Remodelling Frida – from institutional registration to common registration and responsibility across member institutions

Grete Christina Lingjærde, Andora Sjøgren
Centre for Information Technology, University of Oslo, Norway

Summary

Frida is a research documentation system. A national group was established in 2008 to look at the possibility of establishing a common national system (also including all the research institutes and Health Trust hospitals) in Norway. The group recommended that a restructured Frida was the best solution, and the Ministry of Education and Research decided in 2009 that Frida would be the common research system in Norway for 150 institutions. The main purpose of the restructuring was to make more of the data shared / common between the institutions and while still allowing for institutional view, ownership and responsibility of the data. This paper will focus on the restructuring of Frida.

1 Current Status

Frida (Research results, information and documentation of scientific activities) is an information system developed for the four Norwegian universities Norwegian University of Science and Technology, University of Oslo, University of Bergen and University of Tromsø. The system was developed to be an authoritative system for the documentation of research activities at the individual institutions. Frida is currently owned by the four universities listed above, but six other institutions also use the system. Frida consists of five modules:

- Research Results
- Catalogue of projects
- Catalogue of scientists
- Research units
- Annual reporting

Frida is today a system specific to the individual institution. Each institution has its own database for registration and controls their own data. The data are physically collected within one database. The institutions have, however, their own data logically separated. This means, for example, that publications shared by several institutions (where the authors come from different institutions), have to be registered several times, and appear as several independent occurrences. At the same time the institutions share some of the data in Frida. This is for example the case with the authoritative registers for periodicals, publishers and institutions. Several code tables, such as those for country, language and so on are also shared. This solution was originally chosen to simplify the operation of Frida. It is easy to run the system, and it is easy to put new institutions into operation.
Data from Frida are used to generate statistics about research activities within the Norwegian universities. Data about scientific publications are also reported to the Ministry of Education and Research where they are included in the Ministry’s model for financing the institutions. The system must therefore also be regarded as an economy system, and has to fulfill requirements specific to those kinds of systems.

2 Process

Several institutions within the sectors Higher Education, Health Centers and Research Institutes have shown interest in a research documentation system. Society has a need for increased access to and information about research activities and research results. In Parliamentary Report Number 30 “Climate for Research”, the Ministry of Education and Research emphasizes the importance of easily accessible research results. The ministry also emphasizes that research resources should be administered efficiently. A national information system for research activities would be an important step in this direction, and would contribute to a strengthened profiling of the research activities within these sectors. A high level of interaction between the sectors reinforces the need for a common research documentation system.

In 2008, the Ministry of Education and Research therefore appointed a committee to consider the possibility of achieving one common system for these three sectors for registration of scientific publications. The committee also considered the need for a research documentation system that covered other research activities in addition to scientific publications. A premise for the solution was that data needed to be registered only once across the institutions, while the control of and responsibility for the data quality remained at the individual institution. The committee delivered a report in September 2008. A restructured Frida was recommended as the common national system. The main purpose of restructuring Frida is to increase the amount of data shared by or common to the institutions. This is for example relevant for publications with authors from different Frida institutions. A challenge in achieving this goal is implementing and strictly enforcing common routines and common responsibility for such shared publications when different institutions to the Ministry report these.

In the autumn of 2009 the Ministry decided to follow the recommendations and a restructured Frida will soon be the research documentation system for 150 institutions in Norway. Frida has consequently been chosen to be developed into a national database for the documentation of scientific publications for all three sectors (Higher Education, Health Centers and Research Institutes). The database for research publications will be called NVI (Norsk Vitenskaps Indeks - Norwegian scientific index). NVI itself will be included as a component in the national research information system, Frida (see Figure 1).

