DMP title

Project Name G080022N (FWO DMP) - DMP title

Project Identifier G080022N

Grant Title G080022N

Principal Investigator / Researcher Gerda Neyens

Description Research field: structure of exotic nuclei. Research questions: how does the quantum structure of nuclei (isotopes) with an unusual neutron-to-proton ratio change as a function of N and Z. This is investigated by measuring the hyperfine structure of these isotopes. These short-lived isotopes are produced at ISOLDE-CERN, where the experiments are performed using the CRIS and the COLLAPS collinear laser spectroscopy setups. Purpose of the data: get information on the nuclear spin, nuclear moments and charge radius, to compare to predictions by nuclear theories, in order to validate or reject.

Institution KU Leuven

1. General Information

Name applicant

Gerda Neyens

FWO Project Number & Title

G080022N, **Laser spectroscopy near very exotic doubly-magic nuclei to benchmark state-of-the art nuclear theories.**

Affiliation

- KU Leuven

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Type of data	Format	Volume	Source
Laser spectroscopy data	.csv	100 GB	Combined data streams from the various pieces of apparatus in the laboratory
Digital logbook	Microsoft OneNote		
Beam line and ion source simulations	Various Simion and COMSOL outputs	5 GB	Simion and COMSOL
Beam line and ion source design	.stp	1 GB	Autodesk Inventor
Analysis results	Output files from SATLAS and .xlsx	1 GB	SATLAS data analysis software and summary files in excel

The data format at CRIS has been designed to have optimal compatibility with the popular python packages 'pandas' and 'numpy', used for the analysis of CRIS data using the SATLAS package (program written and maintained by our team). The data ressemble the ROOT structure, making them easily convertable to be of use with other analysis codes.

The CRIS data sets combine measurements from all devices relevant to a CRIS experiment: the channels of the wavemeter(s), the ion counts following discrimination and digitization using a TimeTagger4-2G time digitizer card, the acceleration voltage that determines the kinetic energy of the beam, the intensity of the signal propagating through our Fabry-Perot interferometer (FPI), and the fast-skew-time variable voltage supplied to an acceleration electrode used to modify the kinetic energy of the beam when in voltage-scanning mode.

The data is collected in separate .csv files, each of which contains only the data relevant to a single physical device. Each .csv file is organized in columns: the first column contains the real-time timestamp with respect to the beginning of the UNIX epoch, corresponding to the recording of the data line. The remaining columns contain the values of the different components of each device measured at the specified timestamp.

For the WS-U wavemeter, these columns are the measurements of Channels 1, 2, 3, and 4. For the WS-6 wavemeter, the only column contains the values of Channel 1.

For the digitizer, called 'tagger' in our DAQ system, each data line corresponds to a single count. The columns correspond to the number of triggers the digitizer has received thus far, the unique time-of-flight stamp of the count, the total number of counts the digitizer has recorded since the last trigger, and the physical channel on the digitzer that the count was recorded from.

For the FPI, the device-specific columns correspond to the measurement of the voltage generated from the photodiodes at the exit of the interferometer, measured at the given timestamp (four channels are included).

For the general acceleration voltage determined by the extraction potential of the ISOLDE facility equipment, titled as 'iscool' in our DAQ, the only device-specific column is a measurement of the voltage of the extraction electrode, after reduction with a voltage divider. The voltage is measured with a LAN-connected facility voltmeter that is remotely monitored by our DAQ system.

For the variable acceleration voltage, the device-specific columns correspond to the set voltage given by our control software, and the measured voltage as appearing on a calibrated voltmeter following a voltage divider, sitting next to our beamline.

Following an experiment, all the separate .csv files are imported simultaneously and combined into a single python pandas dataframe using each datalines timestamp column to overlap the measurements of each device in time. In the course of each experiment, both using radioactive beam from the ISOLDE facility and using stable beam from our laser-ablation ion source, all experimental conditions are monitored in a digital logbook using Microsoft OneNote, which is then exported into .pdf format and attached to the folders containing the data sets. The data files are then compressed to .zip format and saved on the /EOS network drive hosted by CERN, as well as the specified repositories at KU Leuven and the University of Manchester.

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference:

Short description of the kind of personal data that will be used:

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and

valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

2. Simulation work: the simulation code version and the beam line design version will be associated with each generated data set. The number of particles generated and their initial profile will also be saved in standard SimIon formats.
3. Offline data: the source of the reference material will be systematically noted in the logbook. The laser light generation conditions and laser light operation parameters will also be documented. The detector settings (voltage, connection mode) will be noted. Finally, each data set will be referred to in terms of the isotope mass and laser wavelength.
4. Online laser data: the production protocol including target number, target and ion source currents will be recorded for the entire data set; the laser light production parameters will be described. For each independent file, the accelerator settings (supercycle, proton current), the isotope mass and scanned laser frequency will be written in each file, as well as the scan conditions (from & to wavenumbers, step size, scan speed). The laser light power will be frequently logged for possible offline correlations during analysis. All facility parameters such as separator settings, proton cycles, acceleration voltages (very low-precision), etc are recorded in real-time by the CERN 'TIMBER' service, and are accessible at all times.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

There is no existing standard for the data type that will be collected, besides for the decay spectroscopy data, where the acquisition system (CAEN COMPASS) has a standard format for the trapezoid parameters & pulse shape discrimnator parameters. However, even that metadata is highly specific to the use of that software and is not an industry standard.

