



# Modernizing Workflow for NSR Updates

Benjamin Shu

November 2024



# Current State: Data Model

Limited by its adherence to “exchange” format:

- Entries need to be pre-processed before use in code

```
<KEYNO    >2002RI06
<HISTORY  >A20020708
<CODEN    >JOUR NUPAB...
<REFERENCE>Nucl.Phys. A...
<AUTHORS  >M.V.Ricciardi...
<TITLE    >Experimental Study...
<KEYWORDS>NUCLEAR REACTIONS...
<SELECTRS>T:1H;A. ...
```

# Current State: Data Model

Limited by its adherence to “exchange” format:

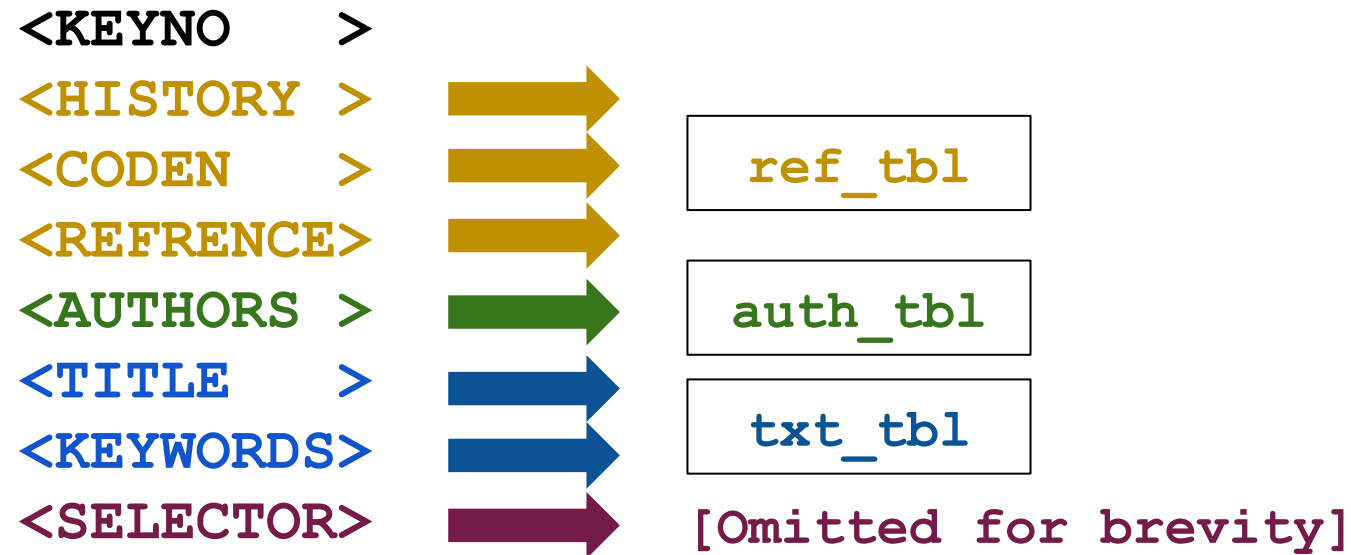
- Entries need to be pre-processed before use in code

<KEYNO >	2002RI06	Unique ID
<HISTORY >	A20020708	Entry/modification dates
<CODEN >	JOUR NUPAB...	Source type, journal
<REFERENCE>	Nucl.Phys. A...	Citation information
<AUTHORS >	M.V.Ricciardi...	List of authors
<TITLE >	Experimental Study...	Formatted title
<KEYWORDS>	NUCLEAR REACTIONS...	Sentences about content
<SELECTRS>	T:1H;A. ...	Structured search keys

# Current State: Data Model

Limited by its adherence to “exchange” format:

- Text files need to be synchronized with SQL tables



# Current State: Data Model

Limited by its adherence to “exchange” format:

- Custom parsing required for LaTeX-like syntax

<b>Title</b>	First half-life measurement of a low-lying isomer in $^{37}\text{Si}$
<b>Keywords</b>	<ul style="list-style-type: none"><li>• RADIOACTIVITY <math>^{37, 38}\text{Al}(\beta^-)</math>, <math>(\beta^-n)</math>[from <math>^{48}\text{Ca}(^9\text{Be}, X)</math>, <math>E=140</math> MeV/ nucleon]; measured <math>E_\gamma</math>, <math>I_\gamma</math>, <math>I_\beta</math>, <math>\gamma\gamma</math>-coin, <math>\beta\gamma</math>-coin; deduced <math>T_{1/2}</math>. <math>^{37}\text{Si}</math>; deduced levels, <math>J</math>, <math>\pi</math>, <math>T_{1/2}</math> of isomeric states at 68- and 156-keV, <math>B(M1)</math>, matrix elements for low-lying ground-state M1 transitions. <math>^{36, 38}\text{Si}</math>; deduced <math>E_\gamma</math>, <math>I_\gamma</math>, transition intensities. Comparison to other experimental data and shell model calculations using the SDPF-MU, SDPF-U-SI, and FSU interactions. <math>\text{CeBr}_3</math> based implantation detector with position-sensitive photomultiplier tube surrounded by 16 segmented Ge detectors (SeGa) and 15 <math>\text{LaBr}_3</math> detectors at Coupled Cyclotron Facility (NSCL, MSU).</li></ul>

