

## Breaking the Subject Index Barrier— A Citation Index for Chemical Patents\*

BY EUGENE GARFIELD †

### I

It is not my intention, as the title of my talk might imply, to attack the excellence of the various indexing services. We all know how indispensable they are. When I speak of the subject index barrier, I have in mind the intellectual limits of a perfectly conceived and executed subject index, wherein the subject relationships of the knowledge recorded is dependent upon the interpretation of authors and/or indexers and must of necessity fail to anticipate all the polyvalent viewpoints of users. Every thought, idea or discovery can take on new meaning depending upon the user's frame of reference. When, as is often the case, the frame of reference of author or indexer is the same as that of the user then conventional subject indexes are quite useful.

The indexer, by judicious arrangements of bibliographical terms, can bring out aspects of a work that the author would overlook. He may also place related articles in closer juxtaposition by editing the index after all entries have been colligated. Keeping the user in mind, the conscientious indexer will translate the terminology and phraseology of the author into a standardized and more usable form. However, the indexer is, of necessity, primarily guided by the subject content which authors provide.

The indexer also faced with a practical economic barrier cannot index with the almost infinite depth to be found in the Citation Index. The Citation Index breaks this "barrier" by presenting subject matter in Bibliographical arrays which are neither alphabetical nor classified but associative.

---

\* Presented at the Minneapolis meeting of the American Chemical Society September 16, 1955.

† Director, The Institute for Scientific Information, 1122 Spring Garden Street, Philadelphia 23, Penna.

Reprinted from the *Journal of the Patent Office Society*, Volume 39, Number 8, August 1957, pp. 583 - 595.

## II

At this point it would be well to describe a Citation Index. Even those familiar with Shepard's Citations<sup>1</sup> may not be able to immediately visualize a similar system for use in patent or literature searching. A Citation Index for patents might indicate under the serial number for each patent those subsequent patents which had made reference to it. An example might look as follows:

Patent No.	Date	Chemical Abstracts Reference	Inventor	Pat. Off. Class
1,970,578	Aug. 21, '34	28, 6332(7)	C. Schoeller	260-458
2,437,261	Mar. 9, '48	42, 4616e	L. L. Lerner	260-397.2
2,450,079	Sep. 28, '48	43, 886d	K. R. Brown	260-234
2,451,149	Oct. 12, '48	43, 352d	E. Boehm	167-30
2,500,349	Mar. 7, '50	44, 5089i	M. DeGroote	252-340
2,518,230	Aug. 8, '50	44, 11, 040i	L. Freedman	167-81
2,541,285	Feb. 13, '51	45, 4892d	A. L. Rawlins	167-81

Figure 1

The citing patents could be classified chronologically, by inventor or by principal subject matter. This sample listing is by no means complete. It is pertinent to cite at the end of a patent, references to the published literature, as well as issued patents. Indeed, it is often just as valuable to *begin* a search with a literature reference rather than a previously *granted* patent. However, I have limited the scope of this paper to issued patents and patent citations.

## III

In conducting research the scientist must be interested in information that is either in anticipation of his own proposed methods or ideas or information which confirms them. Anyone who has done some research will remember that small batch of reprints you kept folded in your lab coat—you neater fellows used looseleaf notebooks, or verticle file folders. These reprints were the few "gems" that you had found when you were still in the preliminary stages of your experimental work and

<sup>1</sup> Adair, W. C. Citation Indexes for scientific literature. *Amer. Documentation* 6, 31-2 (1955).

thinking. Quite often these reprints were so frequently used they did not physically survive the research program. They constituted your bible for a while, and unless you were doing something completely unrelated to any thing previously reported in the literature—a rare phenomenon—it contained a good deal of information relating directly to your immediate interests. During this time any work bearing on the subject of this reprint or having the slightest relevance was extremely important, as a single new idea could mean a whole new line of experimentation. In a certain sense this basic reprint or set of reprints was your base of operations—and bibliographically speaking you might wish that the universe of the index were built around it—a sort of bibliographic solar system.

I have labored this point as a means of pointing out the principal feature of the Citation Index. The access points in conventional indexes are subject headings, together with their sub-headings, and modifications. In the Patent Office the access point is the sub-class, the sub-class name being comparable to the subject heading in an index. In the Citation Index the access point is the individual patent or article, or more accurately the idea or ideas which it presents, implies, or which one can infer. Through the Citation Index one learns of the bibliographical *descendants* of any individual patent or article, of those contributions which for a variety of reasons have cited the work in question. By reference to the patent Citation Index subsequently issued patents (and articles) that had referred to the patent in question could be instantly located.

