Journal Articles Related to the Transplutonium Elements

JERRY B. AYERS
Tennessee Technological University, Cookeville, Tenn. 38501

Received October 5, 1971

Manual searches were made of *Chemical Abstracts* and *Nuclear Science Abstracts* for the period 1962 through 1968 for all journal articles related to the transplutonium elements. A computer-based search of *Chemical Titles* was made for the same period. Comparisons were made of the overlap between the three reference sources, and the effectiveness of the CT computer-based search was determined.

Computer-based literature searches are becoming increasingly popular for current awareness searches and are being used for retrospective searches as files are accumulated. *Chemical Titles* (CT), on magnetic tape since 1962, cites articles from approxmately 650 journals (CT covers only journal literature). In 1966, CT included 58% and 48%, respectively, of all journal articles and total references in *Chemical Abstracts* (CA).[3] The CT magnetic tape version contains edited authors' titles, authors' names, and primary citations.

The purpose of this study was to determine the effectiveness of a CT retrospective computer-based search. The effectiveness of a computer-based CT title search is dependent upon how adequately authors' titles represent the contents of their papers, how well the search query is constructed and coded, and the techniques used in preparing the computer programs. It has been estimated that search questions constructed for maximum efficiency for title searching will retrieve 33 to 39% of the actual journal articles of interest cited in CT.[2]

The method used to investigate the effectiveness of a CT retrospective search was a case study using the literature of the transplutonium elements. Any observations from this tudy are biased by the vocabulary of the transplutonium element literature, and the results of this case study may not necessarily represent retrieval effectiveness of other search topics in CT. The literature of the transplutonium elements falls within the defined subject categories of both CA and *Nuclear Science Abstracts* (NSA). Therefore, comparisons of search results of CT were made with both CA and NSA.

EXPERIMENTAL

A computer-based search of CT was carried out for the period 1962 through 1968 in order to locate all journal articles related to the transplutonium elements that were published during the period 1962 through 1967. Most articles are cited in CT within three months of their original publication. A specific search question was written for each element, as well as one broad search question to cover general synonyms for the transplutonium elements. Figure 1 shows a sample search question for mendelevium, and Figure 2 shows the search question for the general transplutonium terms. All questions were one parameter searches—that is, only one match between a reference term and any term listed in the search question was necessary to retrieve that reference. The search system used allows for both prefix and suffix truncation which is indicated by the asterisk. Since some words are fragmented in CT, it was necessary to fragment terms in the search question. Terms for the general search question were selected to be more applicable to the transplutonium elements than to elements or metals in general. With some terms, such as ACTINIDE*, irrelevant references would be retrieved; however, it was felt these terms were needed to give adequate retrieval. Terms broader in concept, such as HEAVY ELEMENT* or RADIOACTIVE ELEMENT*, were not used since these terms would cause more irrelevant than relevant references to be retrieved.

A manual search of the volume subject indexes of CA and NSA was made for the period 1962–1968 for journal references related to the transplutonium elements that were published during 1962–1967. Entries under the specific names and variations of names for each element, as well as entries under general terms such as transuranium, transplutonium, and actinide were retrieved. Figures 3 and 4, show, respectively, the entry headings which were searched in CA, Volume 65 and NSA, Volume 20.

In addition, the entries to all headings which were variations on actinide or transuranium—i.e.; actinide carbides, actinide series, transuranium compounds, and transuranium elements—were searched. The NSA entries which were modifications of the words actinide, transplutonic, and transuranic were also searched. Abstracts of journal articles were examined for those references in which the index did not specify the exact element(s) being referenced.