The detailed metadata were described above.

5. Data storage and backup during the FWO project

Where will the data be stored?

Data generated at KU Leuven (from data analysis) will be stored on the IKS server. Data from CERN experiments will be saved indefinitely on the /EOS network drive hosted by CERN, as well as the specified repositories at KU Leuven and the University of Manchester.

All members of the CRIS collaboration can be provided with the raw data files as per their request, according to the rules in the CRIS Collaboration Memorandum of Understanding. However, accessing the data folders in the /EOS drive storage is left to the CRIS Data Manager (Dr. Ágota Koszorús) and specific members with approved access.

How is backup of the data provided?

The IKS data server is backed up daily across multiple machines and weekly across multiple locations. The ICTS team from the Department of Physics and Astronomy is in charge of this.

All data from CERN are also backed-up on the IKS data server and undergo the same additional aforementioned backups.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

The IKS continuously upgrades its storage capacity to allways accommodate the growing demand for data storage.

CERN provides unlimited storage capacity to user experiments, and therefore the data collected by the CRIS experiment are permanently stored on CERN network drives accessible only by authorized members of the CRIS collaboration that have been granted access. Access to the file folders is password-protected requiring CERN credentials and access is granted only through the CERN access service 'AdAMS' by the CRIS Data Manager (Dr. Ágota Koszorús).

What are the expected costs for data storage and back up during the project? How will these costs be covered?

The costs for data storage at IKS are covered by all ZAP from the institute and are therefore not evaluated per project. When necessary, the expansion of the storage capacity is performed with funding from the active projects. The relative investment is minimal on the overal budget.

At CERN, access to storage is completely free, offered as part of the basic infrastructure that CERN provides for research.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Data stored on the CERN network drives are accessible only by authorized members of the CRIS collaboration that have been granted access. Access to the file folders is password-protected requiring CERN credentials and access is granted only through the CERN access service 'AdAMS' by the CRIS Data Manager (Dr. Ágota Koszorús).

All members of the CRIS collaboration can be provided with the raw data files as per their request, according to the rules in the CRIS Collaboration Memorandum of Understanding. However, accessing the data folders in the /EOS drive storage is left to the CRIS Data Manager (Dr. Ágota Koszorús) and specific members with approved access.

At KU Leuven, only members of my own research team have access to the data and a clear Data Management Plan guideline has been deployed within the group to ensure that raw data are never overwritten.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All data will be retained up to (and beyond) the 5-year term.

Where will the data be archived (= stored for the longer term)?

At KU Leuven, the raw data and the secondary data are kept in a shared folder for the research team on the IKS data server - accessible even after the end of contract of temporary personnel (PhD, post-doc). Retrieving these data is possible through contacting active members of the research team, either the PI or others.

At CERN data are kept indefinitely.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

IKS regularly invests in data storage hardware for all IKS members. The costs are shared between all 5 ZAP and spread over all projects as required at the time of investment. The costs to this project are a few% of the total budget.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?

The primary data follow an experiment specific format that is not appropriate for sharing broadly but may be shared upon request to the PI. Secondary data of relevance to the community will be shared on Zenodo under a CCBY license.

Where/how will the data be made available for reuse?

- In an Open Access repository
- Upon request by mail

Analysis codes will be released on GitHub.
The primary data will be available on request to the PI.
The secondary data relevant to the community will be made available on Zenodo.

When will the data be made available?

- Upon publication of the research results

Who will be able to access the data and under what conditions?

Due to its experiment specific format, the primary data is only accessible for experts with understanding of the content. Sharing of data with persons not a member of the CRIS collaboration will be possible upon request, either to the PI of the experiment or to the CRIS Collaboration Board.

The secondary data relevant to the community will be freely accessible on Zenodo.

What are the expected costs for data sharing? How will the costs be covered?

Data sharing is embedded in the best practices of the research group members as part of their individual data management plan and does incure any cost.

8. Responsibilities

Who will be responsible for data documentation & metadata?

For preparatory studies (simulations, designs, scheme research), the documentation and metadata are the responsible of the researcher (PhD, post-doc) who performs the research.

For experiments using the accelerator, the experiment is running 24/7 in shifts of 8h. For each shift, a person within the team is named responsible for the logging of all information, including documentation and metadata.

Who will be responsible for data storage & back up during the project?

Within the research team, a data curator is named to ensure that the data are managed in a systematic way. This is currently Michail Athanasakis, PhD student.

At IKS, the data server is maintained and backed up by the ICT team of the Department of Physics and Astronomy, composed of 4 persons.

At CERN, the data storage and back up is the responsibility of the CRIS Local Coordinator, currently Dr Agi Koszorus, CERN Fellow.

Who will be responsible for ensuring data preservation and reuse ?

Data preservation within the research team is the responsibility of the data curator (currently Michail Athansakis). Data servers are maintained by the ICT team of the Department of Physics and Astronomy.

Reuse of the data at KU Leuven is the responsibility of the PI. Reuse of the data at CERN is the responsibility of the CRIS Executive Board, to which the PI belongs.

Who bears the end responsibility for updating & implementing this DMP?

The PI bears the end responsibility of updating & implementing this DMP.