

Online catalogue of scientific papers published before the 20th century - a proposal

Roberto de Andrade Martins
Rmartins@ifi.unicamp.br
Group of History and Theory of Science
Universidade Estadual de Campinas (Unicamp), Brazil

Abstract. This paper discusses the possibility and challenges of creating an international database of scientific papers published before the 20th century. The project would include articles published worldwide on all scientific, medical and technological subjects, from the 17th century to 1900. The starting point for the creation of the database would be a digital version of relevant bibliographical sources, such as Reuss' *Repertorium*, the Royal Society *Catalogue of Scientific Papers*, and several other specific bibliographies. This core database would be then gradually supplemented with contributions from several countries. The final size of the database can be estimated as 3-5 million entries.

Introduction

Periodical articles are one of the primary sources used by historians of science. It is well known that the first scientific journals were published in the 17th century. At that time, they were not as significant as books, as a form of communicating scientific results. The situation gradually changed, and in the 19th century scientific papers replaced books as the main form of presenting new scientific results, in most scientific disciplines.

How can historians of science find primary relevant scientific articles for their research? In the case of contemporary science, they can use several available scientific databases to find articles published in the last few decades, but there are just a few databases including information on older periodicals¹. If a historian of science is dealing with early 20th century science or any older period, he/she should use published bibliographies and indexes to find relevant articles.

Is it desirable to develop a database (or a set of databases) describing the old scientific papers? Is it possible, from the economic point of view?

The output of scientific papers increases in time following an exponential law, as remarked by Sola Price. The number of scientific papers published each year, today, far surpasses the whole production published from the inception of the *Philosophical Transactions* up to the end of the 19th century. It would be easy (and inexpensive) to develop a database describing the scientific papers published in the 17th and 18th centuries. It is much more difficult and costly to include in the database the papers published in the 19th century but, as shown below, this is not beyond the reach of the history of science community. However, attempting to produce a database encompassing the whole scientific production up to the period covered by existing scientific databases would be a far too expensive project, which cannot be undertaken by the history of science

¹ Some scientific databases do include references to old papers, although their main scope is contemporary science. The *GeoRef* database (<http://newfirstsearch.oclc.org/dbname=GeoRefS>) includes information about American Journals back to 1785 (but its international coverage includes only the 20th century). There is also a digital version of the *Jahrbuch über die Fortschritte der Mathematik* (www.emis.de/projects/JFM/), describing mathematical papers from 1868 to 1942. Some other examples will be dealt below.

community.

The following sections describe a possible strategy for developing a database of scientific papers published from the 17th to the end of the 19th century. This is a contribution to the “World History of Science Online: databases of bibliographical and archival sources project”, sponsored by the Division of History of Science (DHS) of the International Union of History and Philosophy of Science (IUHPS). The proposal presented here is grounded upon the author’s experience obtained in the development of the *Lusodat* project² – a database on science, medicine and technology in Portugal and Brazil, from the 15th century to 1900.

Scientific articles and journals: what should be included?

All scientific journals (including technology, medicine, social sciences and humanities) of the 17th, 18th and 19th centuries, published in any country, in any language, should be included in the database. However, what counts as a “scientific journal” is somehow arbitrary. Should journals and papers on astrology and other “pseudo-sciences” be included? Should we include the social sciences – history, anthropology, psychology, sociology, economy, statistics, philology, philosophy, ...? I think that the answer should be certainly “yes”, when the periodical or paper presents research results – independently of the field, the novelty, the methodology and the results. On the other hand, political speeches on the economical situation of a country, and other documents that present mere opinions about economy, politics, etc., should not be included. Mere matter-of-fact descriptions of the historical or political situation should not be included, too: they are a relevant source of information for historians, but they are not the result of historical analysis. Historiography should be included; the whole mass of non-historiographic documents that might interest historians should not be included – otherwise, the whole content of all newspapers would become relevant.

Another way of answering to this question is this: anything pertaining to the subject studied by historians of science (including medicine and technology) should be included. Political history is not history of science, and therefore the documents that are relevant for the research of a political historian should not be regarded as relevant for historians of science (unless there is some other reason to include them).