Needless to say such a bibliographical tool, if current, would permit the scientist or inventor with very specialized interests to keep posted on confirmations, emendations, and refutations of his own published patents or articles as well as those of his colleagues. These bibliographical descendants may be entire articles or they may be portions of articles which in their general context are unrelated to the principal subject matter of the cited patent. This is particularly true when a specialized

technique or procedure is being used in two or more patents that are seemingly unrelated, where the procedures are thereby unknown to the chemist who is possibly interested in the application of the technique and, particularly some small modification in it, which may not attract the attention of the indexer. The old technique may also come up in a new patent claiming the technique for a new compound. A further ramification of this particular point is the help the Citation Index would offer the inventor or his firm in keeping abreast of issued patents which on further investigation are not valid. Examiners do not always have enough time to completely evaluate the very patents they have cited. The conscientious patent man interested in protecting his client and having an intimate knowledge of the patents already owned by his firm can observe more quickly the prior disclosure or proper scope of prior patents.

#### IV

There are two types of references in patents. There are those that are similar to references found in the literature that are provided by the inventor in his application and disclosure. These will not be the concern of my subsequent remarks. However, the references to prior patents made by the patent examiner in his responses to the inventor's claims is a more interesting group of citations. These may or may not result in the disallowance of the inventor's claims. In any event they contain pertinent subject matter found in the examiner's search. These references are to be found in all patents published since February 4, 1947 based on an order issued December 19, 1946 that they be printed at the end of the claims section of patent specifications. Prior to that time references of record are only to be found in the patent "wrapper." These citations are used a great deal in searches and at times an old wrapper will be examined simply to locate these references of record.

Unlike the indexer, the examiner has the legal point of view as well as the subject approach and therefore has the ability and is required to draw conclusions from what is apparent in a disclosure and also what is implied. In

his response he often "stretches" his imagination somewhat in interpreting prior art as being anticipatory. Under the duress of time the examiner may do this with the idea of leaving it up to the applicant to find the mistake in his reasoning. The examiner will often be looking for a combination of ideas and will have to draw on patents in several subclasses to make his point. In so doing he has "indexed" those patents in a way that an indexer could not. His frame of reference is quite different when searching the prior art than it is when classifying claims.

One of the greatest advantages of a Citation Index is that it brings together patents that are seemingly unrelated, i.e. unrelated in terms of their principal subject matter. The closer two patents are in their obvious subject matter the less chance that either one would be overlooked in a routine search. Hence, it would be desirable to organize a Citation Index for all classes of patents so that related patents in different main classes could be found. In order to properly illustrate this advantage one would have to compile a complete Citation Index. The example shown in Fig. 1 was drawn from a sample based on a file of 5,000 accessible chemical patents. It is not a completely random sample nor is the example a complete Citation Index entry. However, even this incomplete sample will illustrate the point.

The principal patent in question is one granted to Schoeller in 1934 on "Wetting and Dispersing Agents for use in the Textile Industries." This patent was originally classified in class 260-98. It has since been reclassified to 260-458, because of the revision of class 260 in 1938.

This class covers patents on carbocyclic or acyclic carbon compounds which are esters (and processes of making same) of the sulfoxy acid type but limited to acyclic sulfuric acid esters.

This particular patent was subsequently cross referenced at the patent office to a multitude of sub-classes including the following:

8- 34 , 93 , 127 , 138, 139  
167- 91 ,  
252- 8.7, 8.8, 8.9, 356, 357  
260-400, 404, 410.6, 468, 488, 563, 584, 614

On the other hand the CA abstract for this patent was indexed under "dispersing agents" and under "wetting agents." There was no organic compound indexing.

The patents which subsequently referred to Schoeller's patent are shown. The first patent was issued to Lerner in 1948 on "Condensation products of Cholesteryl Esters with Polyethylene Glycol & Process for Producing Same." The principal class for this patent is 260-397.2, sterols (including Vitamin D). It was cross-referenced in 252-354. It should be noted that neither of these classes corresponds to any of the numerous sub-classes in which the cited patent was placed. It was not possible to completely determine how this patent was indexed by Chemical Abstracts. It was not indexed under the heading sterols. It was indexed under "Cholesterol Esters" and under "Glycols, Polyethylene."

The next patent was issued to Brown the same year on "Waxy polyol ether-esters." Its principal class is 260-234, Carbohydrate esters. It was cross-referenced in 260-210 and 260-410.6, Sub-class 210 is glycosides. Neither the cited patent nor the Lerner patent fall in this class. The sub-class 410.6 covers "synthetically produced higher fatty esters with acyclic polyoxy alcohols." This sub-class subtends the Schoeller patent, but not the Lerner patent. This same patent was indexed under the following in CA: Waxes, Ethers, Esters, polyol, Hydroxy compounds, Palmitic acid, Stearic Acid, Glycerol, Sorbitol, Mannitol, *D*-Glucose, Ethylene oxide, propylene oxide, lubricants, cosmetics, yarns, sizes, polishing materials, and coatings, plus some undetermined organic compound headings.

