IAEA
International Atomic Energy Agency
Atoms for Peace and Development

EXFOR PDFs archiving improvements

Alejandra Martinez
Nuclear Data Section, IAEA, Vienna, Austria

EXFOR Workshop, 03 – 06 December 2024

IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Objective

Propose an improvement to the current process of the EXFOR PDF database update

Value proposition:

+ Usability
+ Accessibility
+ Maintainability
+ Security

IAEA
International Atomic Energy Agency
Atoms for Peace and Development

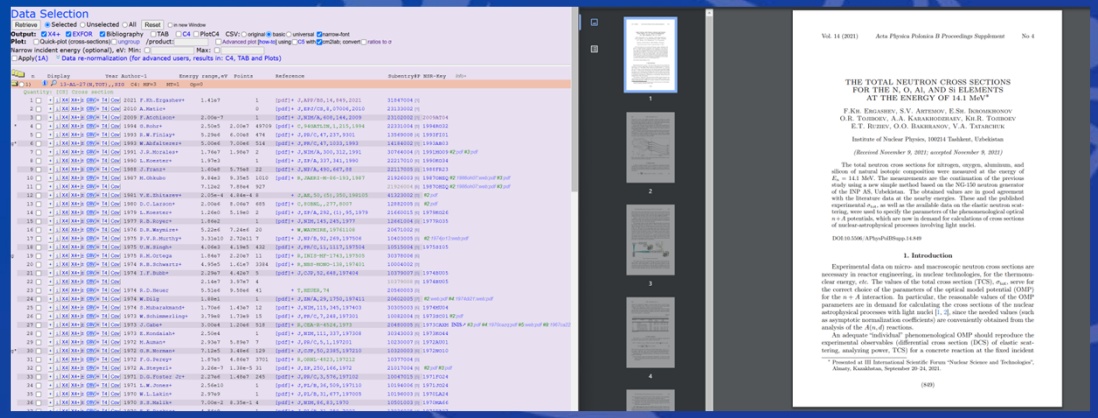
Introduction

The EXFOR web retrieval application currently offers, to a very restricted set of users, the display of publications of various types such as journals, conferences, reports, thesis, etc., in PDF format.

This is a useful feature for EXFOR compilers and some evaluators to check and compare the data.

Sources of the publications:

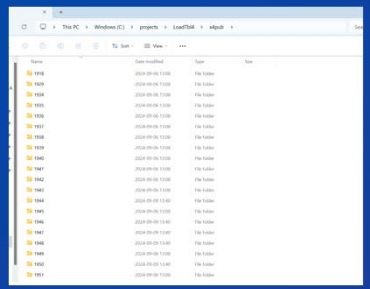
- NSR
- EXFOR



IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Current problem – update of the internal PDF database

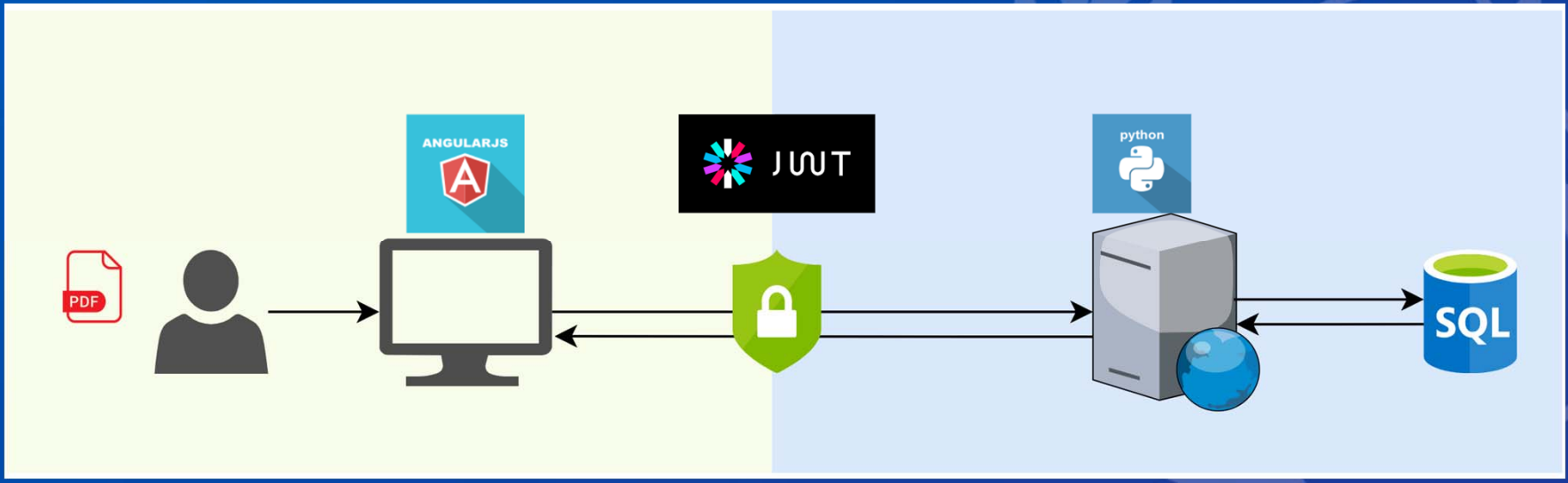
- Some PDFs are updated while others are not.
- Depends on user folders and OS.
- Several scripts involved.
- No version control.
- Low security.
- Not ideal: PDF blobs are stored on the database
  200.57 GB



IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Proposed solution for current system – EXFOR PDFs

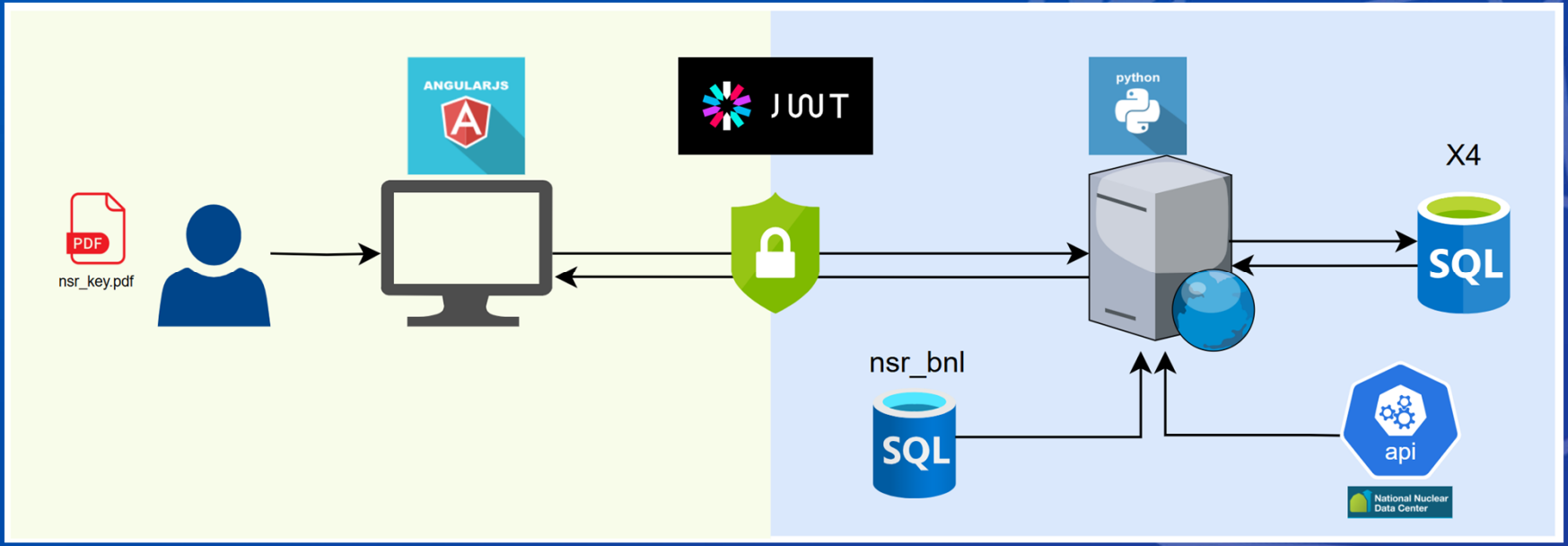
- Angular frontend
- JWT Authentication
- Python backend – APIs



IAEA
International Atomic Energy Agency
Atoms for Peace and Development

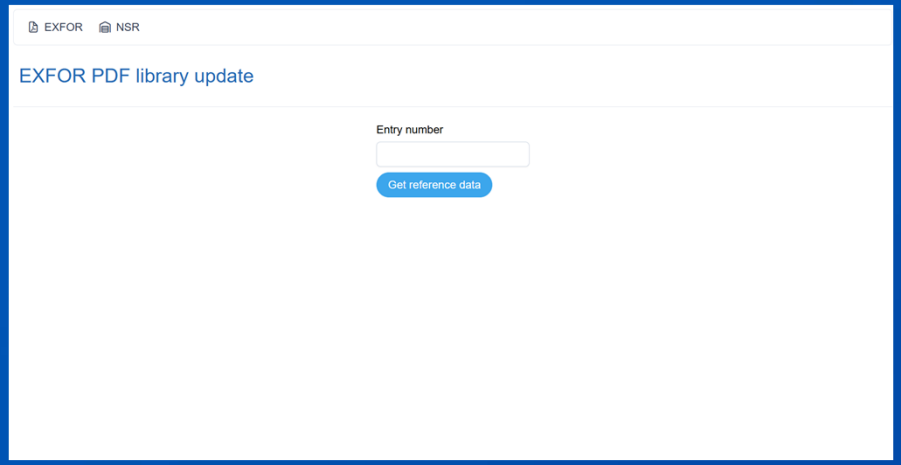
Proposed solution for current system – NSR PDFs

