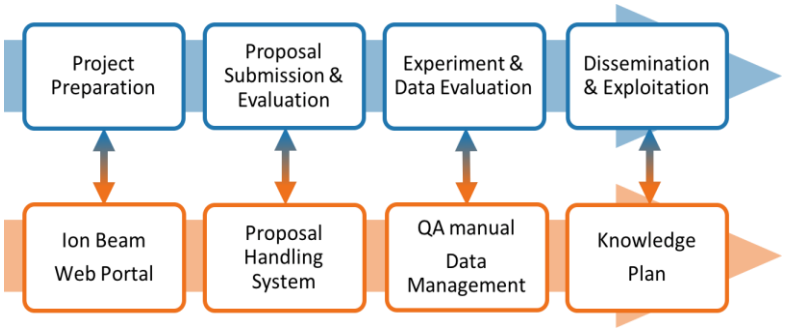

Figure 1: ***User campaign process chain.***

In the "Project Preparation" phase research data may be generated by simulation software, where Monte Carlo (MC) or Molecular Dynamic (MD) codes are used to simulate the ion solid interaction of the planned experiment. Furthermore, software for calculating expected spectra for ion beam analysis techniques are used in this phase to choose the right method and parameters for the experiment or measurement. For this purpose, software will be made available upon request to RADIATE users on the **Ion Beam Web Portal**. The research data generated in this phase is stored locally on the user's computer and can be used later in the data evaluation phase. Conversion to a common data format can be done by software tools which are also available on the **Ion Beam Web Portal**.

In the "Experiment Phase" raw data is produced by the data acquisition systems at the facilities. Raw data is stored in many different formats due to the very different instruments and acquisition systems available at the facilities. In order to transform raw data in research data, software tools will be developed and used for the conversion in a common data format. These tools will be made available to the users.

Finally, in the "Data Evaluation Phase" research data is analysed and further processed to get information about the analysed materials and scientific results. The RADIATE users are strongly encouraged to publish their results and research in **open access journals** and in **open data repositories**, e.g., the open repository Zenodo by CERN (www.zenodo.org). Data available on repositories will be registered to the European OpenAIRE with the corresponding publication and reference to RADIATE and will be listed on the Ion Beam Web Portal.

In addition to research data, which is generated during the user campaigns at the ion beam facilities, data is also generated by the **Joint Research Activities** (JRAs). The JRAs within RADIATE target specific developments in ion sources and beams, in detectors and electronics, as well as in software and data handling. Research data connected to these developments will be archived on repositories and published as open data or can be requested from the facilities, respectively. The contact persons of the RADIATE facilities for requesting data are listed in the Appendix.

3 FAIR data

3.1 Making data findable, including provisions for metadata

The proper management of Research Data is imperative to ensure that scientific findings are findable, accessible, interoperable and reusable (**FAIR**). The requirements for open access data to be **FAIR** will be addresses in RADIATE by adopting a common format for ion beam analysis data based on the World Wide Web Consortium Extensible Markup Language (XML). A universal **Ion Beam Analysis Data Format** (**IDF**) has already been developed in the previous FP7 European project SPIRIT and is implemented in

common computational ion beam analysis codes[7]. **IDF** includes **metadata** on the experimental parameters and information about the samples, as well as data and parameters for fitting.

The **IDF** was developed to allow easy transfer of data and simulation parameters between codes, as well as between experimentalists and data analysts and is implemented in the most popular general purpose IBA data analysis codes, such as RUMP[8], SIMNRA[9], NDF[10]. It is transparent, universal (catering to varying needs), and includes the most common features desired by both experimentalists who collect and archive data and by users who analyse the data. The **IDF** can also be readily extensible in order to include features specific to individual codes and laboratories, as well as being able to incorporate new features and options in the future. Furthermore, an open-source package is available for the implementation of the IDF[11].

For the additional services provided in RADIATE, like MeV SIMS, the **IDF** has been extended to handle data from multi-parameter experiments, e.g., combined IBA and Mass Spectrometry data. Moreover, due to the spreading of IBA imaging techniques the amount of data in one file is greatly increased. Thus, the file format and the ion beam analysis software will be updated to deal with the large data set.

