



NSR Report

B. Pritychenko¹, J. Totans¹, B. Singh², D. Symochko², V.V. Zerkin³, L. Vrapcenjak³, A. Rodionov⁴, G.I. Shulyak⁴

1 National Nuclear Data Center, Brookhaven National Laboratory, Upton, NY 11973-5000, USA

2 Under contract with National Nuclear Data Center, Brookhaven National Laboratory, Upton, NY 11973-5000, USA 3 Nuclear Data Section, International Atomic Energy Agency, Vienna International Centre, P.O. Box 100, A-1400 Vienna, Austria

4 B.P. Konstantinov Petersburg Nuclear Physics Institute, Orlova Roscha, Gatchina 188300, Leningrad District, Russian Federation

Nuclear Science References (NSR)

- Katherine Way (formerly at Met Lab, Chicago) recognized the importance of keywords and convinced several journals to include them.
- Oak Ridge NSR entry consisted of authors, journal and simple nuclear structure keywords; title was often missing.
- References were stored at the ORNL library.
- In 1980 NSR operation was transferred from Oak Ridge to Brookhaven.
- In subsequent years the database was integrated with XUNDL/ENSDF and EXFOR.
- However, evaluators worldwide were still spending plenty of time for references location because not everyone had access to good library resources.



BURBIDGE, BURBIDGE,

- Al57a L. H. Aller, Preprint for Handbuch der Physik (Springer-Verlag, Berlin, 1957).
- Al57b L. H. Aller and J. L. Greenstein (private communication).
- Al57c Aller, Elste, and Jugaku, Astrophys. J. Suppl. 3, 1 (1957).
- Al57d L. H. Aller, Astrophys. J. 125, 84 (1957).
- Al50 R. A. Alpher and R. C. Herman, Revs. Modern Phys. 22, 153 (1950).
- Al53 R. A. Alpher and R. C. Herman, Ann. Rev. Nuclear Sci. 2, 1 (1953).
- Ar53 Arp, Baum, and Sandage, Astron. J. 58, 4 (1953).
- Aw56 M. Awschalom, Phys. Rev. 101, 1041 (1956).
- Ba43 W. Baade, Astrophys. J. 97, 119 (1943).
- Ba45 W. Baade, Astrophys. J. 102, 309 (1945).
- Ba56 Baade, Burbidge, Hoyle, Burbidge, Christy, and Fowler, Publ. Astron. Soc. Pacific 68, 296 (1956).
- Ba57a W. Baade (private communication).
- Ba57 H. W. Babcock, Proceedings of the Stockholm Symposium on Electromagnetic Phenomena in Cosmical Physics (to be published).
- Ba50 C. L. Bailey and W. R. Stratton, Phys. Rev. 77, 194 (1950).



Nuclear Science References (NSR)

A.Bauc N.Amza

First Id

NUCLE/ discuss

NUCLE

Data fro

- Katherine Way (formerly at Met Lab, Chicago) recognized the importance of keywords and convinced several journals to include them. •
- Oak Ridge NSR entry consisted of authors, journal • and simple nuclear structure keywords; title was often missing.
- References were stored at the ORNL library. •
- In 1980 NSR operation was transferred from Oak • Ridge to Brookhaven.
- In subsequent years the database was integrated • with XUNDL/ENSDF and EXFOR.
- However, evaluators worldwide were still spending • plenty of time for references location because not everyone had access to good library resources.



Data from this article have been entered in the EXFOR database. For more information, access X4



- Total number of new compilations: 4,590
- Total number of keyworded compilations: 2,872
- Lack of conferences: ~500 Los Alamos National Laboratory Reports were compiled using the Primo library.
- NSR Dictionaries:
 - 2,565 new authors
 - 14 new journals
 - 59 new decays
 - 164 new reactions
 - 460 new nuclides
- NSR DB updates: 181
- NSR Web: 334,383 retrievals, 97 references/retrieval
- More than 32.5 M references were retrieved, each reference was retrieved 32,556,822/240,414 ~ 135 times
- Work from home was very productive.