Should cultural periodicals be included? Perhaps. There were many illustrated periodicals, in the late 19th century, which included scientific news among its subjects. Historians of science who study the diffusion of scientific ideas among the public would like to have information about those periodicals. Should newspapers be included? Possibly, both for the reason presented in the case of cultural periodicals, and because sometimes there were fierce discussions of scientific issues in the pages of newspapers. However, if one attempts to include *all periodicals* in the database, the project will become unfeasible. A practical compromise could be adopted: journals exclusively dedicated to the sciences (including technology, medicine, social sciences and humanities) should be included, even in the case they do not contain original research papers (e.g., popular science); journals that do not address scientific issues should not be included; borderline cases (periodicals where some scientific papers can be found) can be included if someone thinks they should be

² See <http://www.ifi.unicamp.br/~ghtc/lusodat.htm>. The project is described in: MARTINS, Roberto de Andrade. Building a bibliographical data-base on old science, medicine and technique in Portugal and Brazil. *Quiipu – Revista Latinoamericana de Historia de las Ciencias y la Tecnologia* **11** (3): 311-32, 1994; MARTINS, Roberto de Andrade. Sources for the study of science, medicine and technology in Portugal and Brazil. *Nuncius - Annali di Storia della Scienza* **11** (2): 655-67, 1996.

included and takes to himself the burden of describing and indexing them.

Database of scientific periodicals

As a preliminary step to the development of a database of scientific papers, it is necessary to produce a database on scientific journals. Journals should be described as a whole, both when they are very important (that is, when all or most of their contents are scientific papers) and when they contain a reasonable number of relevant articles (for instance, some cultural journals should be included). Other serials are also to be included here (such as irregular publications, annual publications, etc.).

The main bibliographies that can be used to produce the periodicals database are:

- SCUDDER, Samuel H. *Catalogue of Scientific Serials of all countries including the transactions of learned societies in the natural physical and mathematical sciences: 1633-1876*. Cambridge: Library of Harvard University, 1879. Reprint: New York: Kraus Reprint Corporation, 1965.
- KRONICK, David. *Scientific and Technical Periodicals of the Seventeenth and Eighteenth Centuries: a Guide*. Metuchen: Scarecrow Press, 1991.
- BOLTON, Henry Carrington. *Catalogue of Scientific and Technical Periodicals, 1665-1895; together with chronological tables and a library check-list*. 2nd ed. Washington: Smithsonian Institution, 1897. Reprint: New York: Johnson Reprint Corp., 1965.
- GASCOIGNE, Robert Mortimer. *A Historical Catalogue of Scientific Periodicals, 1665-1900: with a Survey of their Development*. New York: Garland Publishing, 1985.

Using this information, it is possible to produce the core database of scientific periodicals up to 1900. They will include the periodicals where most of the scientific papers of the 17th, 18th and 19th centuries were published. However, this database will be far from “complete”. Two main obstacles will be noticed: (a) those bibliographies have a poor coverage of publications from East Europe, Latin America, Africa and Asia; (b) the social sciences and humanities are usually excluded from the scope of those bibliographies.

The core periodicals database can be complemented using information available from the main online library catalogues. Even after this is done, journals published outside the main European countries will not be fairly represented. During the 19th century, many short-lived scientific periodicals appeared all over the world, and most of them are not described in the bibliographies cited above, and cannot be easily found. Some of them will be only available in the national libraries of the respective countries (or sometimes in other smaller libraries). An international effort will be needed to complement the core database of scientific journals, adding information from all countries.

Articles: 17th and 18th centuries

There is no available database on old scientific papers. There are, however, some very useful printed bibliographies. The oldest one is Jeremias Reuss' (1750-1837) *Repertorium*, published in the early 19th century. It attempted to describe all articles published in journals of scholarly societies:

- REUSS, Jeremias David. *Repertorium Commentationum a Societatibus Litterariis Editarum, Secundum Disciplinarum Ordinem*. Gottingae, 1801-1821. Reprint edition: New York: Burt Franklin, 1961. 16 vols.³

The title of the book means: “Index of Articles published by Scholarly Societies, arranged by discipline”. This monumental work included all scientific disciplines,

³ The same Publisher issued a new edition in 1990.

medicine and technology. Each of the 16 volumes is dedicated to a specific discipline, and articles are classified according to subject. There is also an author index. The subject headings and the editor's comments are written in Latin⁴.

