



The integration of research information sources at university level

Cecilie Ohm, Johanne Revheim, and Jostein H. Hauge

University of Bergen, Norway

Preface

At the University of Bergen, a new system for documenting current research (CRIS) is presently being developed, primarily for administrative and strategic planning purposes. There are several factors motivating the new system development, one major factor being the present revision by our national authorities of current university budgeting objectives and strategies

The new system is a response to a number of needs for extensive information sources on both university and national levels. The system has also been developed with a view to offering active support to the researchers who are meant to be the main data providers at the departments of our university.

Technically, the system makes use of an SQL server with a web interface based on Active Server Pages (ASP).

When building our new CRIS system, we have tried to embed in the system a series of success factors. Accordingly, we believe that some of the most important aspects of the system are mechanisms for minimising as much as possible the registration efforts of the research community, e.g. by providing a series of data entry support tools. Moreover, the ultimate success of the system is dependent on the possibilities of creating flexible and diverse output products in the system's different fields of usage and in various user-defined formats.

1 Introduction and background

At the University of Bergen, a new system for documenting current research (CRIS) is presently being developed as a primary tool in administrative and strategic planning and for a variety of reporting purposes.

There are several motivating factors behind the new system development. Most importantly, our effort is an answer to the present revision of university budgeting objectives and strategies by the Ministry. As a function of this, it has turned out that there is a great requirement for documentation systems that can cover the documentation needs of internal administration and form the basis for the execution of externally defined data services.

One key application area is that of annual reporting of results of research at the university. The contents of these reporting tasks is defined both by our own administration and by national authorities. Over the last years, documentation of research output has become ever more important, and the data produced is being actively utilised by departments and





universities alike as a key tool e.g. to form a basis for budget proposals. In addition, introduction of standardised annual reporting procedures offer the possibility of making effective inter-university comparisons.

Today's annual reports are dependent on data collected from several different registers and database systems at the university. In order to fulfil some of the reporting requirements, data have traditionally been extracted from systems primarily designed for other purposes. Data from different sources will in some cases have to be integrated into a common data set by traditional manual procedures. Given the future importance of research data, this state of affairs is highly unsatisfactory, and brings about a demand for a new research documentation support system. We have, therefore, set ourselves the task of designing and developing a system where different kinds of data pertaining to research will be encompassed in one common documentation system framework.

1.1 Current situation

Firstly, a few words about the current situation at the University of Bergen regarding systems for keeping track of our student body, research activities and production output.

Student-based activities are presently registered in a database system developed jointly by our four universities. Registration includes data about every student and on all levels of study, including post-graduate students. The system also holds data on exams, courses taken, and more. Aggregated data from this database are regularly being exported to a national system for student data, run by our authorities. This database, which was introduced in 1996, is becoming a widely used and powerful tool for statistical information, e.g. in the strategic planning and budgeting of educational resources at universities and colleges of higher education.

We also have a modern database system for accounting purposes, and a new database for personnel information is currently in the pipeline. Both systems are being developed as co-operative projects between three of the Norwegian universities.

Up till now, research-based activities have been held in a national system that has quite a wide user participation at departmental level. However, after having collected and worked on data from this system since 1995, our experience to date has been that not all parts of the system are being actively used and that data turn out to be incomplete and also scarce in quite a few important areas of coverage. The consequence has been, that we traditionally have had to put a considerable amount of intellectual effort into validating and supplementing our data. This is a time-consuming way of operation since it also presupposes direct contact between our central administration staff and more than one hundred contact points at departmental level.

2 Creating a new database

Realising the need for co-ordinating the data production for reporting on research-based activities and output, we became highly motivated to try and develop a new system for the integration of multiple research information data types. Our ambition has not the least been



to create a user-friendly system where data can easily be input by the researchers themselves, and also by department administrators. The latter may also be given the authority to validate the data produced by the primary respondents, i.e. the researchers. However, as it has turned out, our strategic objectives have entailed that we have had to considerably extend the scope of our CRIS system, compared with the more conventional database types in this field.