- Total number of new compilations: 4,590
- Total number of keyworded compilations: 2,872
- Lack of conferences: ~500 Los Alamos National Laboratory Reports were compiled using the Primo library.
- NSR Dictionaries:
 - 2,565 new authors
 - 14 new journals
 - 59 new decays
 - 164 new reactions
 - 460 new nuclides
- NSR DB updates: 181
- NSR Web: 334,383 retrievals, 97 references/retrieval
- More than 32.5 M references were retrieved, each reference was retrieved 32,556,822/240,414 ~ 135 times
- Work from home was very productive.





- Total number of new compilations: 4,590
- Total number of keyworded compilations: 2,872
- Lack of conferences: ~500 Los Alamos National Laboratory Reports were compiled using the Primo library.
- NSR Dictionaries:
 - 2,565 new authors
 - 14 new journals
 - 59 new decays
 - 164 new reactions
 - 460 new nuclides
- NSR DB updates: 181
- NSR Web: 334,383 retrievals, 97 references/retrieval
- More than 32.5 M references were retrieved, each reference was retrieved 32,556,822/240,414 ~ 135 times
- Work from home was very productive.





- Total number of new compilations: 4,590
- Total number of keyworded compilations: 2,872
- Lack of conferences: ~500 Los Alamos National Laboratory Reports were compiled using the Primo library.
- NSR Dictionaries:
 - 2,565 new authors
 - 14 new journals
 - 59 new decays
 - 164 new reactions
 - 460 new nuclides
- NSR DB updates: 181
- NSR Web: 334,383 retrievals, 97 references/retrieval
- More than 32.5 M references were retrieved, each reference was retrieved 32,556,822/240,414 ~ 135 times
- Work from home was very productive.





NNDC ``Paper" Library

- NNDC collects unique references, we acquired Oak Ridge library, Los Alamos, McMaster U. and many other collections.
- NNDC implemented new scanner, mirofiche reader, ...
- Joann Totans is leading this effort.
- Traditional ``paper" library became a bottleneck due to large volume of requests, and change was needed.





- Around 2005 Viktor Zerkin (Nuclear Data Section, IAEA) suggested to extend EXFOR database with published articles (PDF files).
- He started to store PDFs in the EXFOR relational database.
- Data tables are essentially no different than Excel tables: numbers, text and dates.
- Binary Large Objects or BLOBs.
- Databases manage extremely long tables using Structured Query Language (SQL).





- Around 2005 Viktor Zerkin (Nuclear Data Section, IAEA) suggested to extend EXFOR database with published articles (PDF files).
- He started to store PDFs in the EXFOR relational database.
- Data tables are essentially no different than Excel tables: numbers, text and dates.
- Binary Large Objects or BLOBs.
- Databases manage extremely long tables using Structured Query Language (SQL).

