

Increased Data Mining Power for the ICDD PDF-4 Products: Customizable Search Preferences to Select the Properties Displayed in a Database Query

Jerome Bridge¹, Justin Blanton and John Faber
International Centre for Diffraction Data (ICDD)

The PDF-4 Relational Database (RDB) products have integrated search and viewer functions that allow the user to query the database. Over 31 searchable fields are provided and any of these can be combined using Boolean logic to create a cumulative search result. Moreover in most cases, the individual searches (forms) contain a fixed display (PDF#, Quality Mark, Chemical Formula, Chemical Name, d-spaces for the 3 strongest lines, and the crystal system). The searchable parameter is often shown to help confirm the nature of the search query.

To enhance the display information content in PDF-4 RDBs, we have developed customizable user preferences² that operate both individually and globally on search forms. On each individual search form, the user can customize the list of display properties. In addition, the ordering of the display properties can be selected. Global preferences are a set of display properties that can be applied to all search forms. An example is shown in Figure 1.

PDF #	QM	Compound Name	Empirical Formula	D1	D2	D3	RCVol	R Cell A	R Cell B	R Cell C	SYS	I/IC
01-073-2475	C	Ammonium Copper Chloride	Cl4 Cu H12 N2 O2	5.519	2.701	2.759	465.790	7.640	7.640	7.980	T	1.500
01-073-2476	C	Strontium Chloride Hydrate	Cl2 H4 O2 Sr	2.926	1.690	1.885	59.860	3.330	3.989	5.253	M	.520
01-073-2477	C	Cesium Lead Chloride	Cl6 Cs4 Pb	3.805	4.244	6.590	834.430	9.416	9.416	9.416	R	5.800
01-073-2478	C	Cesium Lead Bromide	Br6 Cs4 Pb	2.950	6.865	7.000	942.530	9.807	9.807	9.806	R	4.380
01-073-2479	C	Strontium Chloride Hydrate	Cl2 H12 O6 Sr	3.970	6.876	3.527	224.290	4.108	7.940	7.940	H	2.870
01-073-2480	C	Potassium Tin	K8 Sn46	3.215	2.063	5.380	1740.990	12.030	12.030	12.030	C	6.200
01-073-2481	C	Cesium Molybdenum Compound	Cs2 Mo O S3	3.621	3.845	3.551	864.920	7.242	9.771	12.223	O	2.170
01-073-2482	C	Lead Selenide Bromide	Br6 Pb4 Se	2.801	4.138	2.470	334.450	4.361	9.520	9.520	O	6.510
01-073-2483	C	Lead Sulfide Iodide	I6 Pb5 S2	3.225	3.113	3.411	458.330	4.437	7.504	14.552	M	7.310
01-073-2484	C	Ammine Zinc Nitrogen	H12 N6 O6 Os2 Zn	3.741	6.108	3.190	296.070	7.481	7.481	7.481	C	13.300
01-073-2485	C	Ammine Cadmium Osmium	Cd H12 N6 O6 Os2	3.783	6.178	2.184	306.260	7.566	7.566	7.566	C	15.230
01-073-2486	C	Strontium Lithium Antimony	Li4 Sb4 Sr3	2.845	2.922	3.584	289.870	4.880	7.500	9.096	O	4.660
01-073-2487	I	Copper Vanadium Oxide	Cu2 O7 V2	3.258	3.077	3.215	279.620	5.291	5.291	10.829	O	3.720

Figure 1. The PDF Number search display form using user-selectable preferences.

In this paper, we will describe the flexibility and ease of setting both individual and global search preferences. A laptop demonstration will be used to provide hands-on experiences. Moreover, we will show that with these preferences, we can display almost all of the paper index book product formats.

¹ Student Intern, Computer Science Department, West Chester University

² In contrast, consider the case for our flat file PDF-2/PCPDFWIN database and viewer where the search results are completely fixed and only the PDF card data itself is shown.