
The digital archive of the Swedish East India Company, 1731-1813: a joint project of a university library and a history department

Margareta Benner

The author

Margareta Benner is Head of the Central Library (Humanities and Social Sciences) at Göteborg University Library, Göteborg, Sweden.

Keywords

Digital storage, Indexing, Internet, Document image processing, Sweden

Abstract

The article describes a digitisation project in which Göteborg University Library, Sweden, is publishing a collection of eighteenth century manuscripts on the Internet, in collaboration with its history department. The aims of the project are to make one virtual collection of documents owned by different institutions, to give free and open access to the collection, to save heavily used documents from wear and tear, and to let users enrich the collection by subject indexing and by their own research findings. The texts are the scattered remains of the archives of the Swedish East India Company (1732 -1813). These documents, which are in great demand by academic users as well as by the general public, are being digitised as high quality images. Instead of text encoding, indexing is used to provide searchable entries for words in document titles, names of people, and subjects.

Electronic access

The Emerald Research Register for this journal is available at

<http://www.emeraldinsight.com/researchregister>

The current issue and full text archive of this journal is available at

<http://www.emeraldinsight.com/1468-4527.htm>

About the company

As had earlier been the case in other European nations, Swedish merchants in the eighteenth century wanted to profit from East India trade. A Swedish East India Company was founded in 1731. It was given a Royal Charter and worked as a monopoly enterprise until 1813. In the first decades, its headquarters were not in Stockholm, the Swedish capital, but in Göteborg, the harbour city on the west coast. Co-founder of the company was the Scotsman Colin Campbell, who had the necessary commercial experience from earlier employment in the English East India Company. Many other Scots and Englishmen were active in the company in the years that followed.

Trade soon concentrated on Canton in China, and accordingly the company was, in fact, rather a Swedish China Company. It was a purely commercial undertaking; there was no element of colonisation. The sailing routes normally went north of Scotland and from there straight to Cadiz, where silver was bought, Spanish silver being the only payment accepted by the Canton merchants. From Cadiz ships sailed to the Cape of Good Hope, and from there to the Canton harbour town of Wampoa. Tea was the main commodity bought, but also porcelain, silk, spices etc. Journeys started in winter and lasted about 16 to 18 months, to make it possible to return home in the summer. The cargo was sold at auctions in the company's warehouse in Göteborg. A large part was sold to foreign buyers and so re-exported from Sweden. In all, 127 journeys were made during the approximately 50 active years. Towards the end of the eighteenth century, profits diminished and the company finally became insolvent. It was dissolved during the Napoleonic wars.

The company as a field of research

There has always been a keen interest in the company's activities among students of Swedish economic, social and cultural history, genealogy, navigation history, etc. Historians debate its impact on the Swedish economy. It has been regarded as one of the most profitable Swedish business enterprises ever.

Refereed article received 20 April 2003

Approved for publication 9 May 2003



Recently, however, it has been argued to have been of more importance locally and for individuals. The company always destroyed the account books after closing them, and the reason for this, it is thought, was to conceal the enormous size of profits.

There is greater agreement on the importance of the company for the development of shipbuilding, navigation, overseas trade, science, and general culture. Some of Linnaeus' disciples travelled as ship chaplains or surgeons. Carolus Linnaeus (1707-1778) was a Swedish naturalist, botanist, professor at Uppsala University. He framed the principles for defining genera and species and created the binomial naming system. Many of Linnaeus' students were successful scientists and explorers. There are journals and diaries left by some of them, e.g. Pehr Osbeck, 1723-1805, a naturalist and clergyman.

A few English language studies of the company exist. Some are given in the reference list.

The company archive

There is little left of archival records. Many of the remnants are in Göteborg University Library (GUL), but there are also documents in some ten other institutions. Those in Göteborg are the City Museum, the Maritime Museum, and the Regional Archives. In Stockholm the institutions are the Royal Library, the Library of the Royal Swedish Academy of Sciences, the National Museum of Cultural History, the Maritime Museum, and the National Archives. The University Libraries of Uppsala and Lund also have relevant documents in their collections. For a list of archive sources, see Koninckx (1980, pp. 506-513). Surviving documents include staff rolls, journals, log books, diaries, letters, invoices, and inventories. Many of the surviving documents seem to have been kept as personal property by people involved in the company in some capacity. In the case of GUL, most of the documents have come into our possession through donations in the 1870s. We know of at least one document still owned privately by descendants of company travellers. In 1986, an important item was bought and placed on deposit in GUL by the Göteborg Historical Museum (now City Museum), viz. Colin

Campbell's diary of the first voyage in 1732-33. The manuscript was later published (Campbell, 1996).

