SYSTEMS DEPARTMENT RESEARCH ACTIVITIES CENTER EMBEDDED DEVELOPMENT WELL EDUCATION TECHNOLOGIES COMPUTER FUNDING EXTERNAL TOOLS MOBILE ORGANIZATIONS SOFTWARE SYSTEMS RESEARCH PROFILE EVALUATION DEPARTMENT SINCE COMPUTER 2006-2010 ANALYSIS DANISH SCIENCE SERVICES INNOVATION METHODS CASSIOPEIA WEB-BASED DATA FUTURE

increase Ph.D villum new STUDENT STRATEGIC REAL-TIME PROGRAMMING DATA-INTENSIVE semantics high EMPLOYED GENERAL number FOUR DEPARTMENT'S YEAR PROGRAMS serves ALONG MISSION GROUPS IEEE DISTRIBUTED honorary existing COVERS BROAD PARTICULAR FIELDS DKK USE INDUSTRY AUTONOMOUS amount information FOUNDATION continued FRANCE TOTAL GRANTS activity particularly verification TERMS COME economic INTERACTION DESCRIPTION performs reasoning many ENVIRONMENTS ThIRD TAKES context-based DISTRIBUTED ACCESS DOCTOR main INTEGRATED RESULTS PREPARED languages GRANTS SINCE LOCATION researches APPOINTED intelligent REGIME FOCUS ens highly BUSINESS point OUTSTANDING internet RASMUSSEN'S REPUTED UNCERTAINTY SOFTWARE successful science PROGRAMMING million EUROPEAN focuses SEVERAL UNITS PRIDE AREAS FORMALISMS within computerised ORGANISATIONS KEY ESTABLISHED meet LEARNING companies DESIGN PRIVATE national researching ADAPTION PER scientists TRANSFORMING technologies human-computer COUNCILS PUBLIC intelligence FELLOW

Research Evaluation
2006-2010

Department of Computer Science
Aalborg university

Computer Science, Aalborg University – Research Evaluation 2006-2010

Preface

It was in the spring of 2010 that the Department of Computer Science at Aalborg University decided it would continue the tradition of conducting a research evaluation. The present report documents the fifth research evaluation of the Department of Computer Science and covers the period 2006–2010.

The process that led to the production of this document began in early 2010 when the department's research leaders, Professors Torben Bach Pedersen, Kim G. Larsen, Peter Axel Nielsen, Jan Stage, and Manfred Jaeger, Associate Professor, along with Kristian G. Olesen, Head of Department, initiated discussions on the structure of the process. Early on, an evaluation committee was appointed. Professor Jan Stage agreed to be the local member of the committee. (As the local member, he would not have a vote.)

As a new way of approaching the evaluation, it was decided that each research unit should identify a committee member to evaluate it.

The department was delighted when Professor Jos Baeten (Technische Universiteit Eindhoven, the Netherlands), Professor Ann Blandford (UCL, UK), Professor Yannis Manolopoulos (the Aristotle University of Thessaloniki, Greece), and Professor Petri Myllymäki (University of Helsinki, Finland) agreed to be on the committee.

We felt that this committee of recognized and experienced senior scientists and research leaders would be able to fully address the research areas of the department as well as provide an insightful and wide-ranging evaluation. (Short curricula vitae of the members of the evaluation committee may be found in Chapter 2.)

In January 2011, a two-day research evaluation seminar was held at a conference facility. All research staff, including Ph.D. students, long-term visiting staff, and technical and administrative staff from the department were invited.

Prior to the evaluation seminar, the committee conducted a full day of meetings with department representatives (the head of department, research leaders, unit leaders and

editors) where each unit gave a short presentation to the committee allowing the committee to pose questions, both to the overall organizational structures as well to the other research units.

On the first day of the seminar, each research unit worked with their own committee member. They first gave oral presentations of their research, allowing the committee member to ask questions on both the presentations and the preliminary report. During the afternoon of the first day, the evaluation committee assembled and worked on the evaluation. On the second day of the seminar, the committee presented its evaluation to the department and answered questions.