<https://www.nndc.bnl.gov/nsr/?search-type=keynumber&key=2023OG04>

# Current State: Data Model

Limited by its adherence to “exchange” format:

- Custom parsing required for LaTeX-like syntax

```
<KEYWORDS>RADIOACTIVITY  $\{+37\}$ ,  $\{+38\}$ Al ( $|b\{+-}\}$ ), ( $|b\{+-}\}n$ ) [from  $\{+48\}$ Ca ( $\{+9\}$ Be, X), E=140 MeV/nucleon]; measured E|g, I|g, I|b, |g|g-coin, |b|g-coin; deduced  $T\{-1/2\}$ .  $\{+37\}$ Si; deduced levels, J, |p,  $T\{-1/2\}$  of isomeric states at 68- and 156-keV, B(M1), matrix elements for low-lying ground-state M1 transitions.  $\{+36\}$ ,  $\{+38\}$ Si; deduced E|g, I|g, transition intensities. Comparison to other experimental data and shell model calculations using the SDPF-MU, SDPF-U-SI, and FSU interactions. CeBr $\{-3\}$  based implantation detector with position-sensitive photomultiplier tube surrounded by 16 segmented Ge detectors (SeGa) and 15 LaBr $\{-3\}$  detectors at Coupled Cyclotron Facility (NSCL, MSU)
```

<https://www.nndc.bnl.gov/nsr/?search-type=keynumber&key=2023OG04>

# Current State: Data Model

Limited by its adherence to “exchange” format:

- No way to uniquely identify author names
  - No space for ORCID (or other fields)

```
<AUTHORS >T.Redpath, P.Gueye,  
T.Baumann, B.A.Brown,  
A.Cunningham, P.A.DeYoung,  
N.Frank, C.R.Hoffman,  
A.N.Kuchera, B.Monteagudo  
Godoy, C.Persch, A.Revel,  
W.F.Rogers, M.Thoennessen,  
J.A.Tostevin, D.Votaw
```

```
MariaDB [nsr_test]> SHOW COLUMNS IN auth_both;  
+-----+-----+-----+-----+-----+-----+  
| Field | Type          | Null | Key | Default | Extra |  
+-----+-----+-----+-----+-----+-----+  
| akey  | int(11)       | NO   | MUL | NULL    |       |  
| aname | varchar(32)   | NO   | MUL | NULL    |       |  
| a1    | char(1)       | NO   |     |         |       |  
| a2    | char(1)       | NO   |     |         |       |  
+-----+-----+-----+-----+-----+-----+  
4 rows in set (0.001 sec)
```

# Current State: Database Hosting

Only exists as a MariaDB database instance


- Limited ability to track updates
  - *Who* made the last update?
  - *What* did they change?

At time of writing,  
this information is  
**NOT** being tracked




Aug 28, 2024

---

 Updated json/entries (again) after adding code to check for XUNDL datasets.  
Benjamin Shu authored 2 weeks ago

---

 Updated json/entries after adding code to store EXFOR keys.  
Benjamin Shu authored 2 weeks ago

