The practical implementation of the CRIS system CRIStin and the goals/challenges of bringing 150 institutions into production within a year

Agnethe Sidselrud, Grete Christina Lingjærde
CRIStin and Centre for Information Technology (USIT), University of Oslo, Norway

Summary
FRIDA was the research documentation system for the four oldest universities in Norway from 2003-2011. The Ministry of Education and Research decided in 2009 that a restructured FRIDA should be the common research information system in Norway. The restructured FRIDA was named CRIStin - Current Research Information System in Norway.

In this paper we will give an overview of both the development of the CRIS system and the process of putting 150 institutions in production. Experience, challenges and obstacles are addressed and discussed. The new system will benefit more users and stakeholders such as researchers, research-administrative staff, the government and the public. One of the benefits of having one common national research information system is that there is no longer a need for a separate interface in order to provide a view of national data.

1 Background

The main purpose of restructuring FRIDA was to make more of the data shared among the member institutions while maintain an institutional view, including ownership and responsibility for the data. The transition from institutional registration to common registration and responsibility across the institutions was a challenging task that has not been solved before.

CRIStin is a collaboration between the Ministry of Education and Research and the Ministry of Health and serves as a common information system for all scientific publications in Norway and for all publicly financed research institutions: the universities and university colleges, the independent research institutes and the regional health authorities.

The system is expected to stimulate and increase the societal use of research results. In 2012 CRIStin will serve the Norwegian government’s performance-based budgeting system in three research sectors.

From the summer of 2010 till the autumn of 2011 the CRIS system CRIStin was developed and 150 institutions were put in production during this short and hectic period – a challenging mission. The system is better suited for statistics and as a data warehouse for advanced sophisticated analysis than its predecessors. Closing similar systems and focusing resources on one national system will benefit more users and stakeholders. More benefits and opportunities are discussed later in this paper.

This work is licensed under a Creative Commons Attribution-NoDerivs 3.0 Unported License.
1.1 Common registration

The new features in CRIStin compared to the previous systems used by the various institutions, was that the 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.

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 responsible only for the control and quality-check of their contribution to an occurrence (entry for publications, monograph, anthology, etc.), while other institutions remain responsible for the quality of other parts of the same occurrence. This implicates the necessity for well-structured data. Furthermore, common routines for data registration that are valid for all institutions must be implemented. For the individual institution the task of implementing CRIStin was not only the technical implementation of CRIStin but also of learning and implementing rules and routines for the common annual reporting.

1.2 Rules

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

- The deadline for reporting data must be the same for all institutions
- The deadline for data registration for the researchers must be the same
- All institutions which are reporting their 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 the most essential bibliographical 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 member institutions
- After a record has been reported, super users can no longer change the content

2 Two phases

During the implementation phase it was decided by the ministries that only the Higher Education sector was to report their 2010-data on April 1, 2011 using CRIStin. For the two others sectors (health sector, and research institutes) a test reporting was to be conducted with the deadline on September 15, 2011. It was not possible to have 150 institutions starting up in a few months simultaneously, so this was a reasonable decision and a good solution. Institutions in the Higher Education sector were also considerably more familiar with the predecessor FRIDA and CRIStin would not be that different as far as the user interface is concerned. Although the two-phase-implementation process was a wise and quite logical decision, we met some challenges - both technical and administrative - as system specifications and the system implementation were based on common rules and routines as described above.
2.1 The plan and the execution

To be able to achieve all this in such a short time, planning is essential and one of the most important success factors. A lot of the activities have to start simultaneously and must be administrated in a flexible way. Both technical and administrative processes must be initiated and synchronized. The large number of new institutions and users must be handled professionally. A strong belief and confidence in achieving the goal together is a must all the way in the process.

Since the CRIStin system was a restructuring of a system already in use, it was easier to start many different activities simultaneously. The super users of FRIDA functioned as a resource and a reference group with their experience and expertise that was so important to involve.

The next part of the paper presents some of the main activities in the process.

2.2 Local institutional data

Before an institution can access the system, the system must be populated with institutional data:

- Personal data (names and ID-numbers of all researchers, their titles and contact information)
- Organisational data (names and code-numbers of all the departments and the sub-departments)

The gathering of personal data and organisational data at the institution was an important and crucial task in order to succeed. Preparation and involvement of the new institutions in this process, the dialogue with and instruction of the users were all time consuming and challenging tasks due to the short timeframe.

Guidelines were sent to all institutions and a small team in CRIStin was dedicated to managing this task. Numerous quality checks on the personal data and the organizational hierarchy were necessary before the data was imported into CRIStin. The additional challenge was the level of competence in handling and understanding of this type of data and structure that varied quite a lot at the user institutions.

The exchange format is based on predefined common XML schemes. The automated service was established beforehand (already in use at the “FRIDA” institutions) and was used by some of the new institutions. The smaller institutions were allowed to send their data in Excel over a secure line, this due to need to speed up and complete the process. When full automation is implemented all institutions are expected to deliver the personal and organizational data in the XML schema.

2.3 New Law and the use of social security number/personal identification numbers

A prerequisite for the solution and gathering all of institutions in Norway into CRIStin was that we be allowed to make use of social security numbers as a candidate key for a researcher. 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 crucial and the use of this number was the only way we could integrate the data from different systems and connect automatically information about employees with multiple affiliation etc. The use of social security numbers for
identification purposes demanded a focus on data security and the secure transfer of data between the systems.

Current legislation in Norway allows institutions to submit the social security numbers of their employees, however the submission is not obligatory in case some of the researchers or the leaders at the institutions reject this practice. In order to ensure the completeness of the national data in CRIStin, a new law was introduced forcing the institutions to submit the social security numbers to CRIStin. They law was delayed almost a year, but the institutions were eager to become members and submitted the measure despite the fact that the law was not officially passed by the Norwegian government in 2011.