On the other hand, we have at the same time tried, wherever possible, to adhere to the recently updated European documentation standard for current research activities and results, i.e. the so-called CERIF standard. It is our hope, moreover, that our database system may also stimulate parallel system development at other Norwegian universities or - alternatively - constitute the starting point for a joint development programme based on the current version of our system.

Technically, the system makes use of an SQL server with a web interface based on Active Server Pages (ASP). The system operates in such a way that the users are guided through the different documentation parts of relevance to each individual researcher.

2.1 The components of the database

The integrated system comprise the following main components:

- Documentation of on-going research projects
- Documentation of different forms of results of research activity
- A Ph.D. abstracting and press release service
- A module for documentation of expertise
- A module for web-based annual reporting from primary to top level in the university
- A series of documentation products and data export formats
- A module that is part of a full-text self publishing support service at the university for master and Ph.D. students, and researchers

In the following, we will give a brief description of some of these components.

Documentation of on-going research projects

In the new database, the input is not only research results, but also on-going activities. Documentation of on-going research projects produces important strategic information, and can e.g. tell us much about the degree of external financing of the research at the university and about collaboration among researchers. In some instances, this database provides the only possibility of having a complete overview of all the researchers who carry out their activities at the University of Bergen.

The documentation of on-going research projects does not necessarily have to be performed by the researchers themselves. The university administration, as a service organisation, has a quite good general overview of the different research projects that are financed by external means. On the other hand, in many cases such information can only be found



at departmental level. When using our system, a researcher will easily find out what has already been registered of his/her activities by other academic or administrative partners. Hence, they can easily detect incomplete data pertaining to themselves and at any time make the necessary corrections and supplements.

Documentation of different forms of results of research activity

This is the part of the database we have developed most completely up till now. Here we already have a primary version of the system in use throughout the university. This part covers e.g. a major part of our data requirements in relation to annual reporting (see section 1.1). The research results that researchers are asked to document ranges from traditional result forms such as articles, presentations, participation at conferences and on committees, to activities involved in the exchange of researchers between institutions. All information types are created in such a way that data can be retrieved for a number of uses. We have to this end e.g. made distinctions between national and international activities and publication channels. We have also, as a response to the government's actions to stimulate and document the success of female researchers, been careful to make the proper gender distinctions. This, in turn, paves the way for some very revealing new types of statistics.

Since information is supposed to be given in great detail, we find it to be an advantage if the researchers themselves are willing to input their data. As this will not be possible in all cases, the system at some points also offers the possibility to enter data in an aggregated form, and in some cases by administrative staff. In all our contacts with data providers, we go to great lengths to appear as flexible as possible in matters of forms and levels of data entry. One of our hidden objectives here is clearly that we find it more important to have some data, if only on an aggregated level, than no data at all. As it has turned out, part of this game of documentation also presupposes job qualifications in our staff in counselling and applied psychology!

A Ph.D. abstracting and press release system

As it is today, the university has its own system for retrieving information about Ph.D. candidates prior to the defence of their theses. Information on the doctoral students and their work have up till now been collected in a traditional way by means of written statements made by the students. In our system, the complete preparation of press releases are web-driven where the students contribute information in a standardised way. The Information Department then edits the information, adds information on the web, including a photo, and produces on the fly the complete press release in web format to be utilised by the press. As the research performed by our doctoral students constitutes a major research effort of the institution it is regarded as a priority to communicate effectively the contents of this type of research.

In addition, content abstracts and other scholarly information about Ph.D. dissertations are offered, and in such a way that this information may be exported to international databases in the field, most notably the Dissertation Abstracts of Bell and Howell.



Expertise documentation

As a function of the researchers' use of the new database, the university has, furthermore, at its disposal a suitable web-based tool for communicating effectively to a range of user groups the wide area of expertise that exists at our university. One major target group here is the surrounding public and private community of the city of Bergen that may potentially be interested in the expertise of the university. By linking data from different parts of the database, we will be able to provide a multi-layered information service that will be able to market our researchers also as potential partners of commercial enterprises. In this way the service may function as the starting point for actions of transfer of our research to the community at large. To this must be added that inclusion of a researcher's profile in such a service, must, by necessity, be subjected to the consent of each individual researcher. At least in our country, we find today a rather wide-spread unwillingness to appear in research information overviews directed at industry. The reasons for this state of play are many.

