Annu Jauhiainen,

Linda: The Union Catalog for Finnish Academic and Research Libraries

from / *aus*:

Union Catalogs at the Crossroad

Edited by

Andrew Lass and Richard E. Quandt

pp. / *S*. 101-122

Impressum

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.ddb.de.

Bibliografische Information Der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.

This publication is also openly accessible at the publisher's website. Die Deutsche Bibliothek has archived the electronic publication, which is now permanently available on the archive server of Die Deutsche Bibliothek.

Diese Publikation ist auf der Verlagswebsite ebenfalls open access *verfügbar. Die Deutsche Bibliothek hat die Netzpublikation archiviert. Diese ist dauerhaft auf dem Archivserver Der Deutschen Bibliothek verfügbar.*

Available open access / *open acess* verfügbar:

Hamburg University Press / Hamburg University Press
http://hup.rrz.uni-hamburg.de
Die Deutsche Bibliothek archive server / Archivserver Der Deutschen Bibliothek
http://deposit.ddb.de/

ISBN 3-937816-08-9 (print)

Table of Contents

Part 6 South African Union Catalogs

Chapter 5
Linda: The Union Catalog for Finnish Academic and Research Libraries

Annu Jauhiainen

All Finnish academic libraries and a number of other research libraries have had the advantage of using a joint library management system for over a decade. A unified network called Linnea was created in the early 1990s, consisting of local installations and a common physical union catalog which were all connected by the powerful and reliable academic data transmission network FUNET. A new library system, Voyager, which replaced the VTLS system in 2001, added new features to the union catalog and makes both cataloging and localization easier and faster.

1 History

The academic libraries in Finland have a long history of cooperation in the field of cataloging and library automation. The basic policy has been to follow standards and adopt a joint approach. Since 1977, the libraries have used the FINMARC format and the LSP application purchased from the British Library for offline cataloging and production of printed and microform catalogs. Online databases were already built from these data in the early 1980s. The first union catalogs, one for serials and the other one for foreign monographs, were already created at that time.[1]

---

[1] Esko Häkli, *Off the Record 2. Articles and Papers* (Helsinki: Helsinki University Library, 2000: 98–99).

A new era started in the 1980s. The Ministry of Education funded a project to select a joint automated library system for all academic libraries. The selection process was handled by the Automation Unit of Finnish Research Libraries, a unit within the Ministry established in 1974, which was also responsible for LSP usage on behalf of the libraries. The contract with VTLS Inc was signed in April 1988.[2]

In 1993 the Automation unit with all its tasks and resources was moved to the National Library, where the Division of Database Services continues its work. The unit manages the Linnea network, functioning as a common agency for the academic libraries. In this capacity the National Library is also responsible for the new steps toward Linnea2, as the next generation network is called. The Division of Database Services also maintains the national and union catalog databases.

At the turn of the decade and in the early 1990s, VTLS was implemented in the library databases, one by one. In 1993, when all library databases were up and running, the next step in the Linnea network was to create an online union catalog using the VTLS software. Different options were evaluated. Some people strongly pushed the virtual union catalog option, for they saw the physical union catalog as a waste of money, the money that they would rather have used for local needs. The decision, however, was to go for the physical union catalog, which would be updated by and linked to the local databases. Due to the large number of databases and relative slowness of the FUNET network at that time, a virtual union catalog was not a feasible option. Another reason for establishing a physical union catalog was that the HP3000 servers hosted by libraries would have been heavily overloaded by additional queries generated by a virtual union catalog.

Before the data could be loaded into a union catalog, some customized software development was needed for the VTLS system. For example, a duplicate control algorithm was designed in Finland and subsequently implemented by VTLS. This code was later used in other VTLS-driven union catalog projects, e.g. in Spain and Poland. VTLS also developed

---

[2] Esko Häkli, "A Unified Automation System Using VTLS for Academic Libraries in Finland," *Program* 26/3 (July 1992): 239–248.

features that enabled the libraries to use the Linda union catalog database efficiently for copy cataloging purposes.

The cataloging process was as follows:

1. The record was first searched in the union catalog Linda.
2. If it was found in Linda, it was copied to the local database by entering a single command. In the local database it was possible to do some further editing, e.g. certain fields could be added to the record, etc.
3. If the record was not found in Linda, it was first cataloged there. From there it was copied to the local database, where it could be edited further.

