

SACO and Subject gateways
Prepared by Ana L. Cristán
Acting Team Leader, Cooperative Cataloging Team
Library of Congress

In researching the topic of my presentation it was my task to determine if the SACO Program component of the Program for Cooperative Cataloging (PCC) could be related to subject gateways or if it might be considered a subject gateway in its own right.

Brief Background

SACO is described as the subject component of the Program for Cooperative Cataloging. The PCC is an organization collaboratively led by the Library of Congress and the major cataloging utilities in the United States, OCLC and RLG, which is dedicated to enabling catalogers to create and make accessible bibliographic and authority records that are formulated and structured according to mutually acceptable standards.

That is to say that PCC catalogers create bibliographic and authority records using a set of rules, guidelines, and documentation to which they have all agreed to apply in their cataloging. The user is thereby provided uniformity and predictability when using these records in their cataloging.

Within the PCC framework it is SACO that allows catalogers to submit subject heading proposals for possible inclusion in the list of *Library of Congress Subject Headings* (LCSH) and allows for the modification of existing subject headings and/or modification of the existing subject heading hierarchy. LCSH, in the grand scheme of things, may be called a “boutique” file particularly when comparing its less than 300,000 authority records with that of the shared Name Authority File that contains 5 ½ million records, nonetheless it does not lack in popularity and usefulness. Both authority files reside at the Library of Congress where the staff is responsible for the maintenance and the day-to-day oversight of receiving and distributing these records. To reinforce what has been said by my colleagues at other presentations at this conference, there are a total of 4 component parts to the PCC: NACO the name authority record component, SACO the subject authority record component, BIBCO the bibliographic record component, and CONSER the serial record component. [Slide 1]

The Venn diagram shows the interaction of the component parts with each other and reveals that in some aspects SACO is independent of the others, a characteristic that this presentation will examine later. Information on these programs is available via the World Wide Web on the PCC Home Page at: <http://www.loc.gov/catdir/pcc/>

[Slide 2]

This presentation will focus its attention on the SACO program and grapple with its connection to subject gateways. The definitions and descriptions of subject gateways that may be found by searching the World Wide Web include the following:

[Slide 3]

Subject Gateways [slide 4-6]

The definition that appears to be the most straightforward is the one found on the Desire web site which states that: “subject gateways are online services and sites that provide searchable and browseable catalogues of internet based resources”

The definition goes on to say that “subject gateways will typically focus on a related set of academic subject areas.”

[Slide 7]

The TERENA - Trans-European Research and Education Networking Association focuses attention on librarians or information specialists in the use of subject gateway when stating that: “... subject gateways, select, classify and describe quality resources in a specified subject area. They [the subject gateways] effectively fill the role of information broker for information seekers in that subject [particular discipline] and the people selecting and describing resources are usually subject-specialists, for example librarians. ... gateways can be relied on to identify useful quality online resources, and to be an important resource for anyone working in a field in which there is a significant mass of online source material.”

Further web crawling points to a number of sources including the site by the National Library of Australia (NLA). NLA has developed an online resource that provides a “Best practice checklist for Australian subject Gateways.” The checklist presents a “core set of content, business, and functional characteristics” that is intended “to stimulate further debate on the coherent development of gateways, and to provide Australian gateways with a common framework for participating in the cooperative development of both national and overseas gateways” This site contains a grid which lists subject gateways and their characteristics in addition to a list of standards to be used in order for NLA to accept the site as a gateway. One of the thesauri included in NLA’s grid used to map to a specific gateways is indeed LCSH.

[Slides 8-10]

The levels of the subject gateways vary considerably and certainly the most sophisticated appear to be those developed by the European library community. PINAKES is one of these gateways. This site bills itself as a “subject launch pad” and serves not only a subject gateway but also supplies links to multiple subject gateways. **[Slide 11]**

There is at least one website designed to provide downloadable software for setting up a subject gateway. ROADS, the Resource Organisation and Discovery in Subject-based services is a singular example. It is funded by Joint Information Systems of the UK.