Designing the research evaluation and producing this report was a collaborative effort involving all the researchers in the department. I thank them all for their assistance and support. The editors for each of the groups: Simonas Saltenis, Jiri Srba, Jesper Kjeldskov and Manfred Jaeger, deserve special thanks for their work on the contributions by the research groups.

Henriette Frahm

Contents

1 Evaluation Guidelines

The purpose of these guidelines is to provide an overview of the tasks we expect the expert evaluators to solve at the research evaluation for Department of Computer Science at Aalborg University.

We expect that the expert evaluators contribute in the following way:

- Apply the best of his/her abilities, professional skills, knowledge and ethics, in evaluating the research at the department in accordance with the guidelines and time-schedules provided by the department.
- Read and assess the documentation provided by the department in early December 2010. Each evaluator is expected to focus on the documentation from one specific research unit (within his/her area of competence) but also to have an overall understanding of the results produced in the other research units.
- Work together with the department representative on the evaluation panel.
- Take part in the evaluation meeting on January 3-6 in Aalborg, with arrival late on January 3 and departure around noon on January 6. This will include:
  - A series of presentations by the department and each unit on January 4, in a plenary session.
  - A combination of presentations and discussions of research quality and development strategy on January 5, in a session with only the relevant research unit.
  - A presentation of the evaluation of the department and each research unit on January 6, in a plenary session where the dean will also be present.
  - A panel discussion starting with a 3-4 minute position statement by each expert evaluator on January 6 about "Research Strategy: The Role of IT on the Research Agenda", in a plenary session.
- Contribute to a written evaluation of about one page for each research unit and for the department. These contributions will form an expert evaluator report that is completed before the evaluators leave the meeting or right after.

2 The Evaluation Committee

The Evaluation Committe from left: Yannis Manolopoulos, Jos Baeten, Ann Blandford, Petri Myllymäki and Jan Stage.

2.1 Jos Baeten

Jos Baeten graduated in mathematical logic at Utrecht University in 1978. He has an M.Sc. (1983) and a Ph.D. (1985) in mathematics of the University of Minnesota (USA). He worked at the Technical University of Delft (1983-1984), at the Centre for Mathematics and Computer Science in Amsterdam (1984-1985 and 1989-1991) and at the University of Amsterdam (1985-1991), before getting appointed in 1991 as full professor of computer science at the Technische Universiteit Eindhoven (TU/e). In 2010, he switched to the Department of Mechanical Engineering at TU/e as full professor of systems engineering. He is well-known as a researcher in formal methods and theoretical computer science, in particular in process algebra and its applications. To date, he supervised 26 Ph.D. degrees.

He is scientific director of the national research school Institute for Programming research and Algorithmics (1996-2004 and as of 2009), he was dean of the department of Mathematics and Computer Science of the TU/e (1996-1999 and 2002-2004), initiated the Embedded Systems Institute in Eindhoven (1996), and was scientific director of the TU/e institute of post-master education SAI (2002-2004). He was visiting professor at the Università di Camerino (2004), Università di Pisa (2004), Università di Bologna (2005), all in Italy, and the Universidad Nacional de Córdoba (2006) in Argentina. He chairs the steering committee of the CONCUR conferences and is a member of the Koninklijke Hollandsche Maatschappij der Wetenschappen (Royal Holland Society of Sciences and Humanities).

2.2 Ann Blandford

Ann Blandford is Professor of Human–Computer Interaction in the Department of Computer Science at UCL, Head of Research Department in the Division of Psychology and Language Sciences, and Director of UCL Interaction Centre (UCLIC). She is Principal Investigator on two substantial research programmes on technology in healthcare. One of these (CHI+MED) is studying the design of interactive medical devices from a user interaction perspective, particularly focusing on human error and how to make systems more resilient. The other (Healthy Interactive Systems in Healthcare) is supporting studies of lay people's health information needs, with a particular focus on how people make sense of information related to their health conditions. The need for usable, safe designs is becoming more pressing as people take greater responsibility for managing their own care. A third project, on designing to support serendipity, is also looking at how people interact with information and how technology can better support people's information needs.