H	1		2								4	Acted .						
H													1					
						11,	N		100					1				
10-	-	- AD		31	HU		b		1		land	and i	1					
				13	-		6.4	100	an 1			0	1	11				
	J.	2		17					A.			-						
							1. EI II.					π			J			
	•	•••	AutoSa	ave 🔵 off) 🖒 🖻	· د ب				🖻 upo	dstat ~							Q (
	Но	ome li	nsert	Draw	Page La	ayout	Formula	as Da	ta Re	/iew	View						🖻 Share	e 🖓 Comme
T-AL	7	~ v								ab,	0/			Eormatt	ing v		\cap	
		۵° آ		Arial		~ 10	▲ A [*]	A'	= = 5	≣ ce · ≣ ि ⊽	% *	E E	Format as	Table v	ing -	¥ *	\mathcal{O}	·
	P	aste 🔏		BIL	J • 🗄	- 0	• • A	~		·	Number	12	Call Styles	~		Cells	Editing	Ideas
								-		~ .		H2	g cen orgies	-				
	R3	7	X	√ fx														
		A	В	С	D	E	F	G	н	1	J	к	L	М	N	0	Р	Q
	1	Date	Entries	Mods	KWs	Entries	McMaster Modifies	KWs	Entries	Bratislava Modifies	KW		Berkeley		Entries	IAEA Modifies	KWs	Comments
	3	05/14/09	61	2 21	39												20	
	4 5	05/21/09	83	3 2	45	82		47	,						5)	32	
	6	06/03/09	12	2 0	9	0		0)							0	0	
	8	06/18/09	4	1 2	2 33	0		0	,)	0	
	9	07/06/09		1 1	1 1	104		67)	0	
	11	07/10/09	112	2 35	5 63	0		0	,							Ď	0	
	12	07/14/09	24	1 1	1 11	0		0)						3	3	30	
	14	07/16/09		1 0) 1	0		0)							5	0	
/	15 16	07/17/09 07/20/09	4	1 4 9 1	1 0 1 6	88		70))	0	
	17	07/30/09	16	3 2	2 13	0		0))	0	
-4	18	08/11/09 08/14/09	36	r 1 6 0	22	0		0)						-)	0	
/	20	08/18/09	70) 3	3 47	0		0))	0	
-	22	08/25/09	13	3 1	, 34 I 7	0		0	,							, ,	0	
	23	08/26/09	14	2 10	1	90		62	2)	0	
	25	09/11/09	10) 1	1 8	0		0) 12		12					ó	0	
A	26	09/15/09	15	5 0	0 7	0		0	0 50		27)	0	
2	28	09/23/09	(0 0	0 0	0		0	82		68)	0	
have	30	09/28/09	() 4	4 0	87		57	· 0		0				5)	0	
-	31	10/02/09	28	3 0	0 17	0		0	42		35				1	4	8	
1-	33	10/14/09	58	3 0	36	0		0	0 0		0					Ś	0	
-	34	10/20/09	5	9 0	2 7	0		0	0 0		0						0	
-	36	10/29/09	() 3	3 0	106		81	0		0				2	5	11	
1	37	11/13/09	34	4 0 1 1	20	0		0) 53) n		29				-)	0	
NEDER	39	11/30/09		5 5	5 2	105		68	51		28					0	0	
03	40 41	12/04/09 12/10/09	2.	i 1 9 0	16	0		0	0 0		0)	0	
	42	12/15/09	1	5 0	3	0		0	0 0		0				21	0	14	
	43	12/18/09	18	2	2 11	0		0	32		16				4	3	29	
BAROUT	44	12/23/09	D.	s 0	21	0												



- Around 2005 Viktor Zerkin (Nuclear Data Section, IAEA) suggested to extend EXFOR database with published articles (PDF files).
- He started to store PDFs in the EXFOR relational database.
- Data tables are essentially no different than Excel tables: numbers, text and dates.
- Binary Large Objects or BLOBs.
- Databases manage extremely long tables using Structured Query Language (SQL).