[Slide 12]

This source offers guidelines for using the software provided by ROADS and provides step-by-step instruction for building a more focused searching mechanism. The site features information regarding the usefulness of gateways and cites among the benefits that subject gateways seek to overcome the problems of retrieving huge numbers of irrelevant results. That is to say that when using search engines such as Yahoo, Google, or Lycos, users will get results that often do not provide meaningful or independent descriptions of the sites; thereby requiring the user to sift through an imprecise mélange. Other factors cited to show the usefulness of subject gateways include the points that most subject gateways provide link checking and that because the search is more focused the results are retrieved and displayed more efficiently.

Suffice it to say that subject gateways run the gamut from simple to sophisticated. The Library of Congress has a site called a “global gateway” which provides links to a list of resources grouped by country via its “portals to the world.” While this site does not provide a search engine as used by most subject gateways described earlier this site provides links to resources by country and therein by subject.

[Slides 13-14]

For the purposes of my discourse the definition of a “subject gateway” which best describes how the SACO Program might relate is found in the 1998 article authored by John Kirriemuir, et. al., which states that: “a subject gateway, in the context of network-based resource access, can be defined as some facility that allows easier access to network-based resources in a defined subject area. The simplest types of subject gateways are sets of Web pages containing lists of links to resources.”

The SACO Program has on its Home Page a list of resources for use in the creation of subject headings. This Web page has been developed and is maintained by PCC participant, Adam Schiff, Principal Cataloger at the University of Washington. This list of resource is available at: <http://www.loc.gov/catdir/pcc/saco/resources.html>

[Slide 15-16]

And much like the “Portals to the World” site, the Web Resources for SACO Proposals is a set of Web pages containing list of links to resources thus, according to the Kirriemuir definition it may be considered a simple subject gateway. For the SACO Program the access that this Web page provides is extremely important.

LCSH uses the MARC 21 communication format as its standard for providing a framework to enable data sharing in machine-readable form. Among the hallmarks of an LCSH authority record is the inclusion of the *MARC 21 Authority Format's* “sources found” field, often better known by its numerical equivalent, the “670 field.” This field is a requirement for all subject-heading proposals submitted to the SACO Program. The 670 field allows the cataloger to provide the literary warrant that justifies or authorizes the heading and/or may even be used to justify variant forms of the heading. In a shared environment the 670 field is an indispensable feature. In the case of SACO this field is used by the subject specialists at LC to provide editorial oversight without access to the item being cataloged. The LC subject specialists rely on the information provided by the cataloger in the 670 field for justification of the heading, possibly for justification of the variant forms, and for the determination as to whether or not the heading being created is appropriate to the work being cataloged.

Given the importance of providing evidence of research when constructing a subject-heading proposal for possible inclusion in LCSH it would certainly behoove the SACO Program to investigate building a more sophisticated subject gateway for SACO using this “List of Web resources” as a starting point. However, the ROADS documentation cites “people” as one of the key elements in building and maintaining a subject gateway. Given the current level of staffing of the SACO Program, it would be easier to link from the SACO Web resources page to one of a number of existing subject gateways. Certainly subject gateways are a valuable cataloging tool that catalogers developing

subject proposals for contribution to LCSH could benefit from using. Therefore, how does SACO relate to subject gateways? –Only if one wishes to use the word “gateway” in its most basic meaning that is to say that it provides a passageway from one area into another can one say that the SACO Program is a gateway. Indeed SACO is the “gateway” that offers the opportunity to catalogers and institutions that wish to have subject headings become part of the Library of Congress subject headings authority file. LCSH no longer restricts the development of subject headings that might be applied to catalog those materials housed in LC’s collection, instead via the SACO Program, LCSH incorporates subject headings proposed by catalogers from around the world.