Ann graduated from Cambridge University with First Class Honours in Mathematics, and was awarded a PhD in Artificial Intelligence and Education by the Open University. She started her career in industry, as a software engineer, and gradually developed a focus on the use and usability of computer systems. In 1991, she joined the Applied Psychology Unit in Cambridge as a research scientist. She subsequently moved to Middlesex Uni-

versity, initially as a lecturer, later being appointed as Professor and Director of Research in Computing Science. She moved to UCL in 2002. She has chaired several international conferences, including NordiCHI 2010. She has over 200 international, peer-reviewed publications, including a Synthesis Lecture on "Interacting with Information".

2.3 Yannis Manolopoulos

Yannis Manolopoulos received a 5-year Diploma Degree (1981) in Electrical Eng. and a Ph.D. (1986) in Computer Eng., both from the Aristotle University of Thessaloniki. Currently, he is Professor and Head of the Department of Informatics of the latter university. Temporarily also, he serves as Rector of the recently founded University of Western Macedonia, appointed by the Minister of Education. He has been with the Computer Science Departments of the University of Toronto, the University of Maryland at College Park and the University of Cyprus. He has published over 250 papers in refereed scientific journals and conference proceedings and received over 3000 citations, scoring an h-index equal to 28. He has been principal investigator of ~30 international and national projects, whereas he has supervised 15 Ph.D. graduates. He is co-author of the books: "Advanced Database Indexing" and "Advanced Signature Indexing for Multimedia and Web Applications" by Kluwer, as well as "Nearest Neighbor Search: a Database Perspective" and "R-trees: Theory and Applications" by Springer. Among other events, he served as a main organizer of the 8th National Computer Conference (2001), the 6th ADBIS Conference (2002), the 8th SSTD Symposium (2003), the 16th SSDBM Conference (2004), the 8th ICEIS Conference (2006), the 10th ADBIS Conference (2006) and the 4th Balkan Conference on Informatics (2009). His research interests include data management (databases, data mining, web and Geographical Information Systems, and informetrics).

2.4 Petri Myllymäki

Petri Myllymäki obtained his M.Sc. and Ph.D. degrees in Computer Science at University of Helsinki, where he became a Docent in computer science in 1999. Prof. Myllymäki is one of the co-founders and the current leader of the Complex Systems Computation

(CoSCo) research group, and he has over 20 years of experience in research in the area of intelligent systems and machine learning. His current special research interests are focused on Bayesian and information-theoretic modeling, in particular with Bayesian networks and other probabilistic models. He has been an editorial board member, program chair/committee member and reviewer for numerous international scientific journals, conferences and organizations, and he has published over 100 scientific articles in his research area. He is a co-founder of Ekahau Inc., Cloud'N'Sci Inc. and BayesIT Inc., and has also been responsible for several applied research projects with other companies including Nokia, Elisa, M-Brain, Kesko, Kone, StoraEnso, Space Systems Finland/European Space Agency, ABB and AlmaMedia, and the industrial co-operation has led to a number of fielded applications and patents. Dr. Myllymäki is a professor at the Department of Computer Science of University of Helsinki, and he is also working as a programme director at Helsinki Institute for Information Technology (HIIT), a joint research institute of University of Helsinki and Aalto University.

2.5 Jan Stage

Dr. Jan Stage is full professor in information systems and human-computer interaction at Department of Computer Science, Aalborg University. He obtained a PhD in Computer Science from University of Oslo in 1989. His research interests include methods for usability evaluation, interaction design and usability evaluation of mobile systems, object-oriented analysis and design, and prototyping. He has produced more than 100 publications, including articles in top-level peer-reviewed journals and conference proceedings in human-computer interaction and information systems, and an influential textbook on object-oriented analysis and design. He is leading a national research project on usability engineering in software development and involved in European research on usability. He is associate editor of Behaviour and Information Technology, member of the editorial boards of Journal of Database Management and International Journal of Mobile Human Computer Interaction, and was guest editor for International Journal of Human-Computer Interaction in 2006 and Journal of Systems and Software in 2010, member of the program committees of the Interact 2009 and NordiCHI 2010 conferences, and reviewer for several journals and conferences in the HCI area.