Comparisons were made of the journal references re-

MENDELEV
ELEMENT-101
ELEMENT 101

Figure 1. Search question for mendelevium

TRANSPLUTON
TRANS PLUTON
TRANS-PLUTON
TRANSURAN
TRANS URAN
TRANS-URAN
ACTINID
ACTINIUM SERIES
ACTINYL ELEMENT*
ACTINON*

Figure 2. General search question for transplutonium elements

Americiates (III)
Americiates (VI)
Americium
Americium, analysis
Americium, (nitrilotriacetate),
________, tris(dihydrogen phosphato)tris(phosphoric acid)-(tributyl phosphate)-
________, tris(hydrogen phenylphosphonato)tris(phenylphosphonic acid)(tributyl phosphate)-
Americium barium protactinium oxide
Americium butyl phosphate
Americium compounds
Americium hydride
Americium lithium fluoride
Americium proactinium oxide
Americium sodium fluoride
Americium xenonate (VIII)

Figure 3. Subject headings related to americium in *Chemical Abstracts,* Volume 65

AMERICIUM
AMERICIUM ALLOYS
AMERICIUM COMPLEXES
AMERICIUM COMPOUNDS
AMERICIUM FLUORIDES
AMERICIUM IONS
AMERICIUM ISOTOPES
AMERICIUM ISOTOPES Am-241
AMERICIUM ISOTOPES Am-242
AMERICIUM ISOTOPES Am-243
AMERICIUM ISOTOPES Am-244
AMERICIUM OXIDES
AMERICIUM SULFATES

Figure 4. Headings related to americium in *Nuclear Science Abstracts,* Volume 20

trieved by manual searches of CA and NSA with those references retrieved from the computer-based search of CT. Only those journal articles published in the years 1962–1967 were compared. Some 1961 and 1968 references were retrieved but discarded since all articles for those two years were not retrieved. Journal references which were retrieved in the manual search of CA but not in the computer-based search of CT were located in the printed copy of CT to determine why they were not located. References were not retrieved in the CT search because: (1) the journal was not covered by CT, (2) the journal was covered by CT but the particular article was not included, or (3) the article was in the printed CT but the title words did not match the search questions. Most references were not retrieved by the computer-based search owing to reasons 1 and 3 above.

RESULTS AND DISCUSSION

Table I presents the results of the manual searches of CA and NSA for references related to the transplutonium elements for the years 1962 through 1967. Included in this table are summaries of the number of journal references retrieved for each element and for the total group of elements. CA and NSA contained, respectively, 1030 and 921 references. Of this number 536 were common to both sources resulting in a net total of 1425 unique references. This figure represents a 57% overlap of CA with NSA, and a 51% overlap of NSA with CA. Another study[1] showed that the overlap between all references (books, reports, patents, theses, journal articles, etc.) to the transplutonium elements between CA and NSA was 37%. This discrepancy in overlap can be attributed to the difference in abstracting between CA and NSA. NSA abstracts a substantially larger number of reports, conference proceedings, etc., than CA.[1]

In general, the longer an element has been known, the greater the number of references retrieved for the time span studied and also the higher the percentage overlap between the two abstract sources. For example, a total of 522 distinct references were retrieved for the element americium (first produced in 1944)[6] with a 62% overlap of CA with NSA. In comparison, 22 distinct references were retrieved for the element lawrencium (first produced in 1961),[4] with a 9% overlap of CA with NSA. In the future, it may be necessary to search only one abstract source for information pertaining to some of the elements, if the trend in overlap continues. At the present time; however, it is almost mandatory to search both sources for complete information about any of the transplutonium elements.

The results of a computer-based search of CT and a manual check of CT for references related to the transplutonium elements are presented in Table II. A total of 580 references related to the transplutonium elements, were retrieved using a profile designed for high precision. Of this number, 59% of the references were retrieved by the element name, 33% by use of general terms such as actinide or transplutonium, and 8% by use of another element name (an article may be related to curium but retrieved by means of the search term americium). This latter problem occurs because the title of an article does not fully describe its contents.

In an effort to obtain the number of relevant citations in CT and for calculating recall, all references retrieved by manual search of CA were cross-checked with the printed

Table I. Comparison of Results of Manual Search of *Chemical Abstracts* and *Nuclear Science Abstracts* for Journal References Relating to the Transplutonium Elements (1962–1967)

	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	Total
Total references retrieved by manual search of CA	384	226	61	182	46	65	29	25	12	1,030
Total references retrieved by manual search of NSA	358	215	58	138	41	53	23	24	11	921
Total references common to both CA and NSA	220	131	28	83	18	27	10	8	1	526
Unique references in CA and NSA	522	310	91	237	69	91	42	41	22	1,425
Overlap of CA with NSA, %	62	61	48	60	44	51	44	33	9	57
Overlap of NSA with CA, %	57	58	46	46	39	42	35	32	8	51

