

HORIZON 2020 RESEARCH AND INNOVATION FRAMEWORK PROGRAMME OF THE EUROPEAN ATOMIC ENERGY COMMUNITY

Nuclear Fission and Radiation Protection 2018 (NFRP-2018-4)

Project acronym	: SAND	SANDA				
Project full title:	Solvir Europ	Solving Challenges in Nuclear Data for the Safety of European Nuclear facilities				
Grant Agreemer	nt no.: H2020	H2020 Grant Agreement number: 847552				
Workpackage N°:	WP6	WP6				
Identification N°:	D6.6	D6.6				
Type of documen	: Delivera	Deliverable				
Title:	SANDA	– Data Mar	nagement plar	1		
Dissemination Lev	vel: PU	PU				
Reference:						
Status:	VERSIC	VERSION 1				
Comments:						
	Name	Partner	Date	Signature		
Prepared by:	E. González	1	15-04-2021	E. Jun		
WP leader:	E. González	1	15-04-2021	e lun		
IP Co-ordinator:	E. González	1	15-04-2021	E. Jun		

Table of content

3
3
5
6
7
8
9
.10
-

SANDA – Data Management Plan

1. DATA SUMMARY

The SANDA project will produce several types of data and SANDA is participating voluntarily in the Open Research Data (ORD) pilot according to the H2020 guidelines on data management and the requirements specified in Article 29.3 of the H2020 Model Grant Agreement. Special effort will be made to assure that all data relevant for final use or for analysis of data produced and compiled by SANDA will be findable, accessible, interoperable and re-usable (FAIR). The SANDA data management plan (DMP) implementing these actions is described in the current report.

All the results from SANDA will be published and described publicly but in addition, for the nuclear data produced by the project, either experimental or evaluated, the main dissemination path is provided through well-established international libraries and international nuclear data centres coordinated by the IAEA¹, https://www-nds.iaea.org/, and by the OECD-NEA, https://www.oecd-nea.org/databank. These organizations maintain on their own resources an infrastructure that receives, tests, archives, stores and distribute both experimental nuclear data in the EXFOR format and evaluated nuclear data libraries in the standard formats of ENDF and ENSDF. This mechanism assures that the project output data will be available broadly and well beyond the end of the project.

The main types of data produced by SANDA are:

1) The results of the differential or integral nuclear data measurements performed in WP2 or in WP5. For the description, reporting and use of these nuclear data there is an international standard defined by IAEA in its EXFOR database, and SANDA will communicate all experimental data to the EXFOR nuclear database. Experimental Nuclear Reaction Data (EXFOR) is an open database maintained and operated by the IAEA, https://www-nds.iaea.org/exfor/exfor.htm. It is also a standard format to describe the experimental information which is necessary to facilitate a harmonized presentation and retrieval for the analysis of these "raw" measured nuclear data.

The data types included in this category cover: reaction cross sections, fission yields, decay data, cross sections for the generation of secondary particles (integral or differential) and reaction rates. A complete listing of the data types that can be stored in EXFOR, and a detailed description of the retrieval system and other tools for EXFOR is described in the references². The papers also describe the basic EXFOR format and the structure of the database.

EXFOR data is usually used only by nuclear data experts, producers or evaluators. These experts will analyse the EXFOR data, check for consistency, compare with other data and with

¹ International Atomic Energy Agency is an agency from the of the United Nations to act as the world's central intergovernmental forum for scientific and technical co-operation in the nuclear field.

² The experimental nuclear reaction data (EXFOR): Extended computer database and Web retrieval system. V.V. Zerkin and B. Pritychenko. Nuclear Inst. And Methods in Physics Research, A 888 (2018) 31-43. https://doi.org/10.1016/j.nima.2018.01.045

Towards a More Complete and Accurate Experimental Nuclear Reaction Data Library (EXFOR): International Collaboration Between Nuclear Reaction Data Centres (NRDC). N. Otuka, E. Dupont, V. Semkova, B. Pritychenko, et al. Nucl. Data Sheets **120**, 272--276 (2014). https://doi.org/10.1016/j.nds.2014.07.065.

basic physical laws and derive "evaluated" nuclear data stored in another database with different format, that once verified, are the actual data used by scientists or engineers making simulations or analysis for commercial or experimental nuclear reactors, designing nuclear components or applications and other final uses of the data.

2) The results of the evaluation of reaction or nuclear structure and decay data resulting from the activities of WP4. These evaluated data derive from the interpretation of all the experiments available for one particular magnitude and the requirements of physical laws. These evaluated data are compiled in libraries and are the main source of data used by the final users, both for reactor or fuel cycle design, simulation and interpretation and for other applications of nuclear science and technologies.