Depending on the material, 50–90% of MARC records could be copied from Linda. Inter-library loan (ILL) localization was also very efficient, because Linda contains summary-level serials holdings from over 400 Finnish libraries. But in the old Linda there were no links between the union catalog and the local database for retrieval of up-to-date holdings and item information. The technology of the time did not make that possible. When you searched a title in Linda, you got the bibliographic record plus a list of libraries holding that title. There was no way of seeing how many copies the libraries had and what the status of the copies was. It was necessary to log onto the local database in order to see the status. The link between Linda and the local databases, which permitted easy copying, was available only in cataloging.

2 Linda and the Other Linnea Union Catalogs

The Linda database is the union catalog for the Finnish academic libraries. The numerous libraries of the 20 universities in Finland, along with the Library of Parliament, the National Repository Library and some special libraries, contribute their records to the database. The National Bibliography Fennica, complete from 1488 onwards, is also included in Linda. In addition Linda contains summary-level serials holdings from hundreds of special libraries and polytechnic colleges. Altogether there are over 460 libraries contributing their records to Linda in one way or another.

At the end of 2001, Linda contained 3.7 million bibliographic records, the annual growth being about 200,000 records. The database includes references

on monographs, serials with summary holdings information, cartographic materials, audiovisual materials, electronic resources, multimedia and archives.

Linda does not cover music materials. They are cataloged in Viola, which is the Finnish National Discography and National Bibliography of Sheet Music as well as the union catalog of music materials. Viola contains references to Finnish sheet music since 1977, and to sound recordings since 1901, that is, from the very beginning. Cataloging covers the whole sound recordings and scores as well as the individual compositions contained in them.

In addition to Viola, Linda has another sister database, Manda. Manda is the union catalog of 20 regional central public libraries in Finland. Manda contains references on books, music, visual materials, cartographic materials etc., but not serials, as information on serials holdings of these libraries, as well as many other public libraries, can be found in Linda.

3 Selection of the New Library System

Towards the latter part of the 1990s, it became evident that the VTLS system had to be replaced. VTLS had been a trustworthy companion of the libraries for a long time. The system had been a good and stable housekeeping tool, taking care of most of the traditional activities and functions of the libraries. However, that was no longer enough. Due to the rapid change of technology and the new needs in the library field, a new library system was needed, one that could respond to these new requirements and go beyond being a mere housekeeping tool. In answer to the demands of the market, all library system vendors were developing so-called third-generation library systems with relational databases and client/server technology, graphical user interface and Web gateways, the ability to search multiple databases simultaneously, multimedia support and support for internationally accepted standards such as Z39.50, Unicode and ISO ILL, to meet the growing needs of the users. It was also clear that the classic VTLS system was coming to the end of its life-cycle and would not be developed further, since VTLS, Inc. was concentrating on its new system, Virtua.

Because of the great success of Linnea1, as the old network is now called, there was no need to revise the basic service philosophy when moving to a new system. It was self-evident that we would continue with a joint system. Libraries were satisfied with the system and the workflows and cooperation with one another. However, the libraries were open to totally new technical and organizational solutions if these proved more advantageous both functionally and economically. And the National Library wanted to avoid transplanting old patterns into a totally new environment, and wanted to make full use of the advantages offered by the new technology. Thus, we wanted to explore different options for the future database or network architecture during the software evaluation process. One of the important issues was whether to merge existing databases or to keep the current structure. In the RFP, the vendors had been asked if their system could support other kinds of database solutions, i.e. a single central system with full functionality and no local systems, or a data warehouse-type central system of bibliographic data with local circulation systems and indexes. This was also discussed in detail in the negotiations with the final candidates.

Merging existing databases together was technically possible. Some of the vendors even encouraged it. In some cases, it would have meant a significant saving of money in the software price, as well as in the ongoing maintenance of the software and hardware and the overall maintenance of the system. On the other hand, it would have meant a difficult and time-consuming implementation, plus higher implementation costs. Most importantly, it would have meant losing all the work that had been done in the Linda database over the years, because the new centralized database would have to be created from the local databases, not Linda, since the necessary holdings and items information did not exist in Linda but only in the local databases. Besides, we were not convinced of the functionality or the technical merits of such an action, nor the security of the results. We also had to take into consideration the opinion of the participating libraries, which were quite reluctant to pass the maintenance and configuration of their database to a centralized agency that would not be so well acquainted with local customs and needs.

The conclusion of the discussions and the research in this area was that we would gain nothing by merging databases into one centralized system.