3 The Department of Computer Science

Departmental Management and Secretariat.

3.1 Executive Summary

The Department of Computer Science performs at a high level, and is well prepared to meet future challenges. The department has many activities within research, education and co-operation with the companies and public organizations, and covers the main areas of computer science. The activities are grouped in four research units:

Database and Programming Technologies has a particular focus on data management technologies; the design, implementation, and application of programming languages, environments, and tools; and technologies for location- and context-based mobile Internet services.

Distributed Systems and Semantics focuses on real-time, embedded systems and distributed systems; formalisms for the description and analysis of computer systems; and verification and validation methods and tools.

Information Systems researches development and use of computerised systems in organisations; human-computer interaction, development of software systems, as well as mobile and web-based systems.

Machine Intelligence is researching methods for reasoning, decision-making and learning under uncertainty, autonomous agents, intelligent web-based services, and data analysis.

The department has increased its external research funding in terms of both total amount and the number of grants. Grants come from several sources – from private foundations such as Villum Kann Rasmussen's foundation over the Danish National Research Councils to the European research programs. There has been an increase in the external funding since 2007, particularly since the activities of the Center for Embedded Software Systems were integrated. Since 2008, external funding has amounted to more than DKK 30 million per year.

The department's research groups are well recognized in each their fields and some even highly reputed. The department has a broad h-number profile with most above 10. Only two Danish computer scientists have h-numbers of 40 or more – and they are both employed at the department. The department also takes pride in an IEEE fellow, and an honorary doctor at ENS Cachan, France – all appointed because of outstanding research results.

Collaboration with the ICT industry is a key activity, and the department has been very successful in transforming research into business results. In 2008, Cassiopeia Innovation was established. Along with the existing organizations Center for Data-intensive Systems and Center for Embedded Software Systems, Cassiopeia Innovation serves as a general access point to researchers and a portal to more specialized third mission activities in the department.

The Department of Computer Science is well prepared for the future, and the strategic goals are: continued development of the research profile, adaption to a new economic regime, the Ph.D. education, student recruitment, external funding and third mission activities.

3.2 Background

Aalborg University is a fairly young university, established in 1974 and is characterized as a very dynamic university. In 2007, the Danish Building Research Institute became part of the university, and in 2010, the Faculty of Medicine was established as an independent faculty. At present Aalborg University consists of the following five faculties:

- The Faculty of Humanities
- The Faculty of Social Sciences
- The Faculty of Engineering and Science
- The Faculty of Medicine
- Danish Building Research Institute

The Department of Computer Science is part of the Faculty of Engineering and Science, which is divided into 13 departments. The Department of Computer Science is about average in size, as it receives approximately 1/13 of the Faculty of Engineering and Science's funding for the departments.

3.2.1 New management structure

In 2003, a new university law was adopted in Denmark changing the fundamental management structure of the Danish universities. Up until this point leaders (rector, deans, head of departments) at Danish universities had been elected by colleagues, students and technical and administrative staff. But the new university law abolished the elections of leaders and introduced a management structure resembling the structure of private companies. In addition, a board with a majority of external representatives was appointed. The board appoints the rector, the rector recruits the deans and the deans hire the heads of department.

The process of adapting to this new legislation was quite extensive. In May 2005, the new rector of Aalborg University was appointed by the new board, and in October 2005, the rector recruited the new dean of the Faculty of Engineering and Science. In the spring

of 2006, the new dean appointed his heads of department, and in March 2006, Kristian G. Olesen became Head of Department at The Department of Computer Science.

So at the beginning of the period of this evaluation, new leaders at all levels had just come into office and the implementation of the new leadership has had impacts up until now. Additionally, in the fall of 2010, a new dean of the Faculty of Engineering and Science was appointed. But since the change of dean occurred so late in the period of the evaluation, the impact of the new leadership might be difficult to identify at the departmental level at this point in time.