The evaluated data corresponding to cross sections for nuclear reactions will be stored in the international standard ENDF files and the data will be available from the OECD-NEA Databank and from the IAEA ENDF page <u>https://www-nds.iaea.org/exfor/endf.htm</u>. The details of the ENDF format, procedures, and content are described in a pdf document³ available from the ENDF web pages.

The evaluated data corresponding to nuclear structure and decay data will be provided in the international standard ENSDF files and the data will update the IAEA data repository and distribution <u>http://www.nndc.bnl.gov/ensdf</u>. The details of the ENSDF format, procedures, and utility and analysis codes are described in pdf documents⁴ available from the ENSDF web page.

3) The rest of the results will be prepared as PDF documents included in the project deliverables or in the milestones reports. This includes most of the results from WP1, WP3, WP6, and part of results with non structured data from all the work-packages. All deliverables are public, once they are in their final stage, and can be found in http://www.sanda-nd.eu/, the project web-page, or in the EC web-page associated to the project. The milestone reports are also available from the project web. Both types of documents are reachable from a searchable repository using a Graphic User Interface, GUI, based on the DRUPAL web system.

³ <u>https://www-nds.iaea.org/exfor/x4guide/manuals/endf-manual.pdf</u>

⁴ <u>https://www.nndc.bnl.gov/ensdf/ensdf-manual.pdf;</u> <u>https://www.nndc.bnl.gov/nndc/evalcorner/ENSDF-Procedures.pdf;</u> https://www.nndc.bnl.gov/nndc/evalcorner/

2. MAKING DATA FINDABLE

As indicated in the previous section all the data of types 1) and 2) will be stored following welldocumented international standards. The documentation covers the formats and the details to correctly identify each set of data and the source of the data. This includes, for experimental data recording, an ENTRY and SUB-ENTRY identifier, the facility, detector, analysis, corrections and history of the data set. The complete set of metadata included depends on the type of experimental data, but the metadata are clearly self-explanatory. The details of required and optional metadata and dictionaries for potential values of some of the metadata are available in the documentation and manuals accessible from the corresponding web pages. For EXFOR it can be found at <u>https://www-nds.iaea.org/nrdc/basics/#Table of Contents</u> and the references for ENDF and ENSDF have already been provided in the corresponding foot notes.

As mentioned, data will be stored at international data repositories: for type 1), experimental data, in the EXFOR library at https://www-nds.iaea.org/exfor/exfor.htm and for type 2), evaluated data, for reaction data in the ENDF library https://www-nds.iaea.org/exfor/endf.htm and for type 2), evaluated data, for reaction data in the ENDF library https://www-nds.iaea.org/exfor/endf.htm and for type 2), evaluated data, for reaction data in the ENDF library https://www-nds.iaea.org/exfor/endf.htm whereas for nuclear structure and decay data the data will be submitted to ENSDF http://www.ndc.bnl.gov/ensdf.

For each of these libraries there are GUI that allow searching for the data according to many criteria. In particular, the most common search will look for a specific reaction, including target selection, a given energy range and a specific magnitude/quantity. The search criteria can also be more administrative looking for publication year, authors, institute, DOI, or reference numbers.

As for the documents stored in the SANDA web, the searchable parameters include the type of document (deliverable, milestone report, article, ...), the title, date, authors, work-package, and potential additional keywords.

3. MAKING DATA OPENLY ACCESSIBLE

Ultimately, all the results of the project will be openly available. The only restriction applicable is that data will remain private for the consortium until its quality is validated internally as ready for journal or deliverable publication. Formatting the data is rather time consuming and there will be some delay, up to several months, between the validation and the actual uploading of data to the final repositories.

The data will be accessible from the international repositories (IAEA and OECD-NEA) already indicated in the previous sections. All these repositories allow exploring and searching for data using GUI and downloading partial or full libraries/sets of related data. In addition, there is work in progress to provide application-programming interfaces (better known as APIs) allowing automated retrievals from the repository. Indeed, some draft specifications and prototype tests are available for the 3 libraries (EXFOR, ENDF and ENSDF).

All the software needed is documented and available from the web pages of the repositories and are maintained by the IAEA and collaborating institutions (the International Nuclear Reaction Data Centres – NRDC).

The arrangement for the use of these repositories was discussed with IAEA before submission of the proposal. In any case, the access and the possibility to upload nuclear data is granted and promoted by the mandate of the international organizations (for the European research area the intermediary is the OEDC-NEA Databank). Furthermore, there is a continuous connection by including one representative of the IAEA in all the SANDA general meetings and all the meetings and documents of the Executive Committee. Additionally, some of the teams and experiments included in the SANDA project, like the n_TOF collaboration, have nominated IAEA/EXFOR contact points to facilitate the efficient and early transfer of results to the IAEA repositories.