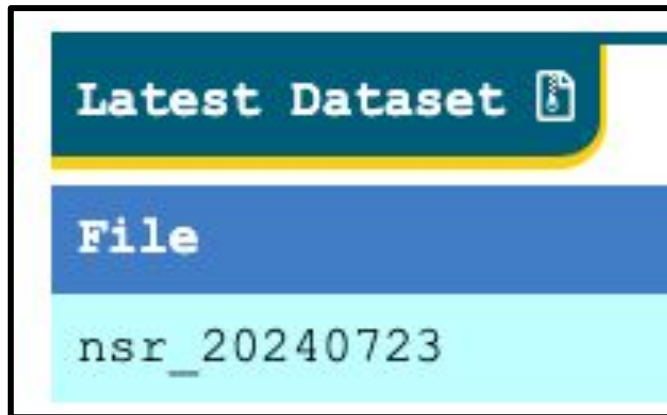
Changelog of updates to NSR



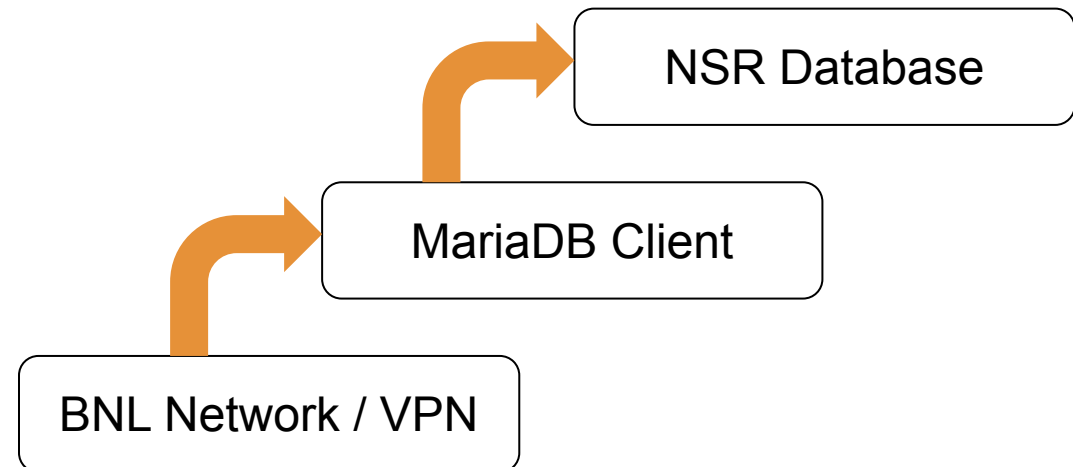
# Current State: Database Hosting

Only exists as a MariaDB database instance

- No offline version for developers
  - Latest archive from 2024-07-23
  - Data files only accessible through SQL client or website



<https://www.nndc.bnl.gov/nsrarchivals/>



# Current State: Database Hosting

Only exists as a MariaDB database instance

- Table schema cannot be changed easily

```
MariaDB [nsr_test]> SHOW TABLES;
+-----+
| Tables_in_nsr_test |
+-----+
| auth_alias          |
| auth_both           |
| auth_dict           |
| auth_tbl            |
| az_tbl              |
| coden_dict          |
| ensdf_link          |
| exch_tbl            |
| nuc_dict            |
| nuc_tbl             |
| part_dict           |
```

```
  reac_conv
  reac_dict
  reac_out_dict
  reac_tbl
  ref_tbl
  sub_conv
  sub_dict
  sub_tbl
  tempalias
  topic_tbl
  txt_tbl
+-----+
22 rows in set (0.000 sec)
```

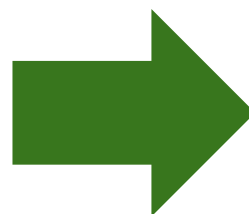
# Step 0: JSON Data Model

[Working]

Based on existing JSON structure used in website

- Preserve text from original “exchange” tags:
  - <AUTHOR>, <TITLE>, <REFERENCE>, etc.

<KEYNO	>	2002RI06
<HISTORY	>	A20020708
<CODEN	>	JOUR NUPAB...



```
{
  "name": "KEYNO",
  "content": "2002RI06"
},
{
  "name": "HISTORY",
  "content": "A20020708"
},
{
  "name": "CODEN",
  "content": "JOUR NUPAB 701 156c"
},
}
```

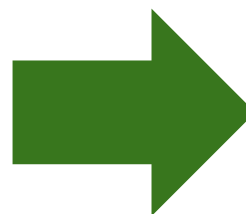
# Step 0: JSON Data Model

[Working]

Based on existing JSON structure used in website

- Parse NSR free text as JSON key-value pairs

```
{
  "name": "KEYNO",
  "content": "2002RI06"
},
{
  "name": "HISTORY",
  "content": "A20020708"
},
{
  "name": "CODEN",
  "content": "JOUR NUPAB 701 156c"
},
}
```



```
"key_number": "2002RI06",
"publication_year": 2002,
"entry_type": "A",
"entry_date": "2002-07-08",
"modification_date": "2002-07-08",
"coden_text": "JOUR NUPAB 701 156c",
"reference_type": "JOURNAL",
"journal_code": "NUPAB",
"index_information": "701 156c",
```

# Step 0: JSON Data Model

[Working]

Based on existing JSON structure used in website

- Add fields for data not stored or made explicit

```
{
  "type": "SINGLE_AUTHOR",
  "text": "M.V.Ricciardi",
  "initials": [
    "M",
    "V"
  ],
  "last_name": "Ricciardi"
},
```

```
"nuclide_selectors": [
  {
    "sub_type": "TARGET",
    "z": 1,
    "n": 0,
    "a": 1,
    "element_name": "Hydrogen",
    "element_symbol": "H"
  }
],
```

```
"is_primary": true,
"pdf_available": false,
"xundl_available": false,
"exfor_keys": []
```

# Step 1: Preserving Existing Data

[Working]