### 2.4 Converting data from the old system

Existing data in FRIDA had to be converted to the new structure in CRIStin. Data records representing the same publication had to be merged. Representations of projects had to be merged if previously registered in different logical databases. Data records that belong to one institution in the current system had to be split into local and global parts.

Institutions that originally did not use FRIDA had their data in other databases. These data were in some cases converted to the new structure in CRIStin. The publications from the other large Cris system in Norway – Forskdok, used in the university college sector - had to be converted into CRIStin. Our experience from this process is that moving data from a system with less structure to a system with more structure needs to be thoroughly tested and quality assured and some of the work in the conversion process must be done manually. The dedication of time and focus for these activities at the member institutions is necessary.

### 2.5 Testing

It was very useful and necessary for the developers to have real data in the system when working with the restructuring of the system. The presence of some of the data in the system made the process faster and enabled the teams to start different activities simultaneously.

A test team was established early and consisted both of the support team in CRIStin and the programmers from USIT. To rewrite an existing system under such time pressure forced certain decisions: we used a larger number of programmers to test more of the functionality at the user level since they had better knowledge of the type and the cause of possible errors than the new institutions.

### 2.6 User involvement

It was very important to involve the users at an early stage and keep them involved, motivated and close all the way.

At an early stage we focused on making documentation for the users:

- Routine description for the annual reporting of the publications
- User manuals
This documentation was written simultaneously with the process of the technical implementation of CRIStin. It was therefore necessary to ensure a close collaboration between the support team and the technicians all the way.

Several start up visits around Norway were organized in order to get in touch with the new user groups. The institutional FRIDA super users were of great help to us and functioned both as mentors and system experts for the super users at the new institutions. Their involvement was essential for our success.

3 New sectors

New sectors were introduced to a CRIS system and the introduction met some sector specific challenges. Some of the sectors were used to annual registration of their institutional data in a Cris system (the Higher Education sector). Independent research institutes and the regional health institutions were not so familiar with the idea of CRIS. Some of the challenges in this implementation process were similar to the implementation and the early days of FRIDA and a lot of different questions were asked again in connection to the performed based funding model. Our prior experience with FRIDA provided a good background for handling some of the issues. In addition, as mentioned above, a FAQ and guidelines were made to help users.

4 Challenges

- Use of social security numbers for the identification of researchers
- Data security
- Dependency of data from other systems
- Technical issues like load and performance
- Logging, authentication
- Converting of data from other system with less structure and less quality

5 Ten most important lessons learned

1 Planning is important. When the implementation process is expected to take only some months (which is extremely short and rather challenging deadline to reach) and several of the activities take place simultaneously, every part of the implementation plan must be described. Timeframe (number of weeks) and human resources (specific team members) should be clearly identified.

2 Being flexible is essential. When the deadline for the system implementation is also the deadline for the national annual reporting to more than one ministry, no delay in the implementation process can be allowed for and reaching the deadline is one of the key success factors. The considerable amount of pressure on the support team, the technicians and the involved contact persons at the new institutions can be moderated by allowing
some flexibility in execution of some of the subtasks – either in regard of alternative technical solution or in regard to assistance to the new institutions.

3 If you depend on getting data from other systems, start the process as early as possible and get help from leaders that follow the process from the top level. Identify the risks of not getting the data on time and schedule alternative solution. Include the risk assessment in the master project plan.

4 Close collaboration between the developing group and the administrator/support team is very important. In a project with high political involvement and much publicity, a high level of expectations from the authorities, the institutions and the researchers – the pressure on both groups (developers and administrators) is considerable. It is crucial to focus on team building and values such as mutual respect and trust in each other’s competence level, flexibility and very close dialogue. The commitment of both groups is one of the most important key success factors. Both groups must share genuine dedication to the project and professional integrity.

5 Make the results visible during the process. Making the components in the project visible gives both the authorities and the user better idea of the overall product, ensures motivation and enables the necessary involvement of the potential user groups.

6 Look everywhere for help. Super users at the some of the institutions were of great help. Use all available competence and invest in building networks of user groups. Such networks – either regional groups (meeting regularly) or e-mail-lists – are of great importance in building competence on the system on the national level. Close collaboration with the user networks could potentially lessen the pressure on the support and training team.

7 Converting data from less strict data structure to a highly structured date takes time and does not automatically improve quality. Make detailed plans for the conversion process and dedicate enough time to test the results. If possible, make the authorities, the institution that owns the data and the users aware of the complexity of the task. Make sure you have resources to work with the data manually in case some corrective work must be done.

8 Provide the user with documentation as soon as possible. The institutions need clear guidelines (especially when a common system is supposed to be used across more than 50 member institutions), detailed user manuals and good support team that they can contact and ask about anything they need to understand better. User manuals for dedicated operations or smaller parts of the system are a good way to diminish the impression of the big size and complexity of a new IT-system.

9 Implementation in new sectors creates sector-specific challenges. Do not underestimate the importance of involvement and acceptance from local leaders, both at the new member institutions, but also at other relevant committees or groups that are responsible for the political decisions that may have influence on the areas the system is currently – or will be in the future - used.

10 To rewrite an existing system under such time pressure made us make decisions that diverge from our usual practices: we used larger number of the programmers to test more of the functionality at the user level since this group had better knowledge of the type and the cause of possible errors than the new institutions.
6 Conclusions

We were able to finish this mission as planned and reached our goal despite all the challenges. It was a great advantage that the CRISin-system was a restructuring of an existing system - FRIDA. Planning and close collaboration with all the involved groups were essential. A strong belief and confidence in achieving the goal together was a must and a prerequisite all the way in the process.
References

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