In March 2021 the RADIATE Consortium discussed in a Workshop on Data Management and Standards the adoption of a common metadata format and the open access publication of data. The RADIATE consortium agreed to use the IDF as a common data format. Furthermore, the extension of the IDF schema for including multiparameter and imaging experiments were discussed. For large and complex data sets, e.g., microbeam images **Root based data structures** (used in CERN and other large infrastructures) can also be used in combination with **IDF**.

3.2 Making data openly accessible

Research data and results which have been generated during the user campaigns will be made available to the users via the Ion Beam Web Portal. To simplify this process the common data format IDF will be used and software will be provided as online web-based tools on the Ion Beam Web Portal.

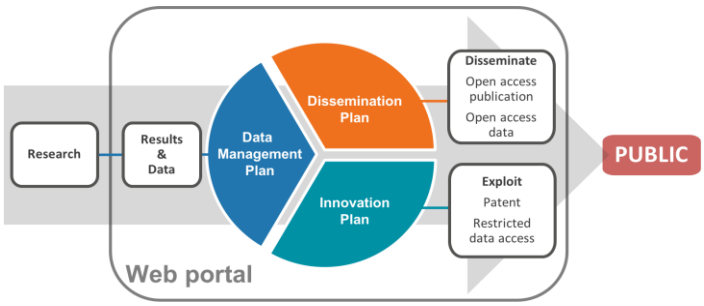

*Figure 2: **"Knowledge plan" of dissemination of results and data in RADIATE.***

Following the requirements and recommendation for Horizon 2020 projects all publications and research data will be published as open access unless there are issues of commercial confidentiality, either because there is a non-disclosure agreement in place or because the data is part of a body of research that is being protected (patent, copyright, etc.).

Figure 2 sketches the two ways how research data can be disseminated in open access publications and open access data or for commercial exploitation. Research data will be archived in local repositories at the facilities and open access to the data will be provided through the EU OpenAIRE

[7] N. P. Barradas, M. Mayer, and M. Thompson, Nucl. Instr. Meth. B 268, 1824 (2010).
[8] http://www.genplot.com/doc/rump.htm
[9] https://home.mpcdf.mpg.de/~mam/
[10] http://www.itn.pt/facilities/lfi/ndf/uk_lfi_ndf.htm
[11] N. P. Barradas, Nucl. Instr. Meth. B 332, 148 (2014).

portal except for results with commercial exploitation potential. Moreover, all data and results which will be made available for open access will be published also on the Ion Beam Web Portal. Except when there are issues of commercial confidentiality data will be freely available on request. At several facilities local repositories already exist, e.g. at HZDR the RODARE system[12] is based on the CERN software INVENIO and is compatible with the metadata format of DataCite[13]. This knowledge and expertise in data repositories will be spread to all facilities in RADIATE, which are in the process to setup such a system. The users will be strongly encouraged to publish their results in open access journals ("gold" open access[14]) or at least to make them freely available by uploading them in suitable open access repositories ("green" open access[15]). In addition, provision of open access to the data is strongly recommended to them.

3.3 Making data interoperable

The RADIATE Consortium will observe OpenAIRE guidelines for online interoperability, including OpenAIRE Guidelines for Literature Repositories, OpenAIRE Guidelines for Data Archives, OpenAIRE Guidelines for CRIS Managers based on CERIF-XML. These guidelines can be found at https://guidelines.openaire.eu/en/latest/.

New tools have been developed in RADIATE, i.e., an **IDF viewer** to view, edit and analyse IDF data and a Python library **pyIDF**, which can be used to handle IDF data in Python programs or in environments like Jupyter or in Python programs. These tools including detailed manuals will made available through the Ion Beam Web Portal to all users and the Ion Beam Community.

Information on data and metadata vocabularies, standards and methodology will be continuously added to the RADIATE DMP to further facilitate interoperability.