![Image](image188x73 to 276x164)

Figure 1: Norwegian science index, a module in Frida
Data can be registered across member institutions, but because the control of and responsibility for data quality remains at the individual institution, a specific institution is only responsible for the control and quality of their contribution to an occurrence (entry for publications, monograph, anthology, and so on), while other institutions remain responsible for the quality of other parts of the same occurrence. This implicates further the necessity for well-structured data. Furthermore, common routines for data registration that are valid for all institutions must be implemented.

The national research documentation system will hopefully give Norwegian research increased social value by providing an opportunity to see correlations between research in three different sectors. The system will also serve as an information source for both researchers and society as a whole. A restructured Frida offers the possibility of functioning as the Yellow Pages of scientific activities in Norway. All researchers using Frida are registered in the catalogue and can be retrieved by users searching for researchers, projects, publications, etc. The report about NVI lists a number of arguments for a database that reflects a common publishing arena where researchers can register their publications once even if the contribution of the researcher should be credited to more than one institution. In the autumn of 2009 the Ministry of Education and Research appointed several committees to specify the task of restructuring Frida. These committees will deliver their reports in February 2010. Planning for the restructuring of Frida has already started and will be finished before the summer of 2010. The year 2010 will be an interim year. The original owners will still be the owners of the system. The Ministry will contribute to the financing of the restructured Frida. The system will be offered to all the relevant institutions from January 2011.

3 The solution – a restructured Frida

Frida is going to be restructured in such a way that we achieve a common unified database. This restructuring will simplify registration for researchers who are associated with several different institutions. The new system will provide a better overview and presentation of Norwegian research and a more suited base for analyzing research data.

3.1 Data model and VPD

A prerequisite for the solution was that Frida was built on a well-founded data model. This makes it relatively simple to change and adapt the system due to new demands, altered needs and a changing future. Another prerequisite for the solution was the VPD- (Virtual Private database) features in Oracle that was taken advantage of in order to simplify maintenance. One physical database provides for all the institutions, but the data are logically separated for each institution. The fact that each person employed by an institution or associated with an institution is uniquely identified by his or her Norwegian social security number, was also integral to this solution. The suggested restructuring of Frida will make the sharing of common data larger and at the same
In the restructured Frida, the publication references will be shared/common among the institutions.

### 3.2 Global and local Frida

The solution requires that we maintain the institutional specific “databases”. However, various components will be shared or partly shared among the institutions. Furthermore, we will introduce two concepts and divide the data between these concepts: A global part that means a common/shared part and a local part that means the institutionally specific part. The global components/concepts are those that do not need to be registered more than once and where it is acceptable to share information both in respect to registration and viewing between the institutions. The local components are those that do not contain duplicate information, but contain institutionally specific information where only a certain institution has the rights necessary to register data or change the data content. Examples of global and local components are:

**Person: Global and local**

Global person contains information that uniquely identifies person, name, title, etc.

Local person contains information about employment at the institution. The institutions themselves decide who is allowed to access this information.
Publications (Scientific results): Global an local

Global_Scientific_publication contains information about the title, references to the journal etc. When an author is referenced from the Global_Scientific_publication it refers to the Global person. Local_Scientific_publication contains mainly information concerning the approval of the post/entry. The institutions themselves are responsible for approving data that are reported to the Ministry. Examples of this type of information:

- Whether the post/entry has been approved or not
- Id of the user who approved the post/entry
- Date for approval.

Other local data connected to publications are for example institutionally specific questions and answers concerning a submit service to institutional archives. Figure 4 shows a high-level model of the relations between global and local concepts.

3.3 Other common tables

Other information objects that are not in the above figure are projects, research groups, organizational units and research centres. These objects will also be global and shared between the institutions.
This restructuring of the Frida system (see Figure 5) will make more of the data shared/common between the institutions while maintaining the institutional view, ownership and responsibility for the data. By restructuring as described, the system will satisfy both institutional and national needs.