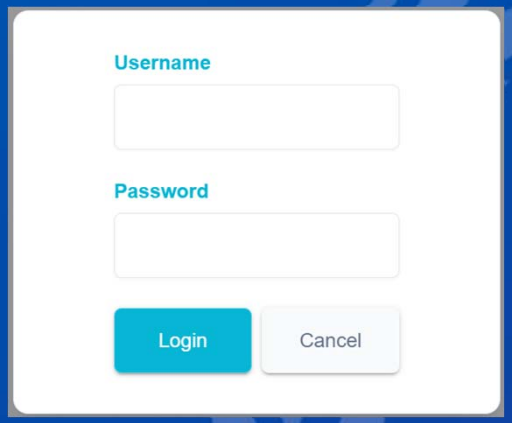
- Angular frontend
- JWT Authentication
- Python backend – APIs

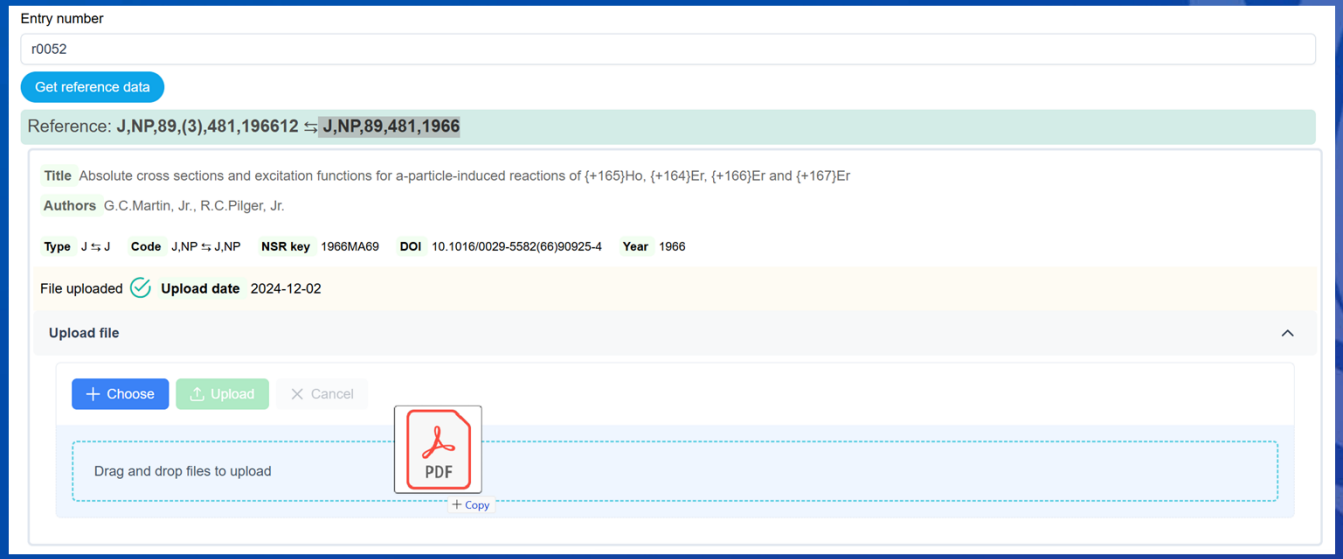


IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Proof of concept – EXFOR PDFs