MySQL Query Br	AutoSave Orr Connection: / exfor	= mulas Data A^ A [×] ≡	e updstat Review View ≡ ≡ ₩ √ %	Conditional Form Econational Form	atting v	Q © v Le ² Share ♀ Comments ↓ ♀ ↓ ↓ - □ ×
e Edit View C	Query Script Tools Window MySQLEn	terprise Help				
back Next Ref	SELECT * FRON 'exfor'.'x4p	ub_pdf`				Execute - Stop
🕈 🕖 Resultset 1						Schemata Bookmarks History
📍 x4pdflD	stdFileName	NsrKeyNo	iAccessFlag	RefYear mypdf		2
1	J,PR,163,1299,1967	1967ME21	1	1967	S	T aufor
29465	R.AAEC AP PR-81,3,1981		1	1981	9.8	• exitin
2	R,INDC(HUN)-14,1978		0	1978 BLOB	۹.FI	allAuthor1
3	R.INR-1773 I PL A.16,1978		1	1978 💷	9.11	all edword
4	R,LA-2177,1959		1	1959 BLOB	9.8	▶ allProduct
29395	C,78ALMAATA,,200,1978		1	1978 8108	٩.8	▶ iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii
29475	J,PR_C,88,034606,2013	2013M026	1	2013 BLOB	9.8	▶ 🛄 allRef
14	J,PZ,19,257,1918	1918ME01	1	1918	9.5	▶ 🛄 allReference
28045	N,NSR-1969VE13	1969VE13	1	1969 BLOB	9.8	▶ 🛄 allSF1
28046	N,NSR-1970AZ01	1970AZ01	1	1970 BLOB	9. FI	▶ 🛄 allSF2
28047	N,NSR-1970BE79	1970BE79	1	1970 8108	9.8	▶ 🛄 allSF3
28048	N,NSR-1970BE86	1970BE86	1	1970 8108	9. Fi	▶ <u>allSF4</u>
27905	J,BAS,33,1889,1970		1	1970	9.E	I allSF5
27906	J,NAP,9,662,1970		1	1970 BLOB	S.E.	
27823	J,0SA,108,185,1971		1	1971 BLOB	ч П	F IIII ARAFZ
27824	J,PRM,3,186,1974	1974R044	1	1974 BLOB	S.E.	Syntax Functions Params Trx
27811	C,73R0CH,2,502,1973	1974H0YY	1	1973 BLOB	<u>م</u>	Contraction Data Definition Statements
27812	C,75GATLIN,4,158,1975	1976LAZE	1	1975 BLOB	۹.E	🛅 Data Manipulation Statements
27784	J,JNE,24,43,1970	1970SI27	1	1970 BLOB	۹.A	MySQL Utility Statements
27785	R,KURRI-AR-7,72,1974		2	1974 8108	ч.П	MySQL Transactional and Locking
27786	J,PKL,2,79,1975		1	1975 BLOB	۹.FI	Database Administration Statements Replication Statements
27730	J,SNP,13,272,1971		1	1971 BLOB	ч н	SQL Syntax for Prepared Statements
27625	R,0RNL-4659,38,1971	1970BAYF	1	1971	<u>я</u> н	
27626	J,RRL,15,29,1973	1973KU17	1	1973	ч н	
27627		1 A Analy of	1	1976 8008	Q ⊆uri	
rows retched so far.	- Edit	Appry Cha	nges 🕋 Discard Chang	ges P4 Hrst P1 Last	✓ Search	



- Around 2005 Viktor Zerkin (Nuclear Data Section, IAEA) suggested to extend EXFOR database with published articles (PDF files).
- He started to store PDFs in the EXFOR relational database.
- Data tables are essentially no different than Excel tables: numbers, text and dates.
- Binary Large Objects or BLOBs.
- Databases manage extremely long tables using Structured Query Language (SQL).





- Portable Document File (PDF) database operation.
- Regular updates of EXFOR and NSR databases extended to PDF files.
- Private collections for existing entries.
- PDF library coverage reflects publishing trends.
- Authorized Web access only due to major publishers restrictions.
- Continuing PDF scanning effort at NNDC.
- Future contributions from LBNL and other labs are more than welcome.



- Portable Document File (PDF) database operation.
- Regular updates of EXFOR and NSR databases extended to PDF files.
- Private collections for existing entries.
- PDF library coverage reflects publishing trends.
- Authorized Web access only due to major publishers restrictions.
- Continuing PDF scanning effort at NNDC.
- Future contributions from LBNL and other labs are more than welcome.





- Portable Document File (PDF) database operation.
- Regular updates of EXFOR and NSR databases extended to PDF files.
- Private collections for existing entries.
- PDF library coverage reflects publishing trends.
- Authorized Web access only due to major publishers restrictions.
- Continuing PDF scanning effort at NNDC.
- Future contributions from LBNL and other labs are more than welcome.