3.2.2 New funding structure

One of the larger changes at the department is the implementation of a new economical model.

Up until 2009, the department's resources consisted of positions, physical resources and funding from external sources. The number of positions, and thereby internal research funding, was primarily determined by the delivered activities (courses and project supervisions) requested by the study board.

In 2010, the Faculty of Engineering and Science introduced a new economic model, where resources are transferred to the departments as money. Funding for salaries is composed of three elements: Teaching (75%), publications (10%) and external funding (15%).

The new model provides the departments with some freedom in the management of resources and gives increased incentives to publication activity and to attract external funding when compared to the earlier model.

3.3 The development in teaching, funding and staff

3.3.1 The impact of student intake

As described above the development of the department has traditionally been closely correlated to the delivered activities in relation to the students. The number of students seeking computer science education was declining in the beginning of the research evaluation period and student recruitment was identified as one of the departments strategic focus areas.

In 2008, a full time communications consultant was hired and a number of initiatives were started. In Figure 3.1 student intake during the period is shown.
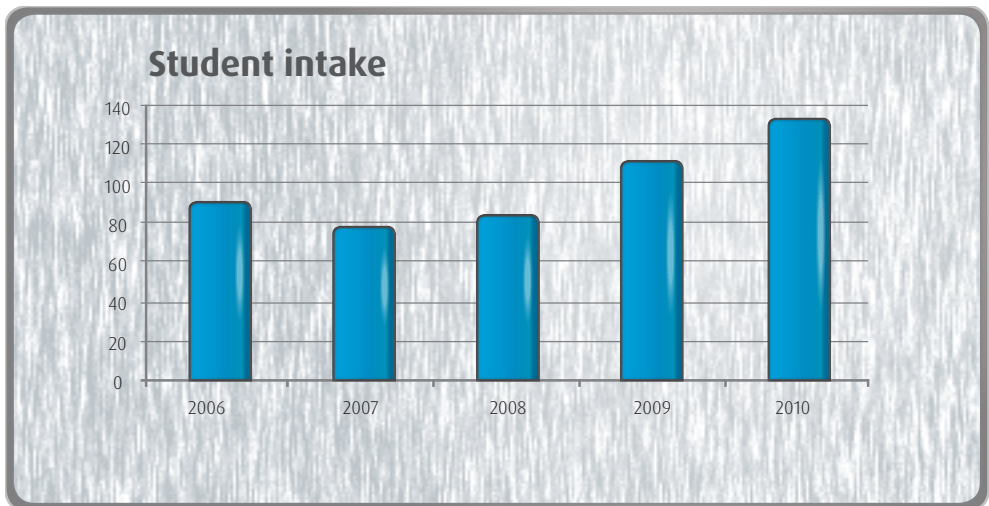
Figure 3.1: The development in student intake

As can be seen, the department has experienced a considerable increase in student intake with a 67% rise during the last 3 years.

This remarkable increase in student intake would be expected to have just as remarkable an influence on the requested teaching hours. But as shown in Figure 3.2 there has been

a decline in the teaching hours requested from the department throughout the overall period.
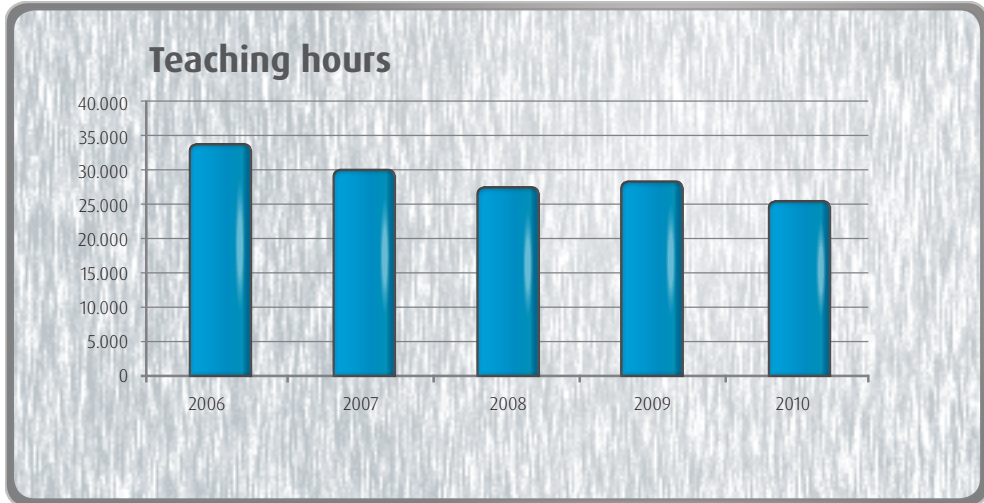
Teaching hours