On the contrary, it would have made life more complicated, and thus it was decided to keep the same number of databases as before. The same result had been envisioned in the future scenario of the Linnea network that had been prepared at the Helsinki University Library in 1997. According to the scenario, the future network would be based on the Z39.50 and ISO ILL protocols. Use of these standards would allow patrons and staff to log on to different library systems, search remote databases in Finland and abroad seamlessly and retrieve records from them online. This would give new scope for the architecture of the network. According to the scenario, it is likely that the three bases of the network, the local services, the central system and the network connecting them (the Finnish Universities and Research Network FUNET) would remain the same, or almost the same, for the next few years.[3]

The future of the Manda database was reviewed during the selection process. The question was whether to migrate Manda to the new system as an independent database or to merge it into Linda. We also considered freezing Manda as it was, in which case new records would subsequently be added to Linda. As the result of research among Manda users and the feedback from the public library sector, it was decided to continue with Manda as an independent database and migrate it to the new system. We have to admit that we were worried about the quality of Manda records and about what the effect of such merger would be on the quality of Linda, since the Manda libraries use various management systems. These systems have not even always used the MARC format for cataloging, which has not been as standardized in that sector as in the academic libraries.

Due to the obvious benefits of the existing physical union catalog, the issue of a virtual union catalog versus a physical union catalog was not seriously considered. After abandoning the centralized database option, it was a natural choice to continue with the physical union catalog, but with the help of various virtual union catalogs, e.g. subject-based, regional, union catalog of union catalogs (Finnish national databases as well as the Scandinavian Virtual Union Catalog, SVUC).

---

[3] Information Technology Scenario for Finnish Libraries, 1997–2006. See http://www.lib.helsinki.fi/skenaario/skenarioe.html.

FINMARC has been the cataloging format of the academic libraries since the 1970s. It has been the basis for cooperation in cataloging. Now that the library system had changed, it was a good time to review the format issue and decide whether to continue with a national format, or to harmonize and go towards a global solution. We saw the advantages of a global option in copy cataloging and in the exchange of records. On the other hand, FINMARC had advantages that we were not willing to give up, most important of which were the ISBD punctuation and field 248. The result of the evaluation was to move towards MARC21 but to keep some of the local features. The new format is a hybrid of MARC21 and is called MARC21-Fin.

The software selection process was arranged according to the European Union rules of procurement. During the final phase, we carried out an extremely thorough evaluation, with system demonstrations, hands-on testing, site visits and reference research, negotiations with the developers of the systems and financial evaluations of the vendors. The goal was to find the most functionally suitable and the most economically advantageous system for the local databases as well as the union and national databases. The essential guideline in the selection process was a fair and objective treatment of all parties involved. Since every step was documented, we would have been able to reconstruct the process, should it have proved necessary.

When the different parts of the selection process were drawn together, Voyager, produced by Endeavor Information Systems, Inc., best fulfilled the criteria. Voyager was found to be a complete, integrated system that was finished in the essential, traditional functions needed by the libraries, but which however is being further developed to meet the new needs and changing technologies. It fits both individual Linnea libraries and the Linnea network well. Local services can be streamlined and their scope extended. But centralized services will also benefit from Voyager via its consortium-driven functions. Increased efficiency is largely based on improved networking, since Voyager supports both Z39.50 and ISO ILL.

Special attention was paid to the union catalog functionality of the four final candidates. The new-generation software was seen to offer several enhancements to a union catalog compared to the old one. For the catalogers, it is easier and faster: the union catalog is updated automatically, as the

system copies new records from the local databases according to the configurations of these databases. For the users it is more informative, since there are real-time links from the union catalog to the local databases, displaying the status of each item. With the help of another Voyager function, Universal Borrowing, the user will also be able to place a request on the item.

The selection process was coordinated by the National Library, but all the libraries were heavily involved in the process from the beginning, when the selection criteria and the RFP were compiled, through the evaluation and testing of the systems until the end, when the decision was made. The directors of the libraries made the final decision by unanimously accepting the proposal made by the National Library. Voyager was selected as the new system for the Linnea network, and the contract with Endeavor Systems, Inc. was signed on February 4, 2000,[4] after the rectors of the universities had also approved the decision.

4 The Network Architecture

The next question was how many servers an optimal solution for the Linnea2 network would require. In the Linnea1 network there were 17 HP3000 servers for the 25 databases. The number of servers was never really discussed during the implementation of Linnea1 because of the limitations of the computer technology of the time. Times were different now, and the consortium license signed with Endeavor enabled the libraries to have any number of databases or server machines. Accordingly, we had a free hand to pick the best network architecture for the Linnea libraries.

