



Internship Title

Internship – Exploration of statistics and Machine Learning techniques for nuclear reaction data curation

Duration of Internship

3 to 12 months

Organizational Setting

Department: Nuclear Sciences and Applications (NA)

Division: Physical and Chemical Sciences (NAPC)

Section: Nuclear Data Section (NDS)

Unit: Nuclear Data Development Unit (NDDU) and Nuclear Data Services Unit (NDSU)

Main Purpose

The main purpose of the internship is to prototype modern methods and software tools for the nuclear data development.

During the duration of the internship with the Nuclear Data Section, the intern will explore the potential of statistical algorithms and Machine Learning (ML) techniques as tools to assist in the curation of nuclear reaction data from databases, such as identifying anomalies and enhancing datasets with meta-data through statistical imputation methods or natural language processing of relevant literature. The intern will be also getting acquainted with general principles of:

- Nuclear data as a foundation to pure and applied nuclear science and engineering
- Statistical modelling in the context of nuclear data
- Data handling with experimental nuclear reaction databases such as EXFOR as well as evaluated nuclear data libraries such as FENDL, TENDL, Neutron Data Standards
- Tools to interact with and update nuclear data

Tasks / Key Results Expected

Depending on progress, interest and priorities, a selection of the following:

- Create prototypes for the validation and augmentation of nuclear databases, such as
 - Creating scripts to flag and visualize anomalies in databases
 - Designing scripts to identify relevant literature with nuclear data from online literatures
 - Integrating these scripts into evaluation pipelines for more robust data evaluations
 - Constructing scripts with Natural Language Processing to improve database quality or generate database entries automatically



Knowledge, Skills and Abilities

- Excellent programming skills in Python
- Basic knowledge of statistics and machine learning techniques.
- A good foundation in linear algebra and numerical optimization is an asset.
- Experience with databases with scientific data is an asset.
- Good communication and teamwork skills; capable of working independently and proactively.

Qualifications and Experience

- University degree in Computer Science, Data Science, Physical Sciences, Mathematics, Engineering, or a related discipline.
- Experience in the application of Python to solve problems involving or related to scientific data.

Potential Institutions/Organizations that can be reach out to in order to attract potential applicants

- Technical Universities.
- Research Institutes.

Internships

The IAEA accepts a limited number of interns each year. The internships are awarded to persons studying towards a university degree or who have recently received a degree (see Internship web pages for further details).

The purpose of the programme is:

- To provide interns with the opportunity to gain practical work experience in line with their studies or interests, and expose them to the work of the IAEA and the United Nations as a whole.
- To benefit the IAEA's programmes through the assistance of qualified students specialized in various professional fields.
- The duration of an internship is normally not less than three months and not more than one year.

Applicant Eligibility

- Candidates must be a minimum of 20 years of age and have completed at least three years of full-time studies at a university or equivalent institution towards the completion of a first degree.
- Candidates may apply up to one year after the completion of a bachelor's, master's or doctorate degree.
- Candidates must not have previously participated in the IAEA's internship programme.
- Excellent written and spoken English essential; fluency in any other IAEA official language (Arabic, Chinese, French, Russian) an asset.
- Candidates must attach two signed letters of recommendation to their application.