A system for web-based annual reporting from primary to top-level in the university

By aggregating the data reported at departmental level (as described in sections 2.1.1-2.1.4), we are able to cover most of the research activities that we are required to report annually to the Ministry. The adoption of one common system for the reporting of all research-based activities, has turned out to be a productive way of validating the different data sources, and also makes it easier to extract data for different types of statistical analyses. We should not, however, at this point leave you with the impression that this is all plain sailing.

There is, clearly, a deep-seated scepticism among many researchers towards our new documentation systems, as these may produce better and more detailed forms of documentation of the work of and output from all individuals in the scholarly community. And rightly so: our new systems have the capacity of being put to a range of uses that may have the effect of monitoring the success or failure of individuals in a number of ways. It is, therefore, of utmost importance to reassure our researchers of the motives behind our system development. We must convince them that the data we collect are for the common good and that the data acquired will, in turn, yield valuable information that may enhance our knowledge base in the field considerably. Such a base is of great value when working in central administration towards improving the financial and other conditions of researchers and their departments and in order to make our institution statistically competitive in the common race of universities for more central funding.

A full-text self publishing support service for master and Ph.D. students and researchers

At the University of Bergen, we have established a project where graduate students are offered to publish their theses on the web in addition to the submission of their work in a hard copy format. Hopefully, this service, which is based on a contract with each individual

author, will in the coming years be extended also to cover doctoral dissertations and the work of researchers.

In this area, the university intends to make extensive use of experience gained internationally when universities set up their web-based publication services for students and researchers. In our case we have decided to start our learning process by concentrating on making web-presentations of master and Ph.D. theses, as these have, up till now, had a much too restricted readership. One reason for this has been that only a very restricted number of paper copies are available.

Within this publishing service, our CRIS database will supply kernel information on each piece of work including metadata formation and offer one single entry point for all uses of research information related to our institution.

3 Introduction of the new system

One of the most difficult and challenging activities when developing a new database, is, as would be expected, the smooth introduction of it to the main user groups. The database we are presently implementing, is in a series of ways directly dependent on the acceptance of it by the community of researchers, in their dual functions as both data providers and users. As described in section 2, the services provided by the database will also to a great extent be used by administrators on all levels at the university, by our national authorities, and, hopefully, by other users as well, outside of the university.

To obtain the most correct data possible, we believe that one major success criterion is to capture data as close to the source of information as possible. This means, as mentioned earlier, that we put a lot of effort into getting the researchers to input their data themselves. Since we already have had in operation a research documentation system since 1993, we have the advantage of a long established contact network with a broad spectrum of users. The challenge today is to get the users of the old system (which they have learnt to use by putting in great efforts!) to switch en masse to the new system, and also to enlist the co-operation of researchers who for one reason or another have not at all so far been documenting any of their activities.

We use a double set of strategies to accomplish this task: Firstly, we make extensive informal contact with a great number of individual users and with officers in charge of university registers. This we do by staging information meetings where we explain our objectives and demonstrate how the system works. We also respond to a variety of questions, some of which tend to be of a more or less unappreciative nature. Also, we provide extensive, on the spot, user support by phone and e-mail. Secondly, we actively involve administrative staff in the departments, both managers and secretaries. Managers are here important actors because they will be able to perceive and communicate to others the immediate use of the data in the database in their daily work. Secretarial staff at the departments have over the years also turned out to be most helpful in inputting data for researchers who will not or cannot enter their data themselves. For the benefit of this group we also arrange special information meetings related to their uses of the system as intermediaries.

Since we for some time have had in place an older system for documenting research activities, we have, as mentioned, already an established a network of contacts that we

may draw on and that comprises our core group of users of the new database. We believe that this will ease considerably the introduction of the new system. To alleviate foreseen transition problems we also offer researchers, as an alternative, the possibility of using the old system for one more year. Data from the old system will then be imported to the new one once the annual reporting for the year 2000 has been finished.