- Portable Document File (PDF) database operation.
- Regular updates of EXFOR and NSR databases extended to PDF files.
- Private collections for existing entries.
- PDF library coverage reflects publishing trends.
- Authorized Web access only due to major publishers restrictions.
- Continuing PDF scanning effort at NNDC.
- Future contributions from LBNL and other labs are more than welcome.





- Portable Document File (PDF) database operation.
- Regular updates of EXFOR and NSR databases extended to PDF files.
- Private collections for existing entries.
- PDF library coverage reflects publishing trends.
- Authorized Web access only due to major publishers restrictions.
- Continuing PDF scanning effort at NNDC.
- Future contributions from LBNL and other labs are more than welcome.





Towards a Natural Language Processing Suite Optimized for Nuclear Physics

The primary objective of this project is to make NSR more accessible and useful to researchers:

- Automation of the process of adding new articles to the NSR database by using NL;
- Development of ML algorithms to automate the identification and application of keywords;
- Expansion of the scope of the articles included in the database to related applied fields;
- Adaption of a nonproliferation-oriented keywords lexicon;
- An updated NSR infrastructure using a modern database and programming architecture;
- An updated and modernized user interface that BM/BK/BMP port both desktop and mobile access.

Cover	^r Page
The project title:	Towards a Natural Language Processing Suit
Applicant/Institution:	Brookhaven National Laboratory
Street Address/City/State/Zin:	Building 490C Upton NY 11973
Postal Address:	P.O. Box 5000 Unton NY 11973
Lead PI name, telephone number, email:	Maia Gemmill Nonproliferation and National Security Department E-mail: mgemmill@bnl.gov Phone: (631) 599-6231
Administrative Point of Contact name, telephone number, email:	Bonny Jean Forte Nonproliferation and National Security Department Email: bforte@bnl.gov Phone: (631) 344-2198
FOA Number:	DE-FOA-0002440
DOE/Office of Science Program Office:	Nuclear Physics
DOE/Office of Science Program Office Technical Contact:	Mr. Keith Jankowski Program Manager, Nuclear Data keith.jankowski@science.doe.gov
DOE Award Number (if Renewal Application):	N/A
PAMS Letter of Intent Tracking Number:	33687
Research area or areas as identified in Section I of	DOE SC NP

Lead Institution: Brookhaven National Laboratory, Maia Gemmill (PI, point of contact), Boris Pritychenko (PI), Benjamin Shu, Alejandro Sonzogni Collaborating Institution: Lawrence Berkeley National Laboratory, Lee Bernstein (PI), Bethany Goldblum, Walid Younes Collaborating Institution: Stony Brook University, Andrew Schwartz

Takeaways

- NNDC NSR compilation efforts are complex and wellorganized: B. Pritychenko, J. Totans (BNL), B. Singh, D. Symochko (Contractors), V. Zerkin (IAEA).
- EXFOR-NSR Portable Document File (PDF) database includes 220 K+ files.
- PDF database effort of V. Zerkin (IAEA), B. Pritychenko, J. Totans (BNL), L. Vrapcenjak (IAEA), A. Rodionov, G.I. Shulyak (PNPI, Gatchina).
- NSR Modernization Proposal.





Takeaways

- NNDC NSR compilation efforts are complex and wellorganized: B. Pritychenko, J. Totans (BNL), B. Singh, D. Symochko (Contractors), V. Zerkin (IAEA).
- EXFOR-NSR Portable Document File (PDF) database includes 220 K+ files.
- PDF database effort of V. Zerkin (IAEA), B. Pritychenko, J. Totans (BNL), L. Vrapcenjak (IAEA), A. Rodionov, G.I. Shulyak (PNPI, Gatchina).
- NSR Modernization Proposal.



Table 1: PDF coverage for EXFOR and NSR references as of October 14, 2021. Reciprocal PDF contributions are shown as # of complementary files.

Database	# of References	# of PDF Files	# of Complementary Files
NSR	236,583	187,617	1,375
EXFOR	$34,\!609$	26,343	1,899
CINDA	$39,\!817$	$14,\!154$	
IBANDL	795	642	