How far can one go in centralization? The answer depends on three factors: the available data transmission network, the capabilities of the software and the state of the computer technology.

The Finnish Academic and Research Network, FUNET, has been a key factor in the Linnea network since the beginning. Without the reliable infrastructure

[4] Annu Jauhiainen, "A New Library System for Finnish Research Libraries Chosen," *Helsinki University Library Bulletin* 2000: 12–19.

provided by FUNET it would not have been possible to use Linda as a cataloging tool in the manner we have since the early 90's. FUNET network allows libraries from all parts of Finland to efficiently access Linda and other union catalogs located in Helsinki. During the last two years the network has not been down even once. Given the extremely robust architecture of the network and reliable maintenance organization (Center for Scientific Computing, CSC), there are good reasons to believe that the FUNET network will remain at least as reliable and efficient in the future as it is now.

A shared server is not feasible for a library consortium if there can only be one database on the server. The Voyager software allows in principle an unlimited number of databases on a single server. However, practical experience from other Voyager consortia made it clear that there should not be more than about 5-7 databases on a single server, since a large number of databases may require much time for Oracle and Voyager updates: it may take several days to update many large databases, and during the process all the databases must be closed. Fortunately this problem has disappeared in subsequent Voyager releases; it is now possible to update databases on a shared server one at a time.

However, if all databases are dependent on the same database application or hardware and operating system process, severe problems would have an impact on every library simultaneously. Fortunately, new server technologies make it possible to have a single server and still avoid this problem: there are servers that can be internally split into several logical (and physical) parts.

Both Sun and IBM, which were the platforms Voyager supports and therefore the only possible candidates for the Linnea2 hardware solution, can deliver cluster-like computers. The high-end models of both the IBM and Sun product family can be separated into logical parts called domains (Sun) or nodes (IBM). Each part has its own operating system process and dedicated hardware from network card to processors. To the operators and users, the server looks like a cluster of computers.

There were, consequently, no technical constraints on choosing the network architecture freely. The National Library was eager to find out whether centralization would save money. The idea was not fully accepted by all at first, for a few computing centers were reluctant to give up

maintaining their own server. Therefore, at the request of the universities three scenarios were analysed:

- centralized model; all databases placed on a single machine;
- semi-centralized model; 3–5 servers;
- decentralized model; the current number of servers.

Cost analysis was based on both purchase price and the total cost of ownership, calculated for five years.

After a thorough analysis of the various options, a decision was made to choose the centralized architecture and buy Sun E10000 as the server system. The decision to go for Sun was based on technical merit and price. Both Endeavor and Oracle use Sun machines as their development platforms; this fact was also taken into account. Large computers such as the Sun E10000 have been optimized for heavy duty database usage and are also very reliable. Our practical experiences have shown that E10000 is indeed a very reliable server. Application-level problems in Oracle or Voyager are far more common than server problems, although still rare.

The Linnea2 server is able to handle 1400 active users, or more than 5000 concurrent users, about twice as many as before, on 17 HP3000 servers running VTLS. Both Endeavor, which did the calculations for the hardware configurations, and we felt that an ample safety margin was needed in order to avoid performance problems.

Immediately after the server was chosen, the decision was made to outsource the maintenance of the new server to the Center for Scientific Computing (CSC), a non-profit company owned by the Ministry of Education. It hosts Finnish supercomputers and maintains the FUNET network. CSC staff have excellent UNIX and networking skills, and are therefore very well qualified to maintain the E10000.

We have good evidence for the claim that an unprejudiced approach to server architecture has enabled us to combine significant savings with important technical improvements. Being a consortium helps a lot: libraries buying systems only for themselves will not be able to utilize new technology with similar efficiency. It is easy to understand from this point of view why library consortia are becoming more common in the US and some European countries. Finland has been one of the pioneering countries

in this area, and our experiences from such cooperation are very encouraging.