Table II. Results of Computer-based and Manual Check of *Chemical Titles* for Journal References Relating to the Transplutonium Elements (1962–1967)

	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	Total
Total references retrieved by computer search of CT	235	133	36	96	23	31	10	13	3	580
Retrieved by element name, %	70	50	41	57	35	58	60	69	67	59
Retrieved by general name, %	25	38	47	34	52	35	40	23	33	33
Retrieved by other element name, %	5	12	12	9	13	7	0	8	0	8
Number of additional references located by manual check of CT	90	42	11	54	15	20	8	8	8	256
Total number of references in CT	325	175	47	150	38	51	18	21	11	836
Recall from information in title, %	72	76	77	64	61	61	56	62	27	69

Table III. Comparison of the Overlap of Computer-based Search of *Chemical Titles* and Manual Searches of *Chemical Abstracts* and *Nuclear Science Abstracts* for Journal References Relating to the Transplutonium Elements (1962–1967)

	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	Total
Total reference common to CT and CA	203	122	27	81	18	26	9	11	3	500
Overlap of CT computer search with manual search of CA,%[a]	53	54	44	45	39	40	31	44	25	49
CA journal references in CT,%	76	73	62	74	72	62	59	76	92	73
Total references common to CT and NSA	182	105	27	67	17	20	6	8	1	433
Overlap of CT computer search with manual search of NSA,%[a]	51	49	47	49	41	38	26	31	9	47
Total references in CT, CA and NSA	156	92	21	56	13	16	5	6	0	365
Overlap of CT with CA and NSA,%[a]	67	69	58	58	58	52	50	46	0	63

[a]Based on smallest number possible.

copy of CT, and an additional 256 references were located in CT. Thus the total number of references retrieved from CT was 836. On this basis, the recall of information by searching journal titles related to the transplutonium elements was 69%. Based on a study of the percentage of key words contained in CT titles, Bottle[2] indicated that the maximum recall for a search such as conducted in this study would be about 33–39%. The percentage recall of information related to the transplutonium elements could have been raised by broadening the search question used in this study. For example, many articles were not retrieved because the following terms were omitted from the search profile: heavy elements, heavy ions, cross-section, half-life, even-even nuclei, even-odd nuclei, radioisotopes, names of the actinide elements prior to the transplutonium elements, etc. However, it is obvious that if the search profile contained these terms, the resulting search would yield an extremely high number of articles that are not related to the subject. This increased output would necessitate a lengthy manual check of the literature to rule out those articles not related to the search question. In addition, a computer-based search of CT is based only on authors' titles, and titles frequently do not reflect accurately the contents of the papers.

In another study[5] using CT as a current awareness service, the precision (relevant references divided by the total references retrieved) was found to be only 11%. However, the search question retrieved 89% of all possible relevant references. This study provides an example of using a very broad search question which has high recall but low precision. In the present study, the profiles provided a balance between precision and recall.

The overlaps of CT, CA, and NSA are shown in Table III. The total number of references common to both CT and CA is 500 with a 49% overlap (total references common to CT and CA divided by total number of references retrieved by manual search of CA) between the references retrieved by computer search of CT and manual search of CA. The percentage of total CA journal references in CT for the transplutonium elements is 73%. According to Chemical Abstracts Service,[3] the material covered in CT represents about 58% of the journals abstracted in CA. This high percentage of transplutonium references indicated that CT covered the articles related to these elements better than the average subject. This fact is probably due to the current high research interest in the transplutonium elements.

Eighty journal references retrieved in the CT computer-based search under specific element names were not retrieved in the manual search of the CA volume indexes. Six references related to berkelium were selected at random and rechecked in the CA volume indexes, for a period of three years past the publication date of the article, against all possible entry points (berkelium, berkelium compounds, inverted entries, actinides, transplutonium, etc.). Only one of the six references was located in the check indicating a reference missed in the original search. Therefore, five references, each of which had berkelium in the title, were not indexed under berkelium in the CA subject index.