History of SACO

The genesis of the SACO Program began with a letter sent in 1981 from the National Library of Australia to LC’s then Director for Cataloging, Henriette Avram requesting a mechanism for including Australian subject headings for local flora and fauna in LCSH. As it was structured in 1981, LCSH contained only those subject headings, which had been used to catalog items in LC’s collections. Experienced gleaned from other cooperative subject ventures and with the advance of technology, in 1992 the SACO Program was developed to allow subject headings that are developed in accordance with LCSH guidelines that may be used in bibliographic records to be added to LC subject authority file. Consequently, SACO now boasts participation from more than 80 institutions and at least 15 of these partners are from the international arena. A current list of those international partners who contribute on a regular basis includes, the British Library, Trinity College Dublin, the universities of Oxford and Cambridge as well as the national libraries of Scotland, Canada, Australia, New Zealand, and Singapore. [Slides 17-19]

It is interesting to note that many of the current international NACO partners began PCC participation by first submitting subject heading proposals via the SACO Program. A mainstream workflow for NACO contribution of name authority headings requires membership in one of the bibliographic utilities; however, for SACO there is no such requirement. For membership in the CONSER and/or BIBCO Programs an institution must first become a participating NACO member, yet there is no such requirement for participation in the SACO Program. For membership in NACO, CONSER or BIBCO an institution must first submit an application and agree to undergo training appropriate to each component program, yet there is currently no such requirement for SACO. It may be that lack of these specific program requirements may have encouraged international participation in SACO, but I offer two additional arguments for SACO’s universal appeal: 1) the use of the MARC authority format and 2) that English is one of the official languages of the catalogs of those participating institutions.

Barriers to participation

These two common bonds more than anything else, have enhanced the SACO Program’s growth. In fiscal year 2002 PCC participants contributed more headings to LCSH than did the staff at the Library of Congress. [Slide 20]

However, this growth has brought increased challenges to the PCC and to LCSH. Many non-English speaking countries are attempting to develop subject headings based on LCSH and find the process of translating LCSH laborious, frustrating, and not cost effective. In a survey conducted by Magda Heiner-Freiling in 2000 for an article in *Cataloging and Classification Quarterly*, on the occasion of LCSH's centennial, it was found that there are over 36 countries in which LCSH is used and that at least 12 of these institutions use "translations or adaptations of LCSH into their own language." Instead of translating the entire LCSH authority file, each institution has developed a mechanism for translating a subset of LCSH when needed for local use and there appears to not be a mechanism in place that allows for sharing of these files outside the immediate region.

In spring of 2002 at LC the subject specialists on staff at the Cataloging and Policy Support Office in concert with the Cooperative Cataloging Team joined to offer a workshop on LCSH to 17 catalogers from Latin American countries. The objective of this workshop was to lay a foundation for Spanish speaking catalogers to understand the principles and underlying structure of LCSH with the expectation that this would facilitate translation of LCSH into Spanish. I served as the coordinator of that workshop and it was because of LC's interest in the development of a Spanish language LCSH and the commitment made to support the work that is being done by El Colegio de México in developing a bi-lingual subject headings list that led to this workshop being held. In March 2001 I attended a seminar on cooperation in authorities in Mexico City sponsored by El Colegio de México that sought to bring together Mexican libraries into the systematic development of a shared database of Spanish and English LCSH. LC and the PCC continue to support cooperative efforts in Latin America and have recently reaffirmed the desire to seek solutions to barriers to international participation through the formation of a PCC Task Group on International Participation.

(<http://www.loc.gov/catdir/pcc/tgip.html>) [Slide 21]

Another effort to facilitate non-English subject headings into LCSH was the workshop presented at a Baltic Conference in September 1997; as a result the National Library of Lithuania has contributed some vernacular subject headings via the SACO process. An example of several of these headings can be seen here:

[Slides 22-25]

Incorporating non-English headings into LCSH is possible, particularly in areas such as geographic and man-made features including rivers, parks, buildings, and/or streets. In other areas the creation of vernacular headings are more problematic or prohibited by rules. If there is a predominance of English language forms of a term or concept found in English language reference sources, the guidelines require that the heading must be established in English. However, the foreign language equivalent term may not be added as a cross-reference unless the item being cataloged is also in English. There is work going on by IFLA committees and others, which you will hear about at this conference that may help to resolve this issue. As noted earlier, the PCC is aware of the need to play a role in the resolution of these barriers to expanded international participation and cooperation.