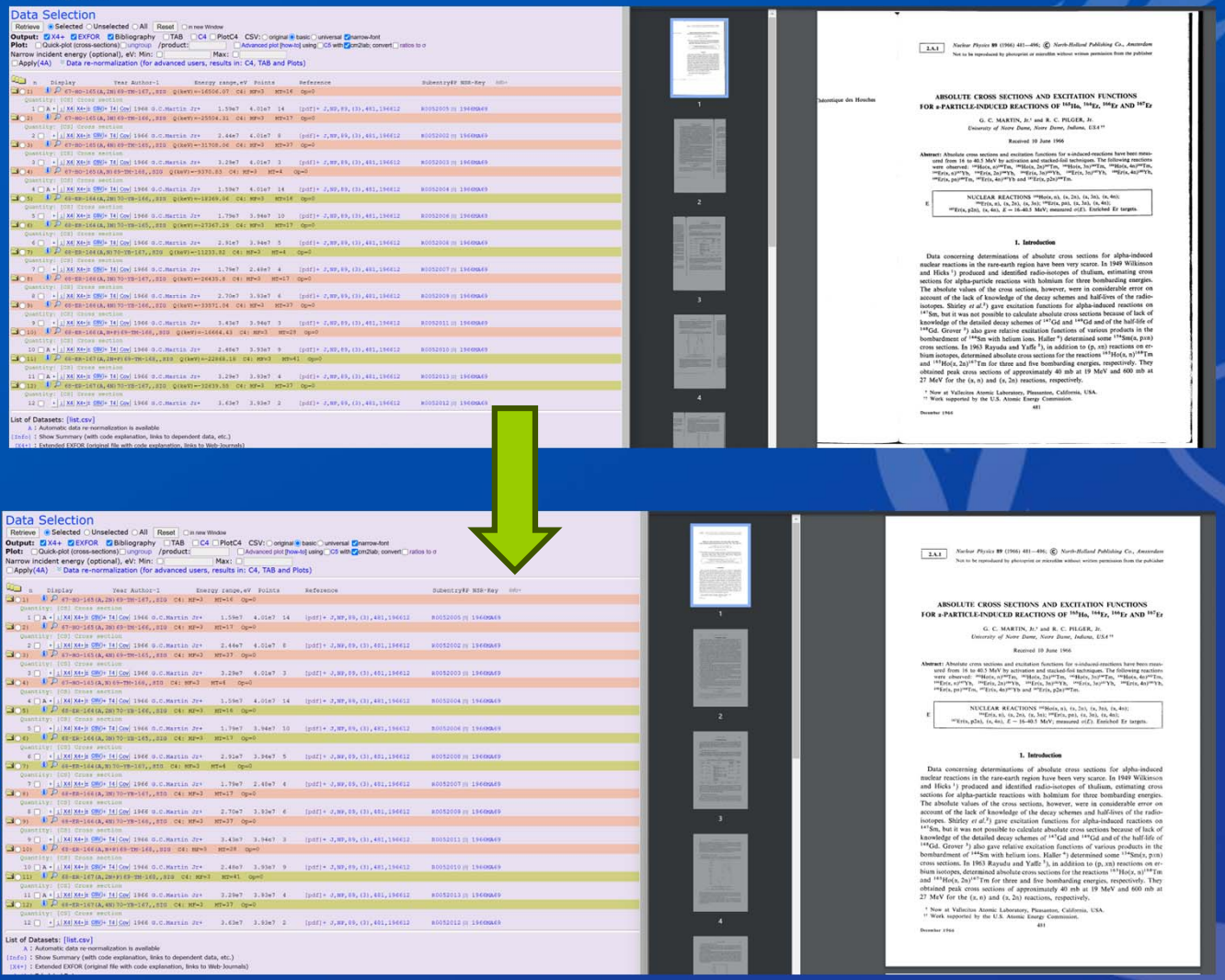






IAEA
International Atomic Energy Agency
Atoms for Peace and Development

PDF update



IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Foreseen advantages

- Clearer and friendlier user process
- Reduction of the time it takes to update a PDF
- Internal web access
- Increased level of security
- Improved maintainability of code
- Reduce database size

IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Next steps

- Change storage

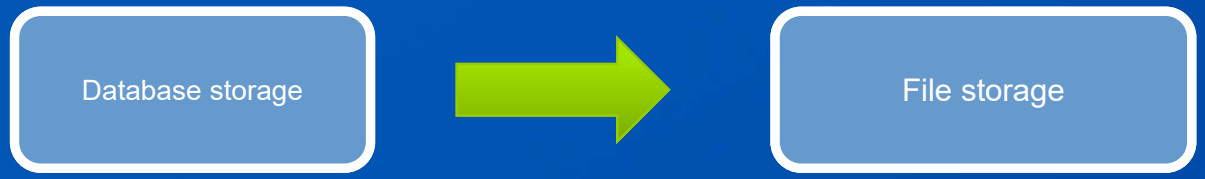
- Improve pipeline for the propagation of updates

- File versioning and traceability

* To be discussed with the team…

IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Change of storage



- Download current PDFs from exfor table (consider the current binary obfuscation)
- Make an inventory of references and PDFs
- Clean duplicates or any other unused files
- Create file reference / metadata table
- Create folder in file system
- Create global parameter for folder (could be a table for multiple folders / folder versions)
- Store PDFs in folder
- Populate new table with file references

* To be discussed with the team…

IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Thank you!