Figure 3.2: The development in teaching hours

The decline in teaching hours in the overall period exceeds the decline in student numbers and does not reflect the increase in student intake in 2008 and onward. This is due to the fact that the appropriations granted for teaching has been declining.

This decline combined with the new economic model demonstrates that the teaching hours are no longer the most effective tool for generating economic growth at the department. On the other hand Figure 3.2 shows that if the department had not acted upon the decline in student intake in the beginning of the period, the numbers of students in 2010 could have induced a dramatic decline in teaching hours in 2010 and put the department's finances under strain since, as mentioned above, teaching comprises 75% of the salary budget.

3.3.2 Externally funded activities

Externally funded activities have historically been a very effective way of developing, expanding and consolidating research activities. Since the internal funding has been declining throughout the research evaluation period, the strategy has been to increase the external funding and thus evolve through externally funded activities. Figure 3.3 shows the external research grants spent in the period.
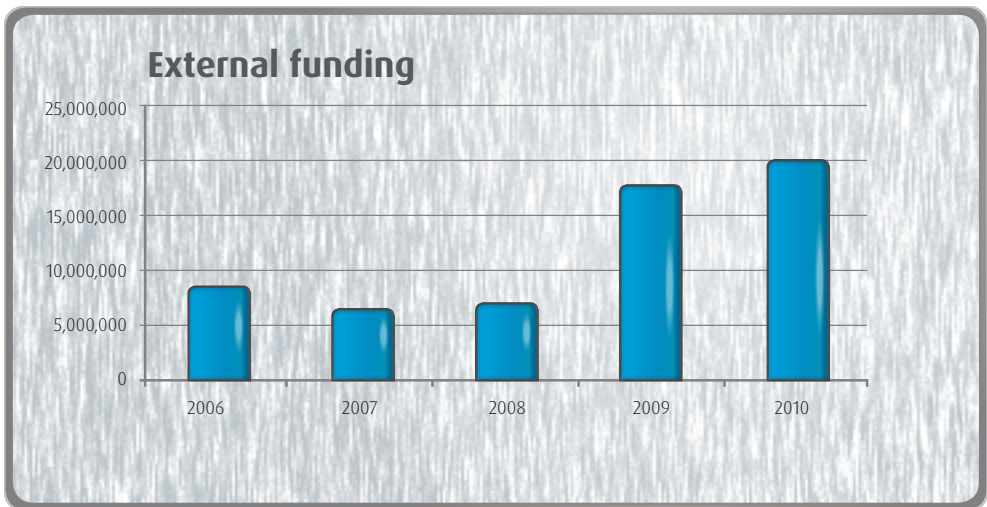
Figure 3.3: The development in external funding (spent)

As can be seen in Figure 3.3 there has been an increase in external funding since 2007 and in particular since 2008 when the activities of Center for Embedded Software Systems were integrated in the department.