Conversion of original exchange files into JSON

- **nsr-pipeline** project - reusable programs for:
  - Downloading exchange files from SQL database
  - Downloading additional information not stored in text files
  - Combining data and printing into new JSON schema

```
Collecting entries: 100% [=====] 249699/249699 (0:09:49 / 0:00:00)

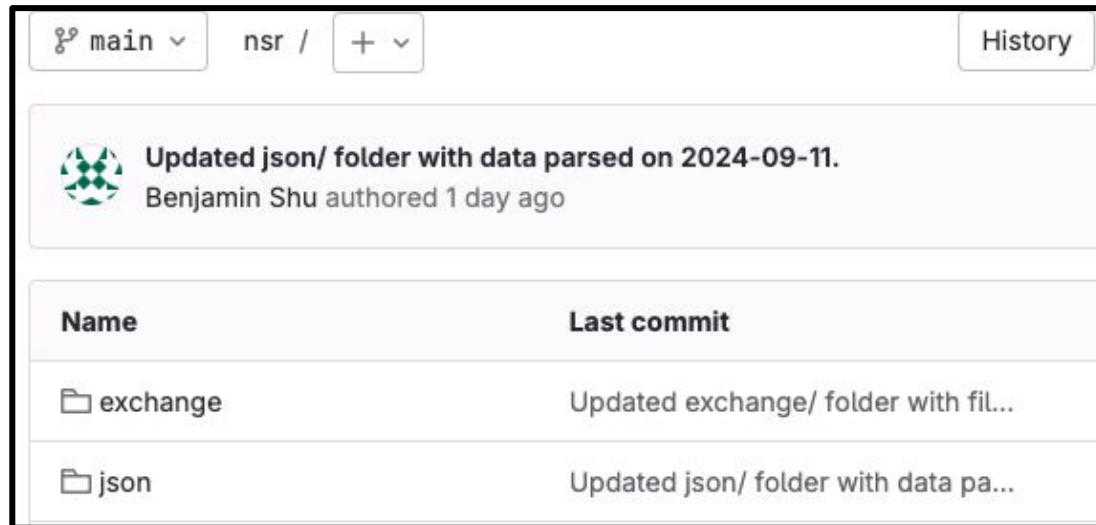
BUILD SUCCESSFUL in 10m 21s
20 actionable tasks: 8 executed, 12 up-to-date
[bshu@development3 combine-exchange-and-ref_tbl]$
```

# Step 1: Preserving Existing Data

[Working]

Conversion of original exchange files into JSON

- Output files stored in GitLab under **databases/nsr**
  - Updating data no longer requires persistent connection
  - Possible public access? (if allowed by BNL)



The screenshot shows a GitLab repository view for the path 'nsr /'. At the top, there are navigation elements: 'main' with a dropdown arrow, 'nsr /' with a plus sign and dropdown arrow, and a 'History' button. Below this is a commit summary: a green circular icon, the text 'Updated json/ folder with data parsed on 2024-09-11.', and 'Benjamin Shu authored 1 day ago'. Below the summary is a table with two columns: 'Name' and 'Last commit'. The table lists two folders: 'exchange' and 'json', each with a corresponding commit message.

Name	Last commit
exchange	Updated exchange/ folder with fil...
json	Updated json/ folder with data pa...

```
databases/nsr
exchange/
...
json/
tags/
documents/
entries/
```

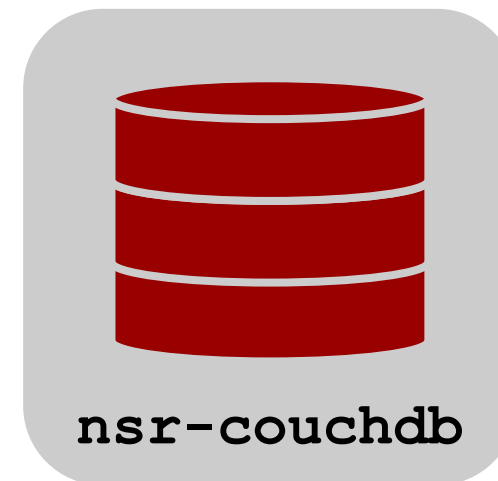
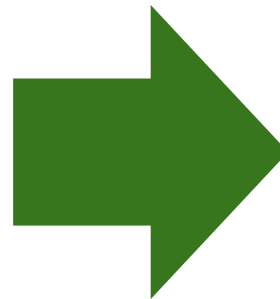
# Step 2: Containerized Databases

[Working]

Use JSON files to build a CouchDB Docker image

- Build container with reusable Python scripts
  - Inserts JSON files and `.js` for database views
  - Can be restarted/rebuilt if necessary
  - Can be redeployed in case of server failure