Figure 4: Part of the new data model of Frida
4 Benefits with the restructuring of Frida

- No need to make a separate view/presentation of data for a national presentation
- The number of duplicates will be reduced, especially in a national perspective
- Data is more suited for statistics and data-warehousing for advanced sophisticated analysis of data
- Reuse of what is already implemented in Frida makes restructuring possible at reduced costs
- Common data is not duplicated in the institutional databases
- Disjoint and private data are stored in the institutionally specific areas/database.
- The institutionally specific view, ownership and control of data are retained
- Rationalization. Closing similar systems and focusing on one system will benefit more users and stakeholders.
5 Consequences

5.1 Log on

The solution requires that a user have to log in via an institution. The registration of publications, projects, etc. only has to occur once even if the user is affiliated with several institutions. When researchers from other institutions are identified with the same occurrence of a registration/data entry, they will only need to register supplemental information. Functionality will be developed so that a user affiliated with several institutions can log on via one institution and go directly to the others by one keystroke. Annual data for this institution may then, for example, be registered.

5.2 Converting data: Merging data/splitting data

Existing data in Frida must be converted to the new structure. Data records representing the same publication must be merged. Representations of projects must be merged if currently registered in different logical databases.

Data records that belong to one institution in the current system must be split into local and global parts.

Institutions that currently do not use Frida may have data in other databases. These data will be converted to the new structure in Frida.

5.3 Quality assurance

In developing and further developing Frida, we have had special focus on quality assurance. Many business rules and functions have been implemented in the system and common rules for registration have been worked out. The fact that data from Frida plays a part in the Ministry’s financing model when allocating money/budgetary funds to the institutions leads us to also regard Frida as an economy system with the following requirements:

- Routines for quality assurance must exist
- Registration and editing of data must be traceable. Who? When? Etc
- Controls and business rules may be obeyed.

Examples of functionality implemented to enhance quality:

- Import of data from ITAR (Import Service and Authority Registers, data from bibliographical sources such as ISI)
- Authoritative registers (from ITAR)
- Code tables
- Structured data model
- Business rules implemented in the database
5.4 Approval of data

All data records that will be reported to the Ministry have to be approved by someone other than the researcher. A lot of functionality has been developed in Frida in this connection, and it is continuously being improved based upon the experiences of the member institutions. When Frida is restructured we have to go through this functionality once again.

5.5 Running rules and common routines

In addition to the new structure there are some conditions that must be fulfilled so that the model can support reporting of the data to the Ministry’s database DBH (Database for Higher Education) in a safe and secure way.

When common data records are to be reported to DBH, all institutions reporting data must follow these rules:

- The deadline for reporting data must be the same for all institutions
- The deadline for data registration by the researchers must be the same
- All institutions which are reporting this data record, must have approved their part of the record
- After a data record is approved locally at one institution, it is not possible to change important data such as the number of authors, publishing channel, etc. by any other institution. Changes at this point must be agreed upon by the super users at the different institutions
- After a record has been reported, super users can no longer change content

5.6 Challenges

- The data model is becoming more complicated
- The institutions must adjust to common deadlines and obey to common rules
- All the institutions depend on each other. Data registration must be done within a reasonable timeframe.
6 Conclusion

A restructured Frida will benefit all the institutions in Norway that are doing research and all that are interested in research. Putting the resources in one system will be more cost efficient and the improvement will benefit more users. Doing things in a common and similar way it will make it easier to search for, retrieve and compare the data between the different institutions. Frida may in the future function as the Yellow Pages of scientific activities in Norway. All researchers using Frida are registered in the catalogue and can be retrieved. The user may search for researchers, projects, publications, research groups, etc.

This restructuring of the Frida system will make more of the data shared / common between the institutions while maintaining an institutional view, ownership and responsibility of data. By restructuring as described above, the system will simultaneously satisfy both institutional and national needs.

Contact Information

Grete Christina Lingjærde
Usit /University Center for Information technology
University of Oslo
Forskningsveien 3B,
0313 Oslo
Norway
g.c.lingjarde@usit.uio.no