As a starting point, it would be possible to produce a digital version of Reuss' *Repertorium*, describing articles published up to the end of the 18th century in the journals of scholarly societies. This database would be the initial core of the 17th and 18th centuries database, but it would be necessary to complement it, in several ways:

- The subject headings used by Reuss were in Latin (and in German, in the last volumes). It is necessary to translate those subject headings into English (and other languages), to facilitate the search of the database. It is also desirable to include modern counterparts to some of the subject descriptions he used.
- Reuss' *Repertorium* did not include papers published in "independent" journals (those that were not published by scholarly societies). Therefore, as a second step, it would be necessary to identify the relevant journals that were excluded by Reuss, to index them and to add this information to the core database.

In the 17th-18th centuries period, the number of scientific journals published outside Europe seems negligible, and it will be quite easy to complement the information of the core database, in order to include those periodicals.

Articles: 19th century

The starting point for the production of a database of scientific papers published in the 19th century should be the Royal Society *Catalogue of Scientific Papers*. This work describes the content of a large number of scientific periodicals (about 3,000) published all over the world, and produced an index of about 800,000 papers. A digital version of this *Catalogue* would become the core database of scientific papers for the 19th century.

During the 19th century and early 20th century, the Royal Society of London sponsored a huge bibliographical project, intended to produce an index of scientific papers published in the main scientific journals of all countries, mostly in European languages. The result was a set of 19 volumes of encyclopaedia size:

- ROYAL SOCIETY OF LONDON. *Catalogue of Scientific Papers, 1800-1900*. London, 1867-1902; Cambridge, 1914-1925. Reprint: Metuchen: Scarecrow Reprint Corp., 1968. 19 vols.⁵

Several large libraries in the United Kingdom co-operated with this project, providing information on the content of their collections of periodicals. Articles published in the most important scientific journals were included, but papers on medicine and technology were not indexed. A large number of referees helped to select what should be included in the *Catalogue of Scientific Papers*.

It is difficult to evaluate the total number of references contained in the *Catalogue of Scientific Papers*. Volumes 1 to 12 contain about 33 references per page, and volumes 13 to 19 contain about 55 references per page. The total number of entries is probably between 750,000 and 800,000, describing articles from nearly 3,000 different serials.

Notice that each entry usually contains information about a single article, but sometimes it contains information about several versions/translations/editions of the same article. Some entries contain correction to previously published information, or remissive

⁴ For additional information, see <http://www.scholarly-societies.org/history/reuss.html>

⁵ The whole set of 19 volumes have been scanned for the Gallica project, and they can be downloaded, in PDF or TIFF format, from the French *Bibliothèque Nationale*: <http://gallica.bnf.fr>.

information.

In each series of the *Catalogue* the papers were classified according to the author name – that is, there was no subject classification. Each entry included the author's name, title of the paper in the original language (except in the case of non-European languages, such as Japanese), abbreviation of the title of the periodical, volume, initial and final pages, year of publication. The first volume of each series contained a list of indexed periodicals, with the corresponding title abbreviations.

The Royal Society project included the production of a subject index for the whole period. Preparation of the indexes began as the last volumes appeared, and it was to include 17 parts. From 1908 to 1914, three subject indexes were prepared: v.1 for Pure Mathematics (1908); v.2 for Mechanics (1909); v.3, pt.1 for Physics: Generalities, Heat, Light, Sound (1912); and v.3, pt.2 for Physics: Electricity and Magnetism (1914). The beginning of World War I interrupted this work, which was never resumed.

- ROYAL SOCIETY OF LONDON. *Catalogue of Scientific Papers, 1800-1900: Subject Index*. Edited by Herbert McLeod. Cambridge: University Press, 1912-1914. Reprint: Metuchen: Scarecrow Reprint Corp., 1968. 3 vols in 4 parts.

After the conclusion of the *Catalogue of Scientific Papers*, the Royal Society began (in 1902) the publication of the *International Catalogue of Scientific Literature*. This is a subject index to the scientific publications appearing in each year, describing the contents of about 1,800 periodicals. This publication was also interrupted due to World War I, and was not resumed.

One severe limitation of the *Catalogue of Scientific Papers* is the lack of a subject classification – except in the case of mathematics and physics, as described above. Therefore, to produce a useful database, it is necessary to ascribe a subject to each of the 800,000 papers, and this is not an easy job.

The best approach seems to be ascribing subjects to each paper according to the title of the paper, the journal where it was published (if it is a specific journal) and the subject of other papers published by the same author. This task could be done by a team of people who are familiar with the several languages used in the *Catalogue* (English, French, German, Italian, etc.) and who are also familiar with the several scientific disciplines, in order to detect the subject of each paper. Of course, this strategy may produce many mistakes, but it is better to have a tentative subject classification than no classification at all.