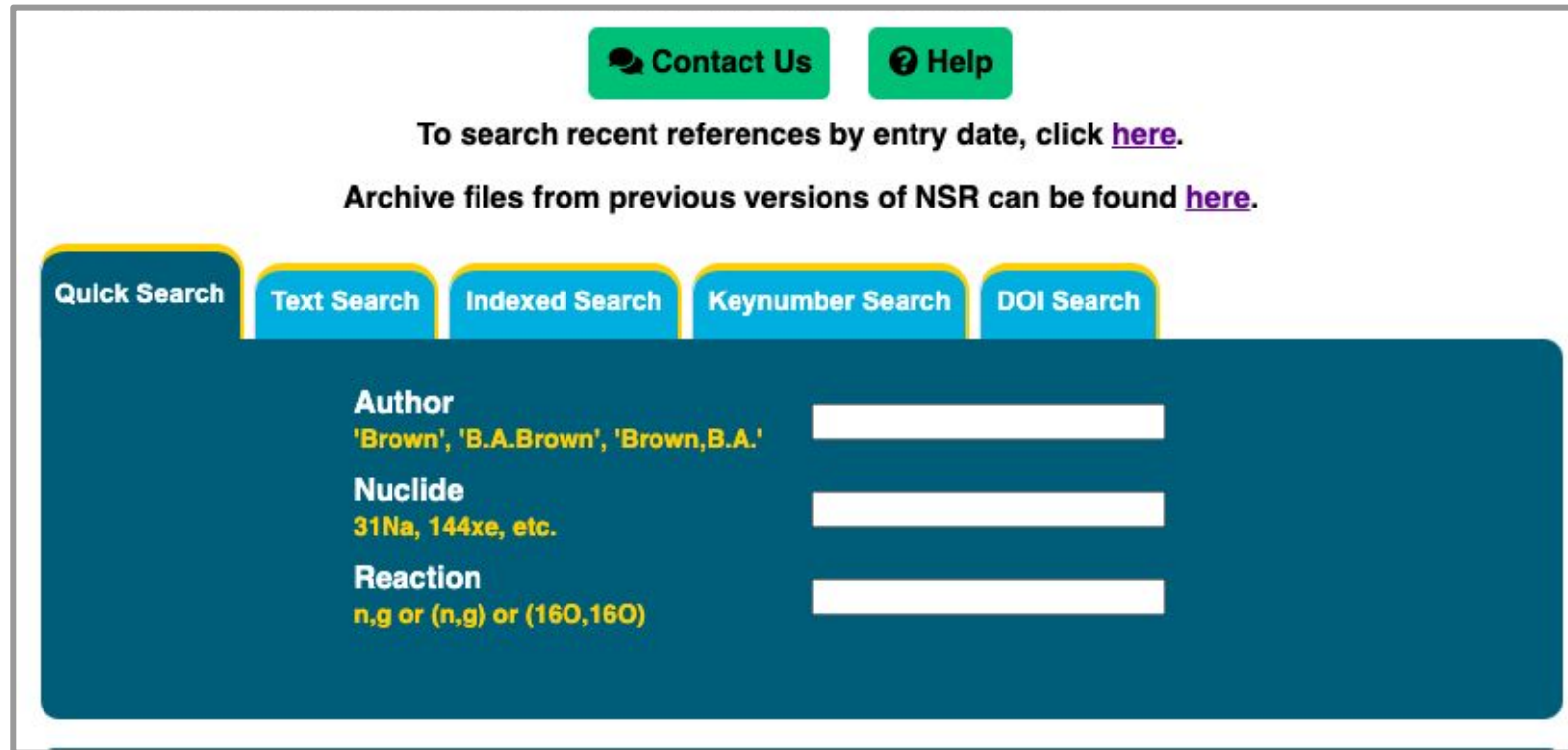
```
databases/nsr/  
design_documents/  
Dockerfile  
load.sh
```