Not only have the externally financed activities increased in terms of spent funding, external funding in terms of both total amount and number of grants has also increased during this period.
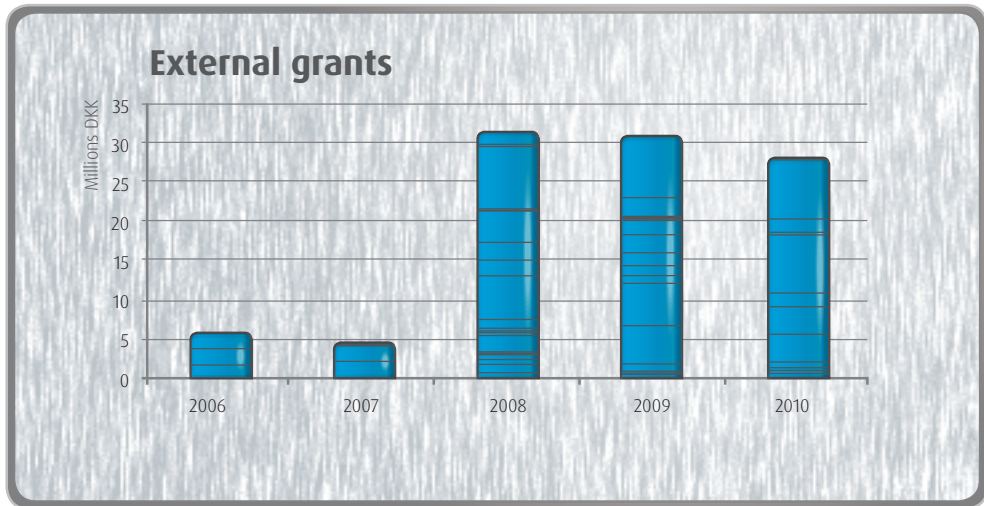
Figure 3.4: The development in external funding (granted)

The increase in external funding is not due solely to a more active approach to applying for external grants. During the period, the department has succeeded in obtaining a number of very prestigious and very substantial grants.

In 2008, the Villum Kann Rasmussen Foundation awarded a grant of DKK 25 million to MT-LAB – a VKR Centre of Excellence. The center has partners from Aalborg University, the Technical University of Denmark and the IT University of Copenhagen. The grant for Aalborg University amounts to DKK 10 million.

In 2010, the Danish National Research Foundation granted DKK 14 million to a Chinese-Danish Research Center for Foundations for Cyber-Physical Systems. The Danish partners are Aalborg University and the Technical University of Denmark, where Aalborg University is the Danish principal investigator. The grant for Aalborg University amounts to DKK 7 million.

The increase in external funding has been a very important financial tool to ward off a decrease in the overall funding of the department and has been crucial for sustaining the level of staff.

3.3.3 Staff

In the period 2006-2010, the department has been under pressure from declining internal funding as well as a temporary recruitment freeze. But in spite of these two obstacles the department has been expanding its staff during this period. This is mainly due to the increase in external funding, as mentioned above, where various projects have been able to compensate for declining internal income. Figure 3.5 shows the total number of staff in the department.

Development in staff
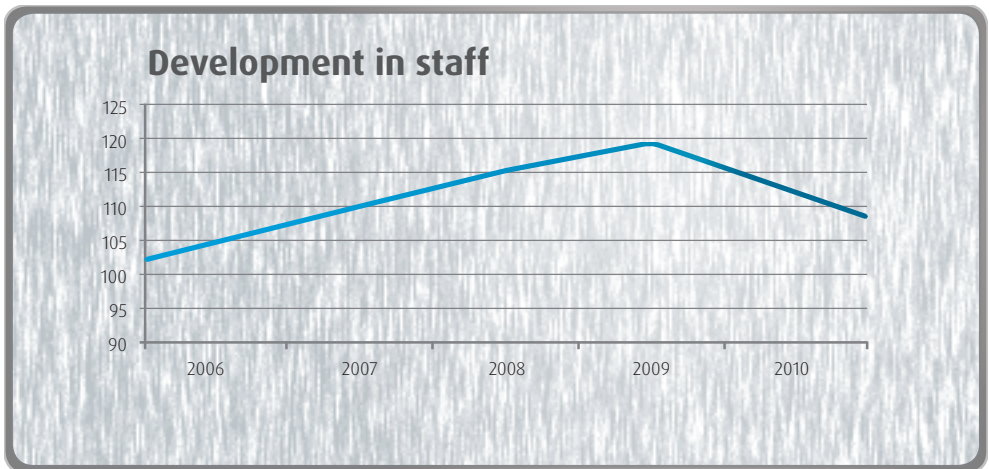
Figure 3.5: The development in the total staff