Current Methods of Contribution to SACO

Improvements in the SACO Program are of concern at LC, consequently, LC authorized an independent study to determine what steps are needed to be taken to decrease the amount of time it takes for a proposed subject heading to work its way through the LCSH editorial process.

Currently the proposal process of contributing new or changed subject headings for LCSH through SACO continues to take a minimum of 4-6 weeks. As noted earlier, NLA requested the addition of Australian subject headings via a letter in 1981; nonetheless, it was not until the early 1990s that LC began to accept electronic mail submission of proposals. Prior to that time contributors were asked to send forward proposals on LC forms via surface mail and later via fax.

In the late 1990s a Web form was developed and made available for participants to use; however, the SACO workflow remains a manual process, while the need for re-keying has been minimized, it has not disappeared. Catalogers are still required to manipulate the proposals in order for these to be integrated into the process of editorial review. The independent LC study concluded that it is a combination of the submission of, the integration into the database, and the subject editorial review process that play a role in the time lag for processing subject headings via SACO. Certainly the requisite editorial review will and should continue; I suggest that the SACO Program needs to develop, in concert with the bibliographic utilities, a mechanism for accepting proposals on a parallel with that of the NACO Program and its FTP transfer of name authorities to the LC master file. Earlier this week my colleague spoke of the great success of the NACO Program. This success can be attributed to the cooperation forged between the bibliographic utilities and LC to enable a contribution/distribution mechanism for processing name authorities in a timely, cost-effective manner.

The PCC Policy Committee (PoCo) is very aware of the need to meet the challenge of institutionalizing SACO. At their annual meeting in November 2002 the PoCo authorized a task force to develop a plan to transform the SACO Program into a full-fledged component of the PCC with guidelines and membership parameters.

This task force is charged with describing a scenario for contribution of subject proposals via the utilities or some other mechanism that would facilitate local review and eliminate re-keying. Parameters for participation, including goals for a minimum number of required proposals per annum will be addressed. The development of courses for beginning, advanced, and continuing subject cataloging education should become incorporated into the responsibilities of the PCC. Of course, the proposed curriculum would include training and documentation to facilitate the development of subject headings. The report from this task force is due in summer of 2003. Alas, it may be easier to describe and recommend what needs to be done than it is to carry out those tasks.

As mentioned earlier, one of the most challenging workflow issues for the SACO Program will be the need to enable LCSH to handle non-English language subject

headings. It is here that the power of the subject gateways may be harnessed or worked to an advantage given that subject gateways are now attempting to provide the capability of multi-lingual searching. The challenge may be met by the use of a “virtual international authority file” (VIAF) as described by Dr. Barbara Tillett in her presentations. As envisioned by Dr. Tillett the VIAF is designed to allow the linking of “authorized forms of names, titles, and even subjects” in order to maximize the benefits of shared authority work. There are models and prototypes of the VIAF currently being planned. The MACS (Multilingual Access to Subjects) project under development by the national libraries of France, Germany, Switzerland, and England also provides a glimmer of hope that these linguistic barriers may be surmounted. [Slide 26]

In the short term the answer may be as simple as expanding the use of the full range of MARC 7XX fields in subject headings as has been done for name headings in the NACO Program. Although implementation of the use of the 7XX field in subjects would not be the most cost-effective mechanism, it would allow the authorized form of headings for other national bibliographic agencies to be linked or made available via the LCSH subject authority record.

In the meantime while developments on the VIAF, MACS and other efforts are brewing the SACO Program continues to provide a mechanism for the development of LCSH that promotes a standardized file structure that can be described and used in a predictable and uniform manner. SACO helps to prevent the duplication of efforts by those that use LCSH in their cataloging. By working collaboratively the PCC continues to welcome the inclusion of new participants and seeks to implement new technologies.