I don't think it is necessary to labor this point any further. A summary of this information will be found in the appended table. It is not intended here to criticize the classifications assigned to these patents by the Patent Office, nor the indexing performed by CA. The sample

does demonstrate, I believe, that a Citation Index will bring together related patents, the principal or apparent subject matter of which are unrelated.

In order to determine more precisely how these patents do relate an examination was made of the patent wrappers in question at the Patent Office in Washington. The records of the Lerner patent show that the examiner did not find the Schoeller patent in his search since the classes he searched did not subtend this patent. He found it in his own files or memory or elsewhere. He states in his response to Lerner's original application "Schoeller describes condensation of various acids including stearic acid and its ester with polyethylene glycols."

In the wrapper for Brown's patent the examiner states "Schoeller discloses the reaction of a carboxylic acid and diethylene glycol and ethylene oxide. There is no invention in substituting a trihydric alcohol, or a similar polyol in place of the dihydric alcohol." Examination of these comments demonstrates the relationship between the three patents involved, the frame of reference in this instance being condensation of acids with polyethylene glycols. Thus, the Lerner and Brown patents were brought together in the Citation Index through the Schoeller patent. One might start with the Lerner patent (found in CA under polyethylene glycols), find the Schoeller patent as a reference at the end of the specification, and through the Citation Index pick up the Brown patent. (One must always remember that patents are not classified by the subject matter of the disclosure but rather according to the claims that have been granted. However, cross references will be made to certain searchable disclosures in the patent specification). Citations are provided for a variety of reasons, but principally to disqualify certain claims. Citations are also made frequently in order to restrict the scope of the application. The examiner will refer to a patent whose scope he thinks covers the subject matter in question. Other citations may be used during discussions of legalities.

Citations in patents often include references to the literature or CA abstracts. The utility of a Citation Index including these citations would be increased but that would be the subject of a more detailed presentation.<sup>2</sup>

I do not intend to make any suggestions at this time concerning the possible compilation of a complete Citation Index to patents. Some time ago a formal proposal was submitted to the Patent Office, suggesting that a feasibility study be conducted.<sup>3</sup> If anyone is interested they can question the Patent Office on the matter. Pending action by the Patent Office, which seems doubtful, I am hopeful that Chemical Abstracts and/or Shepard's Citations may be able to undertake the compilation of a Citation Index to patents. As recently noted in C&EN CA has indicated that Citation Indexes will one day be studied by their new research group.<sup>4</sup> About ten years ago Mr. Harry C. Hart approached Shepard's Citations. Both he<sup>5</sup> and subsequently Mr. Arthur Seidell<sup>6</sup> communicated with the Patent Office again to no avail. Both of these gentlemen presented the viewpoint of the Patent Attorney.

The lack of action in the Patent Office is particularly unfortunate since a system of this type for the exclusive use by the examiners could include references to abandoned applications which are not classified at all and which contain a wealth of citation information insofar as the citations resulted in abandonment. It is not unlikely that many searches have been repeated by examiners in part or in whole when an examination of the appropriate abandoned file would turn up much useful data. Examiners do often remember these abandoned files. However,

<sup>2</sup> Garfield, E. Citation Indexes for science. *Science* 122, 108-11 (1955).

<sup>3</sup> Though the Patent Office considered the citation Index a basically good idea, it was concluded that the time required to compile and use it would not justify acceptance of the proposal. The author does not agree with these conclusions and hopes that future research will resolve this difference of opinion.

<sup>4</sup> Crane, E. J. The Chemical Abstracts Service—Good Buy or Good-by. *Chem. Eng. News* 33, 2752-4 (1955).

<sup>5</sup> Hart, Harry C. Re: "Citation System for Patent Office. *J. Patent Off. Soc.* 31, 714 (1949).

<sup>6</sup> Seidell, Arthur H. Citation System for Patent Office. *J. Patent Off. Soc.* 31, 554 (1949).

it is difficult to see why they have to rely on their memories for abandoned files and not for published patents. Classification is used precisely because the human memory has its limits.

I anxiously seek your comments and criticisms and hope to present data in the future as to how this technique can be used to advantage even in small patent files.

I will be glad to elaborate on machine methods available for compiling a Citation Index during the discussion period if there is time.

In closing I should like to mention that Miss Margaret Courain, Supervisor of the Research Files Division at Merck prepared the thousands of citation cards required in this experiment—and that without her work this data would not have been obtained.

I also want to thank Dr. Charles Bernier who helped in providing information on CA indexing and has made many valuable suggestions. I am also indebted to Mr. Harold Lindenmeyer who helped obtain some of the information on Patent Office cross-references.