# Step 3: Web Development

**nsr-dev** is now the official NSR website

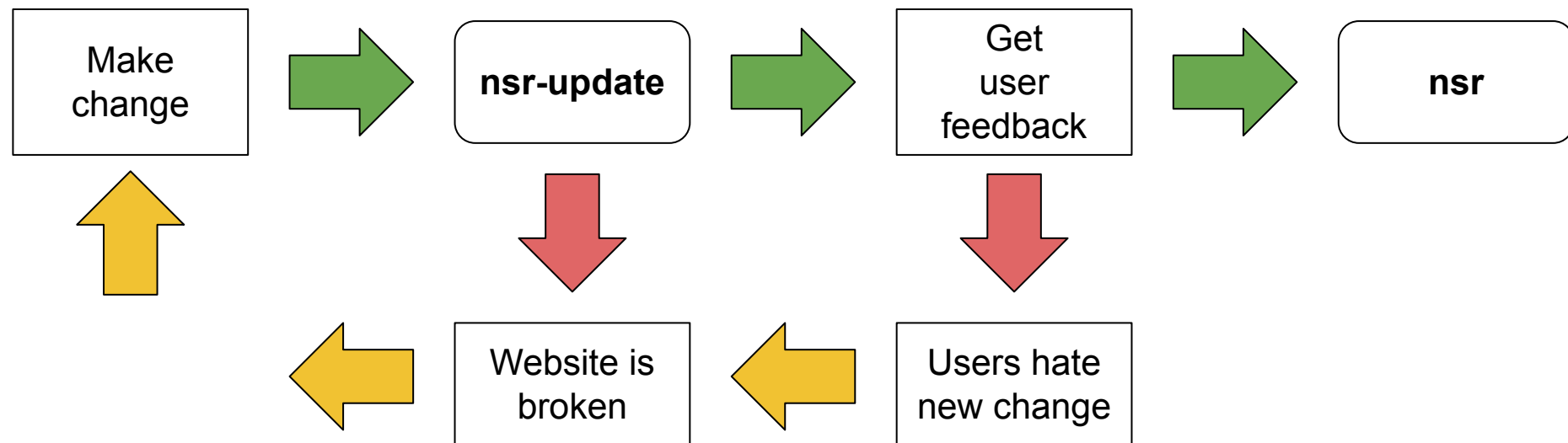


The screenshot shows the top navigation area of the NSR website. At the top, there are two green buttons: "Contact Us" and "Help". Below these, there is a text prompt: "To search recent references by entry date, click [here](#)." and "Archive files from previous versions of NSR can be found [here](#)." Below the text, there are five search options: "Quick Search", "Text Search", "Indexed Search", "Keynumber Search", and "DOI Search". The "Quick Search" option is highlighted with a yellow border. Below the search options, there is a dark blue search form with three input fields. The first field is labeled "Author" and has the example text "'Brown', 'B.A.Brown', 'Brown,B.A.'" below it. The second field is labeled "Nuclide" and has the example text "31Na, 144xe, etc." below it. The third field is labeled "Reaction" and has the example text "n,g or (n,g) or (160,160)" below it.

# Step 3: Web Development

Separate web application for testing CouchDB backend

- Currently live at: <https://www.nndc.bnl.gov/nsr-update/>

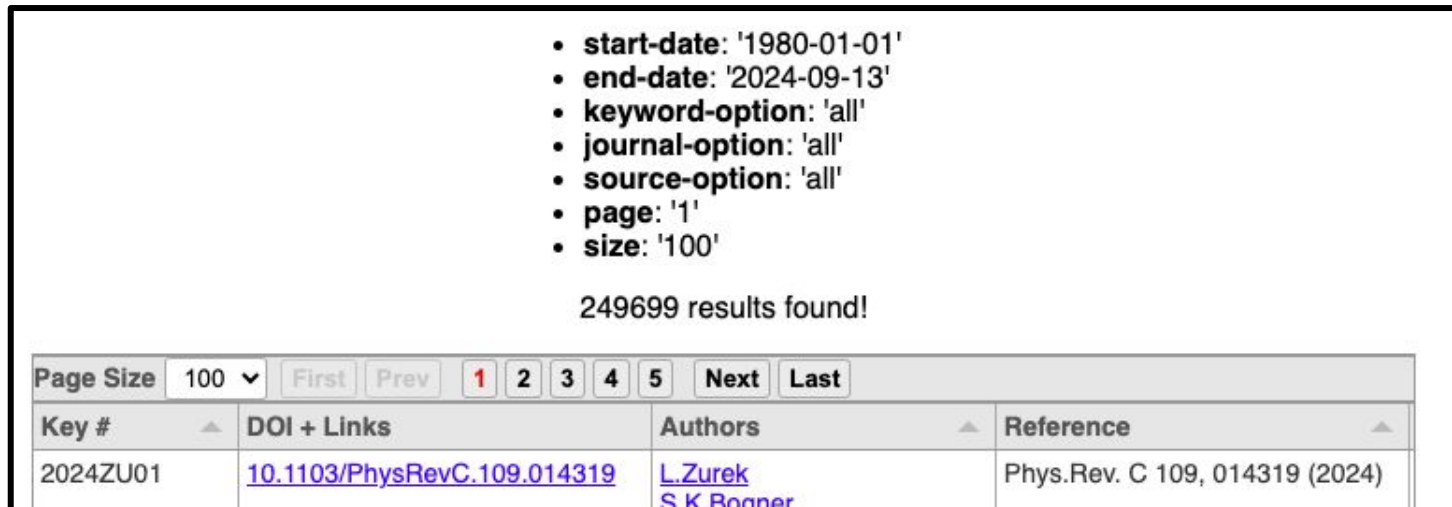


# Step 3: Web Development

[Deployed]

Remote pagination for search results

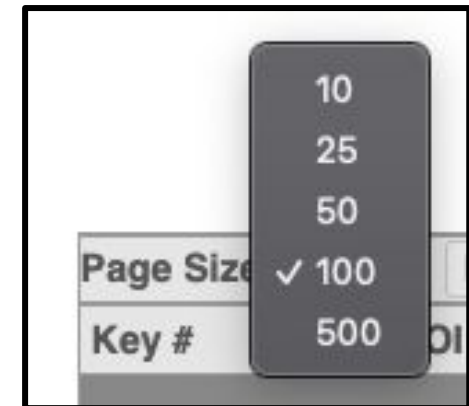
- Shortens response time by sending results in chunks



• start-date: '1980-01-01'  
• end-date: '2024-09-13'  
• keyword-option: 'all'  
• journal-option: 'all'  
• source-option: 'all'  
• page: '1'  
• size: '100'

249699 results found!