The total number of articles common to both CT and NSA was 433 with an overlap between the computer search of CT and the manual search of NSA of 47% (total number of references common to CT and NSA divided by total number of references retrieved by manual search of NSA). The total number of references common to CT, CA, and NSA was 365 or 63% of those retrieved from the three sources (total references in CT, CA, and NSA divided by total references retrieved by computer search of CT).

CONCLUSIONS AND SUMMARY

To obtain a complete search of the journal literature of the transplutonium elements for the period 1962 through 1968, it is necessary to search both CA and NSA. This study has shown that the overlap of journal references of the transplutonium elements found in the two reference sources is only about 50%.

A computer-based search of CT, using a profile designed for high precision, recalled approximately 69% of the material cited in the source related to the transplutonium elements. Approximately 59% of this material was retrieved by using the name of the element, while 33% was retrieved using more general terms. The remaining references were retrieved by using another transplutonium element name. About 73% of the journal references in CA related to the transplutonium elements were also cited in CT. However, only about 50% of the CA journal references were retrieved by a computer-based search of CT. In comparison, 47% of the NSA journal references were retrieved by the computer-based search of CT. Over-all, the search questions used in this study retrieved about 40% of the total journal references related to the transplutonium elements found in CA and NSA searches.

ACKNOWLEDGMENT

Thanks are extended to the staff of the Computer Center of the University of Georgia, Athens, Georgia for their assistance in conducting this study.

LITERATURE CITED

(1) Ayers, J. B., "The Development of the Literature of the Transplutonium Elements," *J. Tenn. Acad. Sci.* **46**, 136 (1971).
(2) Bottle, R. K., "Title Indexes as Alerting Services in the Chemical and Life Sciences," *J. Amer. Soc. Inform. Sci.* **21**, 16 (1970).
(3) Chemical Abstracts Service, "CAS Today," Chemical Abstracts Service, Columbus, Ohio, 1967.
(4) Ghiorso, A., Sikkeland, T., Larsh, A. E., and Latimer, R. M., "New Element, Lawrencium, Atomic Number 103," *Phys. Rev. Letters* **6**, 473 (1961).
(5) Lynch, J. T., and Smith, G. D. W., "Scientific Information by Computer," *Nature* **230**, 153 (1971).
(6) Seaborg, G. F., "The Heavy Elements," *Chem. Eng. News* **24**, 1193 (1946).

Searching the *Nuclear Science Abstracts* Data Base by Use of the Berkeley Mass Storage System*

GLORIA L. SMITH and J. JOANNE HERR**
Lawrence Radiation Laboratory, University of California,
Berkeley, Calif. 94720

Received November 17, 1971

The Berkeley Mass Storage System (MSS) is being used for information retrieval. The MSS has an on-line capacity equivalent to more than 100 IBM-2321 data cells. Advantages of the MSS for information retrieval other than its size are: high serial-read rate, archival data storage, and random-access capability. By use of this device, the search cost in an SDI system based on the *Nuclear Science Abstracts* data base was reduced by 20%. A retrospective search system based on NSA subject categories combines random-access and serial-read search techniques to reduce costs markedly.

Since 1967, the Information Research Group (IRG) at the Lawrence Radiation Laboratory in Berkeley has been operating an SDI system based on the *Nuclear Science Abstracts* tapes. Although the system is operational and has 120 users, it is not static. The IRG constantly endeavors to reduce the costs and to improve the quality of the service. Many of the cost-reducing developments have been associated with the availability of random-access devices. The Berkeley Chipstore is one such device that has dramatically influenced the operation of the IRG's system.

The Lawrence Radiation Laboratory (LRL) at Berkeley and Livermore, California has as a part of its computing facilities two of the largest mass storage devices available to date. These are connected to the largest computers made by Control Data Corporation, the CDC-7600 and CDC-

*This version was presented before the Division of Chemical Literature, 161st Meeting, ACS, Los Angeles, March 28, 1971. Preliminary version was presented at the Association of Scientific Information Dissemination Centers (ASIDIC), Semiannual Meeting, Washington, D. C., Feb. 23-25, 1971.

**To whom correspondence should be addressed.