Title	Number	Date	Inventor	Pat. Off Class & Cross Ref. Nos.
Assistants for the Textile and Related Industries	1,907,578	Aug. 21, 1934	G. Schoeller and M. Wittwer	260-458*
				400
				404
				410.*
				468
				488
				563
				584
				614
				167- 91
				8- 34
				93
				127
				138
139				
252- 8.7				
8.8				
8.9				
356				
357				
Condensation products of Cholesteryl Esters with Polyethylene Glycol and process for producing same	2,437,261	Mar. 9, 1948	L. Lerner	260-397.2* 252-354

Subject Matter of Sub-Class**	Examiner's Comments *Main Class	CA Indexing
<p>Acyclic sulfuric acid esters            Sulfoxy containing fats            Nitrogen containing fats            Synthetically produced higher fatty esters with acyclic polyoxy alcohols            Carboxylic acid esters            Carboxylic acid esters, acyclic acetic series, unsubstituted            Amines            Amines, acyclic, oxo or oxy            Ethers, acyclic            Cosmetics, skin, creams and lotions            Vat or sulfur dyes            Misc. dye compositions, with dispersing agents, organic, oxy or hydroxy compds. or their sulfur analogs            Chem. modification of textiles and fibers, mercerizing, wetting addition or pretreatment            Cleaning or laundering, degumming or desizing            Cleaning or laundering, scouring, degreasing or bowking            Textile treating, org. sulfoxy containing            Textile treating, org. amine containing            Textile treating, ether containing            Colloids, wetting agents, protein or carboxylic compd. containing            Colloids, wetting agents, org. amine, amide or N-base containing</p>	<p>**See entire classification schedules for complete description of sub-classes</p>	<p>Dispersing agents            Wetting agents            CA 28, 6332(7)</p>
<p>Sterols (including vit. D)            Wetting agents, org. sulfoxy, protein or carboxylic containing</p>	<p>Schoeller describes condensation of various acids including stearic acid and its ester with Polyethylene Glycols.</p>	<p>Cholesterol esters            Glycols Polyethylene            CA 42, 461c</p>

Title	Number	Date	Inventor	Pat. Off Class & Cross Ref. Nos.
Waxy Polyol Esters	2,450,079	Sept. 28, 1948	K. Brown	260-210 234* 410.6
Manufacture of materials resistant to or active against microorganisms	2,451,149	Oct. 12, 1948	Boehm	167-30* 31 32
Process for breaking Petroleum emulsions	2,500,349	Mar. 7, 1950	DeGroote	252-340* 260-410.6
Aqueous solutions of Lipoid-Soluble vitamins	2,518,230	Aug. 8, 1950	Freedman	167-81* 99-11
Aqueous solutions of Vitamin A esters	2,541,285	Feb. 13, 1951	Rawlins	167-81*

Subject Matter of Sub-Class**	Examiner's Comments *Main Class	CA Indexing
Carbohydrates, glycosides Carbohydrates, esters See above	Schoeller discloses the reaction of a carboxylic acid and diethylene glycol and ethylene oxide. There is no invention in substituting a trihydric alcohol or a similar polyol in place of the dihydric alcohol.	Waxes; ethers; esters, polyol; hydroxy acids; Palmitic acid; Stearic acid; Glycerol; Sorbitol; Mannitol; D-glucose; Ethylene oxide; Propylene oxide, lubricants; cosmetics, yarn, sizes, polishing materials, coatings CA 43, 886d
Insecticides, etc., carboxylic Insecticides, etc., carboxylic, phenolic Insecticides, etc., carboxylic, naphthalene	See example 10 in patent to Schoeller et al., and particularly the disclosure of Ethylene Glycol Monocresyl Ether.	Disinfectants Ethylene Glycol CA 43, 352d
Resolving colloids, liquid primary dispersants, org. agents, protein or carboxylic containing See above	Nothing specific about S. said.	Emulsions, Petroleum, Stearic acid, ethers (cydic); Stearic acid, esters; Pentaerythritol (esters of Stearic acid); Polypentaerythritols (esters of Stearic acid)—no green marking CA 44, 5089l
Vitamins	S., who shows applicants compds. used for dispersing and solubilizing pharmaceutical agents. No invention is found in applying these ether compds. of S. for solubilizing fat-soluble vitamins.	Vitamins A, D, E No green marking CA 44, 11,040l
See above	Rejected as lacking invention over S. who shows applicants esters used as dispersing and solubilizing agents for fats and oils in pharm. industry. To employ the S. ester as a solubilizer for Vitamin A esters would not involve invention.	Vitamin A, B <sub>2</sub> , C, B <sub>6</sub> , B <sub>1</sub> , E, K; Esters, nicotinamide Pantothenic acid (Na Salt); Stearic acid (monoester with Polyethylene Glycol); glycols, Polyethylene (esters of); Palmitic Acid; Arachidic Acid; Margarin, Linoleic and Oleic Acid CA 45, 4892d