The production of a digital version of the *Catalogue of Scientific Papers* would be a huge work, but a straightforward one. The cost of production of this core database is expected to be about US\$250,000.00.⁶

After the completion of this core database of 19th century scientific papers, it would be necessary to complement the database, because of several limitations:

- The Royal Society *Catalogue* did not include medical and technical papers. This is a severe drawback of the project, because the 19th century was the time when medicine underwent a profound revolution, and technology transformed the way of life of mankind.
- The social sciences and humanities are not contemplated in the *Catalogue*, either.
- Even in the case of the sciences included in the *Catalogue*, the coverage is far from

⁶ A preliminary electronic version of the subject indexes of the *Catalogue of Scientific Periodicals* was launched in December 2004 by Paratext (www.paratext.com/19cm_i.htm). This company intends to produce a database containing the full content of the *Catalogue*.

complete, because there was a *selection* of relevant papers by a group of referees and, besides that, many periodicals have not been indexed.

- Mainstream journals were included, but peripheral journals were seldom taken into account.

The first limitation is a serious one. The output of medical and technical papers, during the 19th century, was probably similar to (and perhaps larger than) the output of scientific papers. If there existed a bibliography similar to the *Catalogue of Scientific Papers* including those medical and technical papers, the production of a digital version of that bibliography would cost as much as (and perhaps more than) the core database. However, since such a bibliography does not exist, it would be very difficult to produce a suitable complementation of the core database in a short time and at a reasonable cost.

This limitation should not be used as an argument against the project, however. It is better to have an initial database of scientific papers without medicine and technology than nothing at all.

In the case of the social sciences and humanities, the output of papers in the 19th century is probably smaller than that of medical papers, but it is certainly considerable. In the case of English language papers, a large fraction of the relevant works can be found in Poole's *Index to Periodical Literature*. This work began by William Frederick Poole (1821-1904) and continued by William Isaac Fletcher (1844-1917) contains about 400,000 citations of articles published in American and British general periodicals from 1802 to 1907.

- POOLE, William Frederick; FLETCHER, William Isaac. *Poole's index to periodical literature*. Boston: James R. Osgood / Houghton, 1882-1908. 7 vols

There is a commercial electronic version of *Poole's Index* that can be accessed by subscription, by Paratext. There is also another humanities and social sciences database comprising both old and new papers: *Periodicals Content Index* (PCI), provided by ProQuest / Chadwyck-Healey⁷. This database contains an index of 20th century periodicals in the humanities and social sciences. Periodicals in the 20th century that extend back into the 18th and 19th centuries are indexed from their earliest volumes (back to 1770). The scope is worldwide and includes journals in English, German, Italian, French, Spanish, and other Western languages.

Both *Poole's Index* and the relevant part of the PCI database should, of course, be incorporated in the project presented here. A negotiation with the owners of those databases could be made, asking permission to use their data and offering as a counterpart the information that will be gathered and that is not contained in those resources.

The core database should be gradually complemented in two ways:

- By directly indexing periodicals that have not been included in the *Catalogue of Scientific Papers*.
- By the use of specific bibliographies containing information about papers published during those centuries.

In the case of journals that have not been indexed by the *Catalogue*, an international effort will be needed to complement the core database, adding information from periodicals published in all countries. The national committees should check which relevant journals had not been included in the *Catalogue*, and next it will be necessary to find complete sets of those periodicals, to examine them and to produce complete indexes of their articles.

⁷ <http://pci.chadwyck.com/>

Towards the middle of the 19th century, Johann Poggendorff (1796-1877) began to develop a biographical and bibliographical handbook. Each entry, in alphabetical author order, presents a short biography and then a list of the works (books and articles) published by the author. Only the exact sciences were included. There is no subject index. The bibliographical information provided by Poggendorff is not full, but nevertheless this is a useful tool.

- POGGENDORFF, Johann Christian (ed.). *Biographisch-literarisches Handwörterbuch zur Geschichte der exakten Wissenschaften*. 2 vols. Leipzig: Johann Ambrosius Barth, 1863.