On the other hand, the responsibility of actually introducing the data in the libraries and verifying the quality and accuracy of the data sent to the IAEA/NEA libraries remains on the teams producing the data, following the tradition of decades of operation of these international repositories.

4. MAKING DATA INTEROPERABLE

As already indicated in previous sections the data of types 1) or 2) will be provided in rigorous international formats (EXFOR, ENDF or ENSDF), using well defined and documented metadata and in many cases with well defined dictionaries of acceptable values for the metadata.

In fact, every day libraries following these standards are read and utilized by different computer codes combining data from different origins (different experiments, different laboratories, different evaluators, different libraries...), in some cases after processing with standard conversion codes available from the same international organizations holding the data libraries.

These standards also allow using the data for widely different applications, from nuclear technologies to basic research like astrophysics, health applications for nuclear treatments or diagnosis, art applications for evaluation of artistic techniques and materials as well as dating and validating the originality of works of art, and many other industrial applications.

5. INCREASE DATA RE-USE

All the data distributed by SANDA to the international repositories and agencies will be publicly distributed, and they are to be considered in the public domain. As indicated above, no embargo is envisaged for the distribution of the data, but the data will not be submitted for publication until its understanding and quality is ready according to the criteria of the research teams, work-package leaders and the project coordinator. Once published data will be submitted to the above mentioned international organizations for dissemination. For data included in deliverables or milestone reports (most of the data to be produced by SANDA), their quality has to be confirmed by the signature, in a SANDA standard cover page, of the author (research team) of the deliverable, the work-package leader, and the project coordinator. Furthermore, there could be typically some months delay for the formatting of data to be send to the IAEA repositories.

The use and reuse of the data after the end of the project has no restrictions and there is no foreseen limit date to be allowed to reuse the data. However, when using the data it will be required to quote the source of the data and the authorship of the data and, as needed, of the associated information (analysis, detectors, interpretation...).

6. ALLOCATION OF RESOURCES and DATA SECURITY

The specific costs of making data FAIR are the costs for quality validation, specific data formatting, and collaboration with the international agencies. The costs for operation, long-term maintenance, support to access and development of access tools for the international repositories are assumed by the IAEA, the OECD-NEA and the collaborating institutions.

The specific costs for the project, were included in the proposal as part of the costs of the experiments, evaluations and other activities proposed in the project, because independently of the agreement with EURATOM for this contract, the dissemination of results via international repositories is considered as a mandatory step of the research in any nuclear data project. The analysis of a nuclear data experiment is not completed until the results have been published and the data has been formatted and submitted to the international repositories.

As indicated in the previous sections, the responsibility of actually introducing the data in the libraries and verifying the quality and accuracy of the date sent to the IAEA/NEA libraries remains on the teams producing the data. On the other hand, the validation of the data quality is a shared responsibility of the research teams, the work-package leaders and finally of the project coordinator.

The long-term preservation of the data stored in the international libraries is part of the mission of the IAEA Nuclear Data Services. They have been operating for decades providing a stable reference with no anticipated date for the end of these services. In addition, once a data set is accepted by the repository its preservation is ensured by the IAEA/NRDC's EXFOR rules .

For the deliverables and other data stored in the SANDA web-page, the coordinator (CIEMAT) that is operating the web server, has committed to operate the SANDA web-page in the present web address, at least, for 5 years after the end of the project. CIEMAT as coordinator of the ANDES and CHANDA EURATOM projects is still operating the CHANDA web in its original address and the web page and document repository of the older ANDES project (ended in 2013) from a CIEMAT web-page supporting nuclear data research (<u>http://win.ciemat.es/</u>).

7. DATA SECURITY

The IAEA support to the EXFOR, ENDF and ENSDF libraries has been structured via a contract including a physical distribution of the storage of their data with several geographical locations of the main server data and alternative geographically distributed servers to access the data. The system is operated on remote cloud servers and professional/commercial support is contracted to preserve the security of the data, including protection versus the loss of data and versus the corruption or manipulation of the data.

In addition, several mirrors of the libraries are operated in different parts of the world (Russia, China,...) but accessible from anywhere in the web.

As for the SANDA web, the web-page and the associated data and document repository is managed by the DRUPAL system, the full set of files and directories are back-up daily and are operated on a dedicated web-server inside a closed internet space. All these elements provide the security of data integrity and a certain level of security versus external intrusions.