Why digitisation?

The surviving documents are in heavy demand all over Sweden by researchers, university students, and members of the general public with an interest in history. They are scattered over several institutions in different cities. They are in danger of wear and tear.

Therefore, in 2001 a digitisation project started as a joint venture of Göteborg University Library and the History Department of Göteborg University. The project is financed by a grant from the Bank of Sweden Tercentenary Foundation for a three-year period. We are now in the last year of work. The Virtual Archive of the Swedish East India Company is available on the Web (www.ub.gu.se/samlingar/handskrift/ostindie/).

The project has three aims:

- (1) to give easy and free Internet access to these heavily used documents, with an image quality good enough for the images to be used instead of the originals and so spare those from damage;
- (2) to collect a virtual archive of documents scattered in different institutions; and
- (3) to let users of the documents enrich the archive with their research findings and do subject indexing.

Giving access

The language of the documents is almost exclusively Swedish. They cannot be used without knowledge of the Swedish language. Therefore, we have not written a bilingual interface to the database. Below, some Swedish words that occur on the web site are given in round brackets and quotation marks after English translations. The translated words are used in search entry headings and as index terms.

The project is confined to manuscripts, leaving printed sources aside. The manuscripts are in ink handwriting, mostly clear and well legible, but sometimes faded or stained, and sometimes the ink has bled through to the back of the paper.

Most of the material in GUL are bound volumes. Therefore, we chose to photograph

the pages from above with a digital camera already acquired by the library, which scans the pages. The images are saved as 300 dpi tif files, of an average size of 30 mB. The tif files are kept as archive images on a server at GUL. For publishing on the Web, copies are made, using the format DjVu by Lizard Tech. Users download a free plug-in for viewing the images and for printing. The format allows magnifying up to $\times 1,200$, and zooming in on selected portions. Transferring the images is a swift process. This is owing to the fact that the programme does not send the whole document in one package, but just the image/page asked for, plus a few preceding and following pages. These are seen as thumbnails in a frame to the left of the image. The average size of files sent is 250 kB.

The format eases the administrative work of keeping the pages of multi-page documents together in one series of files. For the user, browsing the pages is quick and easy.

Up till now (April, 2003) we have been processing only the documents owned by GUL. Today, 98 documents and 6,000 images are displayed, and some more are waiting to be published. Approximately 4,000 pages remain to be photographed. We hope to continue processing documents in other institutions by the end of spring.

Searching the database

Searches can be conducted in three ways: by words in the titles of documents, by persons, and by subjects.

Searching documents (entry heading: "Sök dokument")

Some of the larger documents have proper titles, e.g. "Journal of the ship Adolph Friederich" ("Journal för skeppet Adolph Friederich"). In other cases, there is no title or title page. We then use the title made up in the archive listing of the manuscript department of the owner library.

If words in the "title" are entered, singly or in combination, they will find the document.

A listing will be made up of the titles of all the digitised documents. This list will constitute a virtual archive of the Swedish East India Company.

Searching for people (entry heading: "Sök personer")

In the material so far digitised, there are nine rolls or other staff lists. Five of them have been indexed, yielding 978 personal names of men (no women!). Roll entries typically consist of employment position, name, place of birth, place of residence, age, marital status, and wages.

In the database, postings have been made from the person entries. The first four elements mentioned can be searched. The fifth, statement of age, has been replaced by year of birth; this has been calculated from the statement of age in the year of enrolment. This might obviously be wrong by one year, since the sailor's birthday may be later in the year than the day of enrolment. The (uncertain) year of birth can still be used as a help to identify homonymous men occurring in other documents.

Place names are registered in the rolls at different levels. That is, place of birth is sometimes given as a district in Göteborg (e.g. Masthugget), sometimes as a town or parish (e.g. Skara), sometimes as a region or province (e.g. Finland, at this time a Swedish province). In the database, place names are listed at two levels in drop-down help lists. So, one will find the districts of Göteborg at the same level as the parishes of the provinces.

Searches may be made for single elements or for combinations. You can, for instance, search for helmsmen ("styrman") living in Stockholm, seamen ("matros") born in Finland, or everyone whose first name was Johan.