Aspects of Centralization and Decentralization

Analysis of Sun and IBM hardware and discussions with technical experts led us to some generic conclusions:

- There is a general trend towards centralization, which started in the mid-90s, in commercial companies that are more aware of costs than public institutions. Universities have been slow in reversing their current tendency to decentralize, since the purchase price of small servers is approaching zero. However, the ever-growing number of computers means that operating costs are growing fast. Badly managed UNIX servers have already caused security-related and other problems, and things may get worse if decentralization continues;
- Hardware vendors are reacting to centralization (server consolidation) by developing systems that make it easy to consolidate applications from a large number of existing servers into a much smaller number of large computers via 'internal clustering.' The Sun servers such as 4800, 6800, 10000, 12000 and 15000 and IBM RS/60000 SP are good examples of this trend. In the future we will see even more systems of this kind from Sun, IBM and other vendors. Naturally these machines will be substantially faster than current ones; another prerequisite for centralization;
- Hardware vendors are capable of, and willing to, offer bargain prices for large systems. For workstations and small servers, proportionate discounts will always be much smaller than for large systems. If list prices are used for estimating purchase costs, centralized solutions may seem to be expensive. However, if negotiations are successful, a centralized server may well become the cheapest choice; and
- Never forget to estimate the total cost of ownership. Buying a number of small computers may look like a bargain, but taking all costs into account may change the picture. There are a number of things to remember: maintenance costs (paid to the hardware vendor), license and support costs (to the software providers), operating costs, plus miscellaneous

costs such as floor space occupied by the system and the electricity consumed by it.

5 The Linnea2 Consortium

During Linnea1, the cooperation between the libraries was never formalized. Collaboration was based on mutual understanding, with the National Library as the central agency, giving guidelines and working as an intermediary with the library system vendor, and with the Ministry of Education as the financer of the acquisition of the system and of the implementation. In the Linnea2 project there was no central funding from the Ministry, as had previously been the case; instead, the universities had to find the money from their general budgets. In addition to having a single contract with the software vendor, the members of the Linnea2 Consortium became owners of hardware that they had to administer jointly. It was considered necessary to have a formal contract and bylaws to ensure that decisions, especially concerning money, were handled in a way all members had agreed upon.

After the software, hardware and hardware maintenance contracts had been signed, it was time to legally establish the Linnea2 Consortium. The twenty universities, the Library of Parliament and the National Repository Library are the founding members of the consortium. New institutions can join as associate members that can buy services from the Consortium and from the National Library. According to the bylaws, most decisions, especially those dealing with money, have to be approved by the General Council, based on consensus. The Steering Group consists of seven members. The National Library is the executive body, preparing all the matters for the Steering Group and the General Council and representing the Consortium in dealings with third parties such as the software and hardware vendors (Endeavor Information Systems Inc and Sun Microsystems Finland), the hardware maintenance organization (the Center for Scientific Computing – CSC) and the outside world in general, for example the media. The Library is also responsible for organizing and coordinating cooperation and communication within the Linnea network.

The Linda database is owned by the National Library. The Consortium is not legally or organizationally involved with Linda. However, the

Consortium libraries are the main contributors to Linda, as well as the owners of the shared hardware and the software license. Therefore the National Library feels that it is important to discuss matters concerning Linda openly with the Consortium and have its acceptance in major issues.

6 Implementing Voyager

The implementation of Voyager in the Linnea network took place in the summer of 2001. The process started in April, and all local databases were using the new system by the beginning of the academic year. The implementation in the local library databases was smooth, considering how complex the situation was with so many databases and so many parallel loads. Including test loads, altogether about 35 million bibliographic records were converted from one character set to another, one cataloging format to another and one library system to another. In addition to the bibliographic data, acquisitions and circulation data were also migrated. This required very careful planning, taking into account human resources in the libraries, at the National Library, at the server maintenance organization and at Endeavor. Furthermore, everything was also dependent on the hardware resources. Fortunately we had a powerful server, which is divided into five logical parts, each of which could be used effectively for simultaneous loads.

The biggest challenges in the implementation were the size of the conversions (15 million bibliographic records and 26 databases), the tight schedules, the different conversions, multilingualism and different character sets (Cyrillic and Scandinavian characters), and communication among all parties. Thanks to sharing a physical union catalog, the data of the libraries were relatively homogeneous, which helped the conversion process.

The schedules were made tight on purpose. When Linnea1 was built, it took several years to implement VTLS in all databases. That was possible at the time, because most functions were manual and could continue that way as long as the implementation was finished and the new system was ready to be used. The situation was quite different now. The changeover to the new system had to be planned carefully in each database to make sure

that the functionality and the services in the libraries could continue seamlessly. The main reason for the tight schedule was, however, the union catalog. The libraries were dependant on Linda for copy cataloging and ILL localization. We could not afford to cut that tie for a very long time. Therefore the strategy was to migrate the library databases first and then the union catalog immediately after that. However, this plan did not quite work out as expected.