By not forcing our users to enter data in the new system from day one of its introduction, we offer a kind of flexibility that we think users will appreciate. Our somewhat hidden strategy here is to rather let them be convinced by word of mouth from colleagues and others that the new system is much to be preferred. In many ways we believe to have won the battle if researchers learn from their colleagues that by using the new system they will be able to do their mandatory reporting tasks with substantially less work and with a much higher degree of quality. Also, we hope that they will be even more co-operative when they find that they may make use of a variety of reporting and exporting functions in relation to their personal entries in the system.

4 Experience

In this section we will discuss briefly some of the experience we have drawn so far from designing a new system for documenting research.

As mentioned in section 1.1, the University of Bergen has a suite of documentation systems covering students, teaching programmes, accounting and personnel functions and research. In designing a new system, we have found it important to make maximum re-use of the data already held in these systems. Wherever possible, one should try to adjust the new system to the ones already in use, as this will probably facilitate the development of the system, and later on, the data capturing processes. We have, accordingly, put quite a lot of effort into getting core data from the other key systems in question - but with varying degrees of success. The final solution, we think, is to organise information that will be shared by several database systems in the form of mirror databases specifically designed for the easy export of data to other systems.

In the first phase of the launching our new database during this winter, we have, as previously, found it easiest to co-operate with the researchers on documenting *research results*. On the other hand, when it comes to documentation of on-going *research activities*, we are often met with a series of challenges from the research community. The research activity, as viewed by the researchers, is simply a means to an end, and hence they often do not see the point in investing valuable time in describing their activities per se, even if this is also an important information source to university administration and for university's strategic planning activities. This is, nevertheless, an area of documentation that we will be focussing heavily on once the new system has been fully developed.

The final point we will draw attention to here, are the problems we face when we introduce a new system before it has been fully developed. Our experience is here both positive and negative. Positively, we in some ways facilitate the development of a flexible system by the chosen procedure, since users' responses are given and may be taken into account as we go along designing and developing the system. Also, the users may acquaint themselves with

the system at an early stage of its development and therefore will have the advantage of getting to know the different aspects of the system over a longer period of time. Negatively, we get thrown back at us a series of problems that normally appear only when a new database has reached a proper stage of introduction - in our case to about a thousand users. Such problems we have to tackle in parallel with fundamental development activities. Here we are, as is normal in this kind of work, faced with a wide set of errors ranging from simple system errors that are easy to correct, to important missing functional requirements that have not so far been considered or developed.

5 Success criteria

So far in our work, we find reasons to believe that our new system for documentation of research at the University of Bergen will be a success.

We think that the following elements will be instrumental in securing the general adoption of the system:

A system for registration of research results and on-going research activities must be developed and presented in such a manner that the researchers find the system to be very cost-effective. In addition, it must cater to other documentation needs than the conventional registration and reporting regimens traditionally imposed on researchers by their peers. The researchers must be made to see not only the gains by offering a labour-saving data inputting facility, but also the advantages of a system that may make active use of the data for a series of documentation functions. Some of these must be shown to be vital for securing the necessary financial foundation of the institution.

Most importantly also, the researchers must be given the opportunity to flexibly use the system and their own registered material for own documentation purposes, e.g. when making personal bibliographies and overviews of their research activities.

Another important element is to ask only for data that we cannot retrieve from other sources. The golden rule here is never to ask twice for the same data from our respondents. We also have to bear in mind that activities that steal time from research and education will instinctively be met with scepticism from researchers - and rightly so.

Furthermore, we find that spending a lot of time with our users, and treating them in a professional way, is important if we want to achieve an active co-operation on data capturing at department level. This may seem trivial, but we find that getting to know our users personally and giving them the necessary, ad hoc information about the database, is a very important success factor for getting both extensive and valid data returns.

And last, but not the least, we think that support from the top management at the university is a major success factor. For example, in our institution we have found that the single most important change of view in the research community, regarding our documentation work, turned out to be when the university started to actively use our database to automatically produce publication reports etc from the various departments. Previously, more or less quality-controlled paper forms were in use for the collection of aggregated data on scientific publishing. This change induced the departments to document per item all their scientific contributions instead of just providing the traditional numbers of publications.



As time goes by researchers will be able to explore more in depth the different uses our new system can be put to. As an outcome, we hope to be able enlist their co-operation to an even higher degree for the benefit of university as a whole.