Poggendorff's *Handwörterbuch* was continued in successive series. Two of them include information up to the end of the 19th century and should be used by the present project:

- FEDDERSEN, Berend Wilhelm & OETTINGER, Arthur J. von (ed.). *J. C. Poggendorff's biographisch-literarisches Handwörterbuch zur Geschichte der exakten Wissenschaften (1858 bis 1883)*. Vol. 3. Leipzig: Johann Ambrosius Barth, 1898.
- OETTINGER, Arthur von (ed.). *J. C. Poggendorff's biographisch-literarisches Handwörterbuch zur Geschichte der exakten Wissenschaften (1883 bis 1904)*. Vol. 4. Leipzig: Johann Ambrosius Barth, 1905.

The 19th century: specific bibliographies and indexes

Around the middle of the 19th century, as the number of articles increased, scientists felt the necessity of specific indexes (classified by subject) describing the literature published in each year. The first specific indexes, published in the 19th century, described mathematical and physical papers:

- *Die Fortschritte der Physik* (began in 1844)
- *Beiblätter zu den Annalen der Physik* (began in 1877)
- *Revue Semestrelle des Publications Mathématiques* (began in 1893)
- *Abstracts of Physical Papers from Foreign Sources*, published by the Physical Society of London (began in 1895)
- *Science Abstracts – Physics* (began in 1898), later called Physics Abstracts

In 1890 there appeared the *Readers' Guide to Periodical Literature*, followed in 1907 by the *International Index to Periodicals*, using keywords (instead of subject classification) and presenting an index of papers from cultural and philosophical journals, besides some general scientific journals.

Bibliographies on *scientific books* on several subjects began to be published in the 17th century. Some of the oldest are William Cooper's *Catalogue of Chymicall Books*, and Cornelius Beughen's *Bibliographia Mathematica*, that included also mechanics and astronomy, besides pure mathematics.

- COOPER, William. *A Catalogue of Chymicall Books*. London, 1675.
- BEUGHEN, Cornelius A. *Bibliographia Mathematica et Artificiosa Novissima Perpetuo Continuada, seu Conspectus Primus*. Amstelodami: Janssonio-Waesbergios, 1688.

Many other important scientific bibliographies were published during the 17th and 18th century, including only books.

Up to the 18th century, scientific papers were not as relevant and books, but in the 19th century the scientific bibliographies began to include articles, besides books. One nice example is Houzeau and Lancaster's *Bibliographie Générale de l'Astronomie*.

- HOUZEAU, J. C. & LANCASTER, A. *Bibliographie Générale de l'Astronomie jusqu'en 1880*. London: Holland Press, 1964. 3 vols.

The authors used a lot of bibliographic instruments that were already available (such as Lalande's *Bibliographie astronomique*) and also the *Catalogue of Scientific Papers*, and complemented those sources with the direct analysis of many periodicals. Only half of the articles included in Houzeau and Lancaster's book can be found in the *Catalogue of Scientific Papers*. This shows both that Houzeau and Lancaster did a very careful work, and that the *Catalogue* is far from complete.

As has already been commented, the number of scientific papers produced in each year increases exponentially in time. If one were to include in the database the articles published from 1901 to 1920, this would double the size (and cost) of the project. It seems desirable to limit the project to the 17th, 18th and 19th centuries, leaving the 20th century to future projects.

Database structure

A useful database should contain a huge amount of information together with suitable search instruments. The usual bibliographical databases are built using as a model the old card catalogues and can only be searched using words from the title, author name and subject fields. The database proposed here should have a suitable structure to allow more sophisticated searches⁸. When a historian of science searches a database of scientific papers, he might have different questions in mind, and the database should be devised to help him to answer those questions. It should be possible to make searches in the article database using several different criteria:

- title of the article
- title of the periodical where it was published
- author(s) of the article
- editors, and other people related to the periodical (editorial board, etc.)
- institution that sponsored the periodical
- place of publication (city, country, region)
- language(s) of the paper
- subjects of the article
- year of publication
- place where the document can be found (library, city, country...)

Truncated search terms should be allowed, of course. (that is, searching for medic* should find all entries containing terms beginning by "medic"). Any combination of the above criteria (using Boolean operators) should also be allowed by the database. Therefore, without entering any personal name and title data, a historian should be able to find articles published in Germany, in Latin, in the decade of 1780, on astronomy or astrology (or containing the truncated keyword astro*).