Implementing Voyager in the Linda database was not as easy as in the local databases. The reason for this was the fact that Voyager Universal Catalog was planned for consortia which had not had a union catalog before, but the catalog was created from the participating local databases at the same time as the data were migrated from the previous system to Voyager. The dynamic links between the Universal Catalog and the local databases were created during the load. However, in our case the union catalog already existed: we had Linda, a union catalog that was in very good shape. There had been a lot of duplicate records as the result of the initial loads in the early 1990s, in spite of our sophisticated duplicate detection algorithm. Those duplicates had been cleaned up little by little, and by the time we were ready to start the Voyager implementation, all duplicates had been taken care of. We did not want to lose all the work that had been done over the years and start from the beginning again. Endeavor was willing to do some development for us to enable Linda to be migrated and the dynamic links to be built differently from other UC sites. This development work was, however, more complex and more time-consuming than Endeavor had anticipated, which caused unfortunate delays in the implementation.

As of the fall of 2002, the implementation is still not complete. Endeavor has finished the initial loads, but we are still loading to Linda the material that has been cataloged into the local databases since the VTLS system was closed down in the summer of 2001. We are about to start ongoing life with the UC, which means constant real-time updating from the local databases.

In a large project like this, with a great number of libraries involved, communication is vital to a successful outcome. Communication between the libraries and the vendor had to be organized and coordinated. Communication and cooperation among libraries was equally essential. In

Linnea2, there was one new partner in the communication triangle. The arrangement of outsourcing the server was a new challenge to all partners. CSC had not worked with library databases before. The fact that the operation of all academic libraries is dependent on the server being up and running continually from early morning till late at night has required changes in their thinking and daily routines. For Endeavor, our solution is novel as well, in spite of their large number of customers worldwide. There are some centralized Voyager systems, but not to this extent. The maintenance organization being separate from the libraries or universities was also unknown to them. The change has, however, been most significant for the libraries. Until now, all except two of them have had their own server, maintained by their own people or by the computing center of their own university. By the time the common server was chosen, the libraries were more than willing to give up the maintenance of their own hardware. The long implementation period gave us a good opportunity to learn what living with a shared server really means. Each library always has to remember that in every respect they are not on their own, but must take their fellow libraries in the same E10000 domain into account. One configuration error, such as too long a timeout period, may cause problems in all libraries sharing the same domain, in spite of the safety margin in the server resources. We have unfortunately had some problems, but these occasions have taught us valuable lessons, and all parties should now be aware of how to avoid such incidents in the future. However, the fact that libraries now avoid the trouble and cost of maintaining their own UNIX servers is a significant improvement. A big help in the new situation was however, the strong tradition of cooperation in the library system area for more than a decade. [5]

[5] Annu Jauhiainen, "Voyager Implementation in the Linnea Network", *Helsinki University Library Bulletin* 2001: 10–13.

7 How the UC Linda Works

The Voyager Universal Catalog (UC) is a physical union catalog with real-time links to holdings and item information from the contributing libraries. Bibliographic records are the core of the database. Each bibliographic record has an attached holdings record, or several of them, indicating which local library database is holding the title. If the same bibliographic title belongs to several databases, the same number of holdings records is attached to the bibliographic record.

The records in the Universal Catalog are deduped. The deduplication process occurs when records are loaded into the UC, based on the duplicate detection profile, which is up to the library to establish. Voyager's duplicate detection algorithm does not fulfill the needs we have in Linda. We need to be able to separate almost identical records where only e.g. record types, languages, etc. differ, but that is presently not possible. The basic philosophy of duplicate control of this system needs to be changed in order to make that possible. Neither does the merge function in bibliographic duplicate detection work as a proper merge function should. This feature will be enhanced in the near future.

The holdings records are generated and attached to the bibliographic records when bibliographic records are loaded into the UC database. The 014a field of a holdings record contains the identification, which links it to the associated bibliographic record in the local database. The 852b field indicates to which local library database the record belongs. The UC holdings record only functions as a pointer or stub record in the dynamic connection to the local libraries' databases. As a search result, detailed holdings and item information is retrieved in real time from the holdings and item records stored in the local libraries' databases.

During Linnea1, the catalogers were actually working in Linda. They cataloged everything directly to the union catalog and then copied the records to their own local database. Now the workflow is the opposite. Nothing is supposed to be done directly in the universal catalog. Records are cataloged (or in most cases copied from Linda or from some other bibliographic utility) into the local database. The system takes care of the rest. The cataloger need not know anything else; all that has been taken

care of by the system administrator, who has set up the necessary configurations.