In conclusion, it is not the SACO Program that interacts with subject gateways, but it is LCSH that may be used for mapping to its headings from other authority files and/or subject gateways in order to provide more focused searches. Thus, it stands to reason that the more complete and inclusive that LCSH is, the better it can be used as a vehicle for searching large number of resources and the easier it will be to fit those files into any new technology. It is via the SACO Program that this development is taking place.

A visit to the SACO Home Page will reveal that currently any institution may participate in SACO by agreeing to follow the current LCSH guidelines and submit a completed subject Web proposal. To help facilitate the process, sample forms, guidelines, and a subject gateway will enable prospective participants to fulfill the requirements for research and to speed that subject proposal through the process. The main requirements for creating a subject proposal that will be submitted to SACO are:

1. The use and application of the guidelines found in a current edition of LC’s *Subject Cataloging Manual: Subject Headings*. (4 vols. Available from the Cataloging Distribution Services at LC) [Slides 27-28]
2. The ability to search to a current copy of the LCSH authority file. This is done to assure that a heading is not a duplicate of an established heading or a heading that is in the process of being proposed. Note that a search via World Wide Web of the LC Web authorities module at: <http://authorities.loc.gov/> will find current

- valid subject headings as well as any new heading or modifications of existing headings being proposed without the need of an intervening vendor.
3. Access to a current copy of the *Free-Floating Subdivisions: An Alphabetical Index*. This document is necessary, as it will help prevent the proposing of an unnecessary heading given that it is constructed by means of a subject—subdivision formulation.
 4. Access to a current copy of the *MARC 21 Authority Format*. In order to be able to use LCSH properly it is necessary to recognize the MARC codes and fields used in subject authority records especially when checking to see if a heading or its reference structure is valid, etc.
 5. A MARC 21 identification code. This code is used in the MARC 21 Cataloging Source (040) field to identify the institution that is submitting a proposal. This field is necessary for statistics kept by the Cooperative Cataloging Team and may be obtained online at the LC's Network Development and MARC Standards Office at: <http://www.loc.gov/marc/organizations/orgshome.html#requests> [Slide 29]
 6. While not a requirement (at this time) attendance at a LCSH workshop is strongly encouraged. LCSH workshops may be provided “on demand” by contacting CPSO@loc.gov by contacting the Cooperative Cataloging Team at acri@loc.gov or by attending the SACO workshops held in conjunction with the biannual American Library Association conferences. Note that the workshops are free but that institutions requesting training may be asked to pay for the travel expenses of a trainer.

Thank you for your attention and I will be glad to answer questions or discuss with you on any aspects of the SACO Program or the PCC. [Slide 30]

Sources cited and consulted

Desire <http://www.desire.org/html/subjectgateways/subjectgateways.html>)

Heiner-Freiling, Magda. Survey on subject heading languages used in national libraries and bibliographies. *Classification and Cataloging Quarterly*, v. 29, no. 1/2 (2000): p. 189-198.

Kirriemuir, John, et. al., "Cross-searching subject gateways" *D-Lib Magazine* online at: <http://www.dlib.org/dlib/january98/01kirriemuir.html>

Library of Congress Web site <http://international.loc.gov/intldl/intldlhome.html>

MACS Web site <http://infolab.kub.nl/prj/macs/>

National Library of Australia (NLA) Web site
<http://www.nla.gov.au/initiatives/sg/index.html>

PINAKES Web site <http://www.hw.ac.uk/libWWW/irn/pinakes/pinakes.html>

ROADS Web site <http://www.ukoln.ac.uk/metadata/roads/what/>

Terena Web site <http://www.terena.nl/library/gnrt/websearch/gateways.html>

Tillett, Barbara B. A virtual international authority file (VIAF). In: *Record of a workshop on Authority Control among Chinese Korean and Japanese Languages (CJK Authority 3)*, March 2002, p. 117-139.