3.4 Increase data re-use (through clarifying licences)

RADIATE is expected to produce a substantial volume of research data and knowledge through ion beam-based analysis and materials modifications that will be presented to the scientific community, industry, policy-makers and society at large through dissemination actions. Data published through the repositories or made available upon request from the facilities will be accessible after an embargo period determined per dataset if required.

As RADIATE progresses and data is collected, further information on increasing data re-use will be outlined in subsequent versions of the DMP. In specific, information on how data will be licenced to permit the widest reuse possible, when the data will be made available for re-use, whether the data produced and/or used in the project is useable by third parties, a description of data quality assurance processes and specifications of length of time for which the data will remain re-usable will be provided.

[12] https://rodare.hzdr.de
[13] https://datacite.org
[14] "Gold" Open Access: Authors make a one-off payment to the publisher so that the scientific publication is immediately published in open access mode.
[15] "Green" Open Access: Due to the contractual conditions of the publisher, the scientific publication can undergo an embargo period up to six months since publication date before the author can deposit the published article or the final peer-reviewed manuscript in open access mode.

4 Allocation of resources

Each RADIATE partner has to respect the policies set out in this DMP. Datasets have to be created, managed and stored appropriately and in line with applicable legislation.

The Project Coordinator at HZDR has a particular responsibility to ensure that data shared through the Ion Beam Web Portal are easily available, but also that backups are performed and that proprietary data are secured. Backing up data for sharing through open access repositories is the responsibility of the partner possessing the data.

Validation and registration of datasets and metadata is the responsibility of the facility where the research data has been created during the user campaigns. Metadata constitutes an underlying definition or description of the datasets, and facilitate finding and working with particular instances of data.

5 Data security

As specified by the "rules of good scientific practice" we aim to preserve data for at least **ten years**. Approximated end volume of data collected in RADIATE is 1000 GB. Associated costs for data preparation for archiving will be covered by the project. Long-term preservation after the end of the project will be provided and associated costs covered by the facilities. During the project, data will be saved with backup on a separate server.

Procedures that will be implemented for data collection, storage, access, sharing policies, protection, retention and destruction will be in line with EU standards as described in the Grant Agreement and the Consortium Agreement, particularly Section 8 – Results, Section 9 - Access Rights, and “Annex I – Description of Work” of the Grant Agreement.

6 Ethical aspects

Data protection and good research ethics are imperative for the RADIATE consortium. Good research ethics meets all actions to take great care and prevent any situation where sensitive information could get misused. Research data which contains personal data will just be disseminated for the purpose for which it was specified by the consortium. Furthermore, all processes of data generation and data sharing have to be documented and approved by the consortium to guarantee confidentiality and anonymisation.

7 Further support in developing your DMP

The RADIATE DMP has been elaborated based on recommendation of Horizon 2020 using the DMP online webpage (dmponline.dcc.ac.uk). Furthermore, the DMPs and Data Policies of the project partners have been partly considered.

8 Appendix

Ion Beam Facility	Open Access Contact Person		
	Surname	First name	Email
HZDR	Facsko	Stefan	s.facsko(at)hzdr.de
CNRS CIMAP	Rangama	Jimmy	rangama(at)ganil.fr
CNRS INSP	Ganem	Jean-Jacques	ganem(at)insp.jussieu.fr
ETH Zürich	Vockenhuber	Christof	vockenhuber(at)phys.ethz.ch
INFN	Chiari	Massimo	chiari(at)fi.infn.it
IST	Silva	Rui	rmcs(at)ctn.tecnico.ulisboa.pt
JSI	Bucar	Klemen	klemen.bucar(at)ijs.si
JYU	Sajavaara	Timo	timo.sajavaara(at)jyu.fi
KU Leuven	Vantomme	Andre	andre.vantomme(at)kuleuven.be
RBI	Siketić	Zdravko	Zdravko.Siketic(at)irb.hr
SURREY	Costa	Catia	c.d.costa(at)surrey.ac.uk
UNIVIE	Steier	Peter	peter.steier(at)univie.ac.at
UU	Primetzhofer	Daniel	daniel.primetzhofer(at)physics.uu.se