There are several configurations that must be set on the system side before any records can be loaded to the UC. The settings include definition of each local library that the UC server connects to for detailed holdings and item information, duplicate detection profile, bulk import rules, cataloging policy definitions and security setups.

Dynamic retrieval and display of holdings and item information requires certain configurations on the local library side as well, in order for servers to connect to each other. Database definitions and connection information have to be set up in each contributing library database. In addition, there are some policy issues that need to be discussed, e.g. decisions have to be made whether to exclude certain records from the UC load. For example, such records might be acquisition records for titles that have not been received yet.

Once the configurations are set on both sides—the Universal Catalog and the contributing databases—every change in any of the local databases is updated in Linda. Records can be added, deleted or modified, and the change is reflected in Linda. The ongoing updates are bulk-loaded to the UC on the basis of the schedule set in the configurations. The bulk load schedule has to be defined separately for each database. The loads can be carried out every ten minutes or once a day, or even once a week, or at any interval in between.

8 Universal Borrowing

Voyager's Universal Borrowing (UB) function provides a structure for unmediated, reciprocal borrowing in a Universal Catalog setting. It allows the libraries to use their collections in integrated circulation and share the patron data. According to its basic philosophy, UB is patron-initiated and unmediated. Patrons of participating libraries can request and borrow material from any library within the Consortium. The material can also be returned to any library. All transactions are tracked in real time and patrons can follow the status of their requests, loans and possible fines and fees through the Web interface.

The use of Universal Borrowing requires a fair amount of technical work, in other words, a lot of configurations in each participating database. However, the technical part is easy, in spite of all the work. The technology allows almost anything, as long as you have taken care of the necessary settings. It is the politics that is the hard part. A lot of political decisions have to be made in order to get a sensible and usable functionality. That naturally takes time and requires agreements among the participating libraries.

The Linnea libraries have in principle decided to implement Universal Borrowing. At present a few libraries are starting to test it, in order to see how it fits our workflows and customs. The general trend within the Consortium is to encourage resource sharing and to help the users to get the books they need as fast and as cheaply as possible, even if that will most likely change the guidelines used within inter-library lending. One strict rule has been that users are not allowed to order from elsewhere a book that is held by their home library or any other library within the same city. This will inevitably change because the system does not yet offer a way to check the local holdings before the request is sent to another library.

Simultaneously with the testing period, we are supposed to agree on the political issues. First, there has to be an agreement on which libraries will participate in reciprocal borrowing. Is it going to be all libraries together, so that requests may be sent to any library in the Consortium? Or is the National Repository Library going to be a unilateral companion to each library, in its role as the repository for all of them? Each participating library will have to decide whether it wants to exclude certain collections from this function, preventing access by other libraries' patrons. Each library also has to decide whether it is going to allow requesting and borrowing to all of its patrons, or only to certain patron groups. Libraries together have to agree on the blocking of patrons (when and for what reasons) as well as on fines and fees. They have to decide whether they want to collect overdue fines or any other fees, and how the fines and fees are handled. Sending books from one library to another means costs, as Finnish universities no longer have mailing service free of charge. Since requesting is unmediated, the result will at least initially be a lot of books mailed from one place to another. A lot of books will be requested and never picked up for loan. Who will pay the mailing cost when a book is

sent back to the library where it belongs? The only solution seems to be to make students pay the mailing costs. It is also anticipated that books will at times be returned to 'wrong' libraries, even when it is not a universal borrowing loan to begin with. It is simply handy for a traveling student to return a book to the nearest library. Who will pay for the mailing of those books?

So there are a lot of open issues to be solved before this functionality is ready for use in the Linnea network. However, it is a marvelous way to encourage resource sharing in the tight economic situation.

9 Linda and the Polytechnics

The Finnish polytechnic libraries are at present in the process of implementing Voyager. There will be 28 Voyager databases after the implementation is over by the end of 2003. The polytechnic libraries have been using various systems and have until now not cooperated in the library system field. Nor have they had a union catalog of their own. Their serials holdings are included in Linda, but not their monographs. Now, as their implementation is moving forward, they are facing the union catalog question.

The polytechnics have three options at least in theory: to use a virtual union catalog, to have a physical union catalog of their own, or to join Linda.