Final comments

The historical database of scientific papers proposed here would be enormous. As described above, the *Catalogue of Scientific Papers* of the Royal Society contained about 800,000 references of 19th century journals. However, the number of astronomical papers

⁸ The technical details of the proposed database structure are described elsewhere. See the following documents, available in the Internet: <http://www.ifi.unicamp.br/~ghtc/sources/articles2.htm>; <http://www.ifi.unicamp.br/~ghtc/sources/articles3.htm>

published in the 19th century described by Houzeau and Lancaster is twice the number contained in the *Catalogue*. It is likely that the same situation occurred for the other disciplines covered by the *Catalogue*. However, the proposed database would contain much more than 1,600,000 references, because the *Catalogue* did neither include medical and technical papers nor articles concerning the social sciences and humanities, and that its coverage of peripheral countries was poor. Hence, it is possible to estimate that a more complete database of 19th century scientific papers would include certainly more than 3 million entries, and less than 5 million. For previous centuries, the number of papers is much smaller. Reuss's *Repertorium* describes about 100,000 papers, including medical and technical papers. Therefore, the size of the proposed database of scientific papers, from the 17th century to 1900, will contain from 3 to 5 million entries.

Let us compare this proposal with the current *Periodicals Contents Index*. This database of contemporary papers in the humanities and social sciences includes nowadays 15 million articles, and about one million new articles are added each year. Compared to the *Periodicals Contents Index*, the present proposal of a historical database of scientific papers does not seem so huge and unfeasible as it might appear at a first sight. Let us also remark that the historical database will not increase forever. Its size will steady approach an upper limit, and it will have a lasting value.

The cost of production of the database is expected to be about US\$1,500,000.00. Part of its production (around 70% of the final database) could be easily done because it would employ published indexes of periodical literature. However, it will be costly to add information about periodicals that have not been previously indexed. The inclusion of information about publications of peripheral countries would require an international collaboration.

The expenditure could drop to about US\$500,000.00 if it is possible to make use of selected parts of existing databases.

One might wonder if it is really necessary to engage in such a project. The number of commercial databases increases very fast, and in a few years there will be available resources containing more than 50% of the proposed historical database of scientific papers. However, there are some drawbacks in the current situation:

- Those databases are *commercial* ones, that is, only the researchers who work at institutions that can pay for their use can access their information. The largest and richest universities do subscribe to all or to most databases, but this does not apply to most universities worldwide. It is desirable to produce a free of charge database that could be used by all researchers, from all countries.
- It is unlikely that the companies that develop commercial databases will include information about the old periodicals of peripheral countries. It is desirable to produce a database containing information about scientific papers published in all countries, in order that historians of science may gain a more balanced view of the international scientific development.

There are some other minor shortcomings of the current commercial databases.

- It is impossible to make a joint search using databases from different companies. This practical problem will probably *increase* in the future. It is desirable to have a single database (or a set of databases that can be searched jointly) to simplify the work of the historian.
- The historical database of scientific papers should have richer searching strategies than the usual scientific databases. It is usually impossible to make searches, in the commercial databases, using criteria such as the country of publication, year of

publication or the publication language, for instance.

- The commercial databases of scientific papers do not contain information about the libraries where the journals can be found. This is not a serious limitation in the case of mainstream scientific journals, because they can be found in the large university libraries of first-world countries. In the case of old obscure and short-lived periodicals or journals from peripheral countries, it is desirable to include in the database the information about the libraries where they can be found. Historians of peripheral countries would also welcome information about the libraries in their countries where the periodicals can be found.

In conclusion: the project will not be superseded by the development of commercial databases in the near future. It will not merely duplicate current efforts. Some of its specific features will not be available in the databases that are under development. Hence, it will be a helpful and irreplaceable tool for the historian of science.

Although the project is feasible, it requires careful planning, skilful negotiation, adequate funding and an international joint action. A long-term commitment of the Division of History of Science (DHS), the International Union of History and Philosophy of Science (IUHPS), Unesco and other international institutions will be necessary to attain the proposed goals.

Acknowledgement

The author is grateful to the Brazilian *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP), that supported the development of the *Lusodat* project⁹.

⁹ *Lusodat* is a database on science, medicine and technology in Portugal and Brazil, from the 15th century to 1900 (see <http://www.ifi.unicamp.br/~ghc/lusodat.htm>). The proposal presented here is grounded upon the author's experience obtained in the development of the *Lusodat* database.