Page Size	100	First	Prev	1	2	3	4	5	Next	Last
Key #	DOI + Links	Authors	Reference							
2024ZU01	<a href="https://doi.org/10.1103/PhysRevC.109.014319">10.1103/PhysRevC.109.014319</a>	<a href="#">L. Zurek</a> <a href="#">S. K. Bogner</a>	Phys. Rev. C 109, 014319 (2024)							



<https://www.nndc.bnl.gov/nsr/recent-references.html?start-date=1980-01-01&end-date=2024-09-13>

# Step 3: Web Development

[Deployed]

Bookmarkable URLs & query parameters

- Page updates its own URL after each search

1)

```
• search-type: 'quick'  
• author: 'b.a.brown'  
• nuclide: '31Na'  
• page: '1'  
• size: '100'  
  
5 results found!
```

2)

```
• search-type: 'quick'  
• reaction: 'n,g'  
• page: '1'  
• size: '100'  
  
6706 results found!
```

3)

```
• search-type: 'quick'  
• nuclide: '40Ca'  
• reaction: 'n,g'  
• page: '1'  
• size: '100'  
  
20 results found!
```

- 1) <https://www.nndc.bnl.gov/nsr/?search-type=quick&author=b.a.brown&nuclide=31Na>
- 2) <https://www.nndc.bnl.gov/nsr/?search-type=quick&reaction=n,g>
- 3) <https://www.nndc.bnl.gov/nsr/?search-type=quick&nuclide=40Ca&reaction=n,g>

# Step 3: Web Development

[Deployed]

Indexed Search “Browse” pages now open in a modal window

National Nuclear Data Center

Brookhaven National Laboratory

### Browse Nuclides

**Search By Z (Required)**

Element:  
Neutron

Proton # (Z):  
 →

Search By A

**Mass # (A):**  
 →

Search

Close

# Step 3: Web Development

[Deployed]

Search page now automatically scrolls down to results table

Printer-Friendly

- **search-type:** 'quick'
- **author:** 'b.a.brown'
- **page:** '1'
- **size:** '100'

908 results found!

Note: Mac users may need to [enable right-clicking](#) to use table features.

Page Size	100	First	Prev	1	2	3	4	5	Next	Last
Key #	DOI + Links	Authors	Reference							
2024VY01	<a href="https://doi.org/10.1103/PhysRevC.110.035803">10.1103/PhysRevC.110.035803</a> <a href="#">PlumX Metrics</a>	<a href="#">E.C.Vyfers</a> <a href="#">V.Pesudo</a> <a href="#">S.Triambak</a> <a href="#">M.Kamil</a>	Phys.Rev. C 110, 035803 (2024)							

# Step 3: Web Development

[Not on main NSR yet]

Auto-complete inputs for Quick Search

- Suggestions filled in using lists of all authors, nuclei, reactions



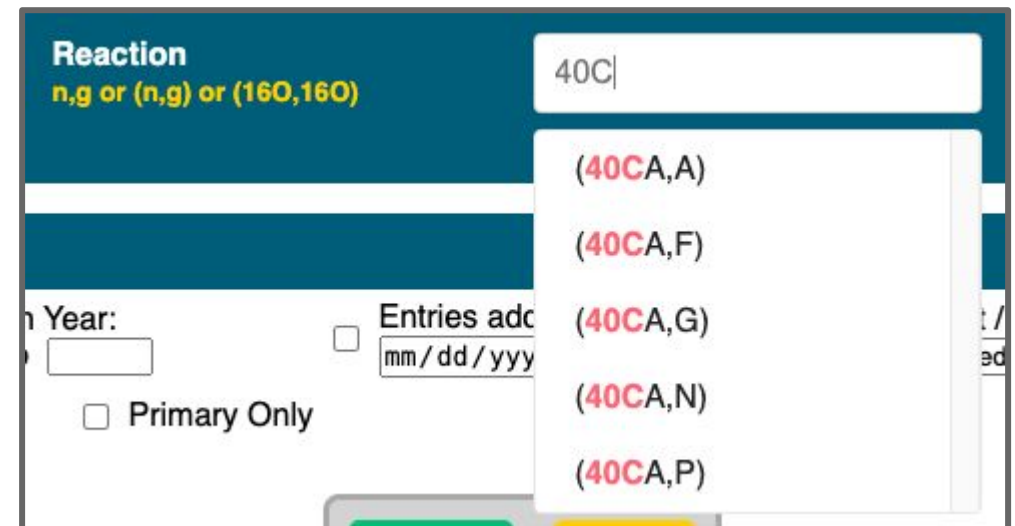
Author  
'Brown', 'B.A.Brown', 'Brown,B.A.'

Nuclide  
31Na, 144Xe, etc.

Reaction  
n,g or (n,g) or (16O,16O)

B.A

B.AI  
B.AN  
B.AAS  
B.A.LI  
B.A.LI



Reaction  
n,g or (n,g) or (16O,16O)

40C

(40CA,A)  
(40CA,F)  
(40CA,G)  
(40CA,N)  
(40CA,P)