The virtual union catalog is a suitable interim solution during the implementation phase when there are only a few Voyager libraries among the polytechnics. Once all 28 databases are up and running, the load on the server would be too high. The polytechnic libraries followed the example given by the Linnea libraries and purchased a shared server for all of their databases. The server is configured for the 28 databases only, and simultaneous search on all of them would be too much for it to handle. The number of the databases would also cause difficulties in duplicate detection when, at the maximum, records from 28 databases were displayed.

A separate union catalog for the polytechnic libraries only is a noteworthy option that has to be considered seriously. Creating such a union catalog would be relatively easy. However, the main problem with

this option is the cost: it would be necessary to purchase a new server, since the shared server the polytechnics libraries now have would not be able to cope with the union catalog database. The libraries should also buy a Voyager UC license and establish a maintenance organization for their union catalog.

From the point of view of costs, adding the polytechnics libraries to Linda is an attractive choice. There are also obvious functional benefits. It is estimated that the polytechnic libraries have a relatively small number of titles that are not yet in the Linda database. So the number of bibliographic records would not grow much if the data from the polytechnics were loaded to Linda, whereas the number of stub holdings records would be comparative higher. The use of the database would not be affected significantly either, since the polytechnics are already using Linda for searching as well as copy cataloging. The centralized server of the Linnea consortium has the resources to accommodate the growth in the number of records and also the increased use. Besides, it is possible to expand the server by adding CPU and memory, should that be necessary.

If the polytechnics' data were added to Linda, the number of libraries contributing to the database would be more than double what it is now. Furthermore, the new libraries do not have the same experience of collaboration as the present Linnea libraries, and they do not share the same practice in cataloging, nor the same level of standardization. That would mean an increased need for support. The Database Services within the National Library, the former Automation Unit of Finnish Research Libraries, is maintaining Linda and supporting the contributing libraries. The unit would have to be strengthened with new resources. However, that would be an easier and cheaper option compared to establishing a completely new support unit, even in the case of a separate union catalog for the Polytechnics.

These three options for the union catalog are under discussion at present. It is expected that decisions will be reached at the beginning of next year.

10 The Portal Project

The National Library has started a project for procuring software for the National Portal and Digital Library. There are two separate procurements, one for the Portal and the other for the Digital Object Management Software. According to our vision, the national network will in the future consist of three modules: Integrated Library System (Voyager), the Portal software (application to be chosen) and the Digital Objects Management System (application to be chosen). These three applications will have to communicate and work seamlessly together, as well as with other applications, via APIs and using open standards, to the extent that the patrons will see a single service.

According to a definition established at the workshop "Portals: Is There a Role for Libraries?" at ELAG, the European Library Automation Group, Semantic Web and Libraries, Rome, 17-19 April 2002:

> A LIBRARY portal is an application which allows one-stop-shop access/searching and discovery via a unified single-point interface to organized heterogeneous resources and enabling services to a pre-defined community (users).[6]

In the Finnish Academic Network, we see the portal as a gateway to the library databases, the union catalog Linda and other national databases, electronic resources, and collections, as well as remote databases which may be open to anyone, or commercial databases licensed by FinELib, the National Electronic Library. As of 2002, FinELib licenses cover about 120 databases and approximately 8200 scientific journals. With the help of the portal, Linda will be part of a huge virtual union catalog that connects all databases the user wants to include in the search.

The portal software must enable efficient searching of remote databases via Z39.50 or other means; it must be possible to exchange patron data between applications using the NCIP protocol and/or application dependent APIs, and all systems must support OpenURL for context-sensitive linking. OpenURL will have a direct impact on cataloging into Linda, for it is expected to solve the difficulties in maintaining the URLs of the electronic

[6] See http://www.ifnet.it/elag2002/workshop.html.

journals that are cataloged to Linda. But this is only one issue in the field of electronic material. Discussions on how to handle all electronic resources in Linda have only just started.

The procurement for the portal software is at the final stage. The decision will be made at the end of October. The plan is to implement it in a very short timeframe and be in production at the beginning of next year.

11 Conclusion

The Linnea libraries have been using the physical union catalog for nearly ten years. We have strong evidence of its advantages. We did not want to have a separate union catalog, the maintenance of which would require extra work. The aim has been, since the very beginning, to have a union catalog that is integrated into the local catalogs, in order to save resources in cataloging and to ensure homogeneity of data. The aim was already reached during Linnea1 and the first joint system. Linnea2, and the new generation system gave us a union catalog that is linked to the local databases in detailed holdings information and offers its users a lot of functional advantages. The next step will be a union catalog that will be a portal to the entire library network, and the basis for new services.