EXFOR search by author using aliases

V.Zerkin, IAEA-NDS, 22 April, 2014

Although EXFOR Web interface offers search by authors using wildcards if necessary, it does not have such a comprehensive authors-alias system as NSR does. (NSR database has a system of aliases for persons, i.e. initials + last names.) This working paper describes an extension of EXFOR database retrieval system implementing simplified alias tool dealing only with last names of the authors. The system is implemented in two parts: formation of an internal Dictionary-349 and Web search using this dictionary.

Dictionary-349 is formed from in three steps: import from NSR (1256 groups, 2510 names) and from EXFOR using special algorithms (112 groups, 248 names), and merging them to common dictionary-file (358 groups, 820 names). Algorithm for automatic identification of aliases from EXFOR names translated from Cyrillic is implemented for the following patterns:

*ski, *skii, *skij, *skiy, *sky,*skyy

*skaya, *skaia, *skaja

*ian, *jan, *yan

Examples of aliases from the Dictionary-349:

#	Source	Name	Alias	Alias	Alias	Alias
1	NSR	Abdel-Kawi	Abdel-Kawy			
2	NSR	Abouzeid	Abuzeid			
3	NSR	Afanasev	Afanasjev			
12	NSR	Aleksandrova	Alexandrova			
16	EXFOR	Amarian	Amaryan			
52	EXFOR	Chuvilskaia	Chuvilskaja	Chuvilskaya		
67	NSR	De Oliveira	De Oliveira Santos	Oliveira		
217	NSR+EXFOR	Muradian	Muradjan	Muradyan		
258	EXFOR	Preobrazhenskaja	Preobrazhenskaya			
322	NSR+EXFOR	Tokarevski	Tokarevskii	Tokarevskij	Tokarevskiy	Tokarevsky
327	NSR	Tsoupko-Sitnikov	Tsupko-Sitnikov	Zupko-Sitnikov		
357	NSR+EXFOR	Zheltonozhskii	Zheltonozhskij	Zheltonozhskiy	Zheltonozhsky	

Web access

The system works automatically when user searches data by author(s). Example [16] on: http://www-nds.iaea.org/exfor/ can be used for testing where search by Author=[Vysotsky] generates result:

Conclusion

The system is functioning in semi-automatic regime using information from external database NSR. Depending on interest from other NRDC members, it can be improved and/or released within NRDC community.