Searching for individuals has turned out to be popular with users. This is not surprising, since family genealogy is a very common research subject for the general public. There are study courses and genealogical societies with many members. They have abundant material for study in the parish registers that ministers had to keep from the seventeenth century. These registers have usually survived and are now widely available in public libraries thanks to the microfilms made by the Mormon Church. The information on individuals in the company documents is appreciated as a welcome complement to the church registers. A large number of people are eager to find out if an ancestor sailed with an East Indiaman. Researchers, on the other hand, also use data on individuals, e.g. for demographic and economic studies.

Searching subjects (entry heading: "Sök ämnen")

This is the field where the main cooperation between the university library and the history department takes place. A faculty member, Reine Rydén, began by indexing 53 documents, assigning 596 terms to them. This was done to test the index method described below.

Index terms are normally chosen from the wording of the documents. Nouns, generally in the singular, are used. Spelling and inflexion are changed to modern usage when necessary. Terms are not translated from Swedish. The method is more or less the same as when preparing an index to a printed book. It can also be described as a very selective free-text word indexing.

In the test period, we started by formulating more general, descriptive terms, but abandoned this method. Members of the history department advised us that we should keep as close to the text as possible, and avoid the danger of interpreting concepts according to modern usage. Using the actual words was also thought to be a help for readers to recognise the terms on the image/page. There, the terms obviously cannot be highlighted as they would be if it were a system that used transcription and encoding.

There is also a searchable field for comments. This is meant to be used for explanations of the indexing target. For example, if another term is used rather than the word in the text, the comment field can be used for a citation. It can also be used for a short description of the context, to spare the trouble of the reader's looking up the page/image; e.g. for the term "Death" ("Död"), a note of the circumstances.

The user has access to the index terms in a drop-down alphabetical list. Terms are also arranged on a second level in broad subject groups (at the moment 16). Subject groups include goods and food ("varor och proviant"), geographical names ("geografiska namn"), nautical terms ("skeppstermer"), events ("händelser"), life onboard ("livet ombord"), business ("ekonomi"), and scientific observations ("naturobservationer"). We thought this grouping of terms would help users to get an overview of them. For indexers it is a tool to find what terms are already used and so avoid coining synonyms. As indexing continues, the number of subject groups will probably grow.

The idea of building a thesaurus with hierarchic relations was rejected. We did not think the labour and cost of doing this would correspond to the utility for users. We do not know how much indexing will be done after the end of the project. Therefore, it seemed to be excessive. Also, since terms were chosen from the actual wording of the documents, it might be impossible to force them into the rigid structure of a thesaurus.

When searching for a term, the user is given the number of hits and is referred to the pages where the term is to be found. Clicking on the page number will display the image.

There is no option of combined searches. Since hits are to page/image, we did not think this relevant. The terms would probably not be adjacent or even in context with each other.

The idea is that undergraduates, as part of their examination papers, will do subject indexing under the supervision of a tutor. This is beginning in spring 2003. Three students have got passwords for access to the administrative interface, and they will register subject terms for chosen documents.

For subject searches to be worth while, there must be a reasonable number of hits. We hope that the undergraduate programme will get a few thousand indexings registered.

Post-project time

When the project is finished, students, researchers and volunteers will hopefully take over the building up of the index, if they find it a useful and interesting tool. People studying a special subject might do topic indexing. There is much to be found in the fields of meteorology, history of food-stuffs, medicine etc.

We hope that the future division of labour will be as follows. The library will keep the database available, complementing it with more digitised documents and developing the database structure. There will also be the task of administering access for indexers and managing the term list. On the other hand, users from the academic world and outside will be expected to enrich the resource by adding new indexings as well as references to their own results, comments, bibliographical notes, etc. If this vision can be realised, we think it will be a somewhat novel model for a database system and a very rewarding activity for library staff.

Staff

The project team consists of Reine Rydén from the history department of Göteborg University, and from Göteborg University Library: librarian Helena Havner; photographers Andras Banovits and Jonas Lundström; technicians Lennart Stark and Rolf Johansson; and myself as project leader. All take part in the project on a part-time basis.

References

- Campbell, C. (1996), *A Passage to China. Colin Campbell's Diary of the First Swedish East India Company Expedition to Canton, 1732-33*, Royal Society of Arts and Sciences, Göteborg.
- Koninckx, C. (1980), *The First and Second Charters of the Swedish East India Company (1731-1766): A Contribution to the Maritime, Economic and Social History of North-Western Europe in its Relationships with the Far East*, Van Ghemert, Kortrijk.