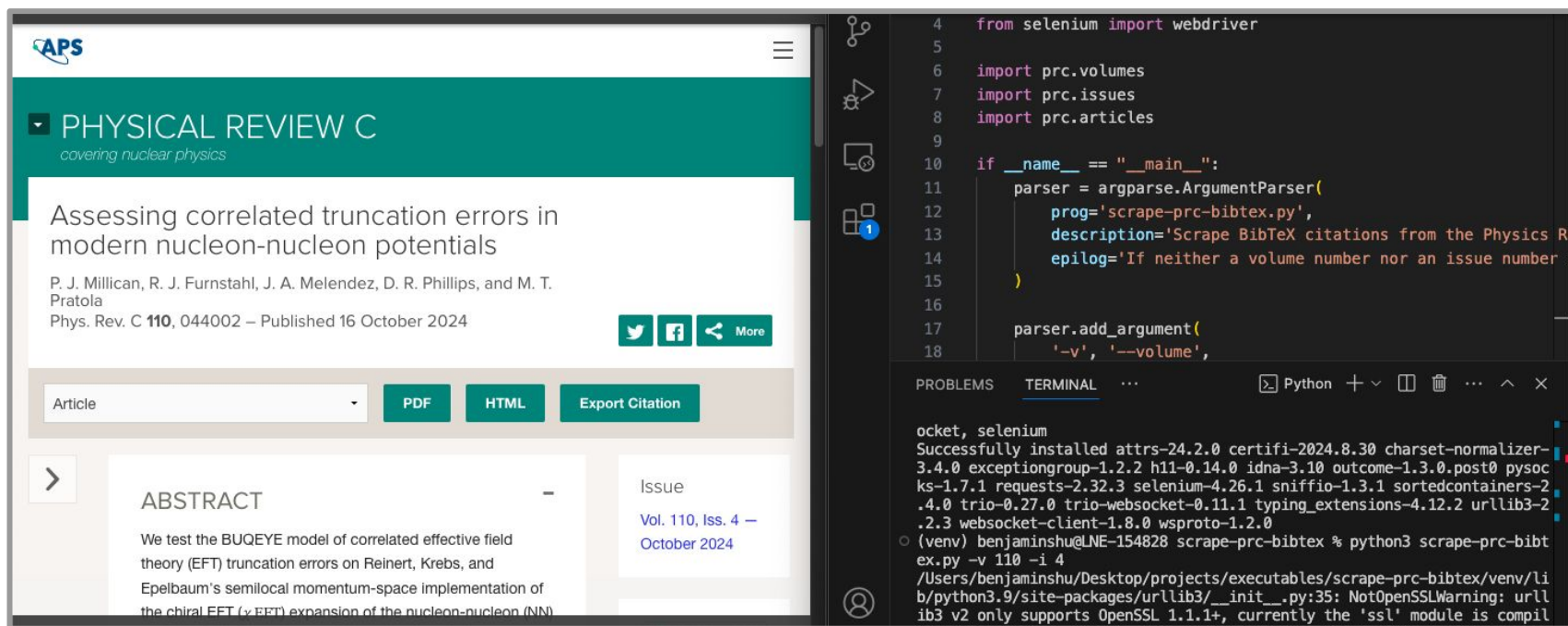
Year:  Entries added:  mm/dd/yyyy

Primary Only

# Step 4: Supporting Software

Automated script for collecting PRC articles

- Visits web page for each article and saves BibTeX citation



The image shows a split-screen view. On the left is a web browser displaying the APS (American Physical Society) website for a Physics Review C article. The article title is "Assessing correlated truncation errors in modern nucleon-nucleon potentials" by P. J. Millican, R. J. Furnstahl, J. A. Melendez, D. R. Phillips, and M. T. Pratola, published in Phys. Rev. C 110, 044002 on October 16, 2024. The page includes options to view the article as PDF, HTML, or to export a citation. On the right is a code editor showing a Python script that uses Selenium to interact with the website and argparse to handle command-line arguments for volume and issue numbers. The script imports selenium.webdriver, prc.volumes, prc.issues, and prc.articles. It defines a main function that sets up an argument parser with options for volume (-v) and issue (--volume). The terminal output shows the script being executed successfully, installing various dependencies like attrs, certifi, charset-normalizer, etc., and running the script with arguments -v 110 -i 4.



# Step 5: Distribution, Collaboration

Goal: Make NSR as accessible as possible

- Changes have already enabled collaboration with INSPIRE
  - “INSPIRE now links to over 100,000 records in the NSR.”



<https://blog.inspirehep.net/2024/09/inspire-and-the-nuclear-science-references-database/>

# Step 4: Distribution, Collaboration

Goal: Make NSR as accessible as possible

- Possible future data model updates
  - ORCID's for authors?

```
{  
  "type": "SINGLE_AUTHOR",  
  "text": "S.F.Mughabghab",  
  "initials": ["S", "F"],  
  "last_name": "Mughabghab",  
  "orcid": "INSERT ORCID HERE"  
}
```

The JSON schema makes it easy to add this field.

Finding and filling in ORCID's will be much harder.

# Step 4: Distribution, Collaboration

Goal: Make NSR as accessible as possible

- Possible future software
  - JSON editor for entries?
  - Offline versions of code tables for nuclides, subjects, reactions, etc.
- Contact us:  
<https://www.nndc.bnl.gov/contact-us/>
- Email: [bshu@bnl.gov](mailto:bshu@bnl.gov)