The drop in 2010 is partly due to a restructured IT organization, where some of the department's IT department was relocated to the faculty. Another reason, as shown in Figure 4.6, is a fall in assistant professors and other scientific staff.
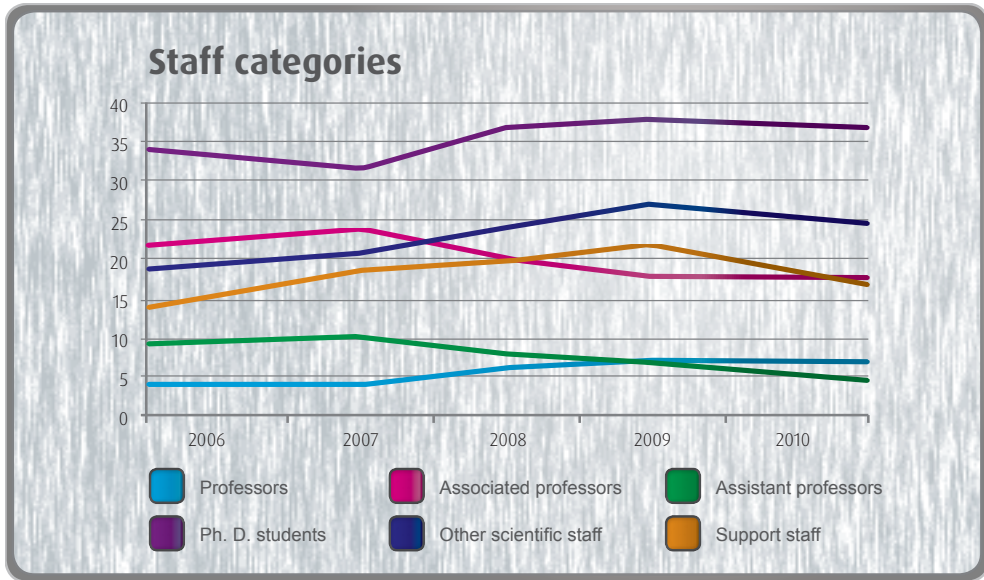
Figure 3.6: The development in staffing

Figure 3.6 shows the development in staff in more detail. The number of full professors has gone up, but as the professorships have been obtained by internal candidates it is counterbalanced by a decrease in associate professorships. Another reason for the decreasing number of associate and assistant professors is the mentioned recruitment freeze that was in effect in 2008 and 2009. We are currently getting back to a situation where we continuously refill assistant professorships as these temporary positions become vacant.

3.4 World renowned research

The research approach at the department is fundamentally constructive and embraces analytical mathematical research, experimental research with algorithms, systems, techniques and methodologies, as well as analytical empirical research.

The research covers software, use and performance of software, as well as information and data. In particular there is research on the use of software in organizations, software engineering, management of software engineering, human-computer interaction, programming and languages, data management, data analysis and data mining, techniques for decision support, machine learning, autonomous agents, networks and protocols, techniques and models for distributed, embedded and parallel software, and tests.

The department's research profile is reflected in a division of the department in four research units: Database and Programming Technologies (DPT), Distributed and Embedded Systems (DES), Information Systems (IS) and Machine Intelligence (MI). The scientific staff is partitioned in these four units.

Besides reflecting the department's research profile, the division into the four units reflects the decentralized responsibility for research, planning and teaching. In each unit the head of department designates a research leader and a unit leader.

The research leaders, typically full professors, formulate the research related goals and strategies within the unit's field. The group of research leaders forms an advisory board that meets with the head of department on a monthly basis.

The unit leaders take care of day-to-day management and with the head of the department form a management team of that meets on a weekly basis.

In addition to the four research units the department comprises supporting staff, which are divided in two. The administrative tasks are accomplished by the secretariat and the technical tasks are carried out in the IT group

The organizational chart of the department is shown in Figure 3.7.
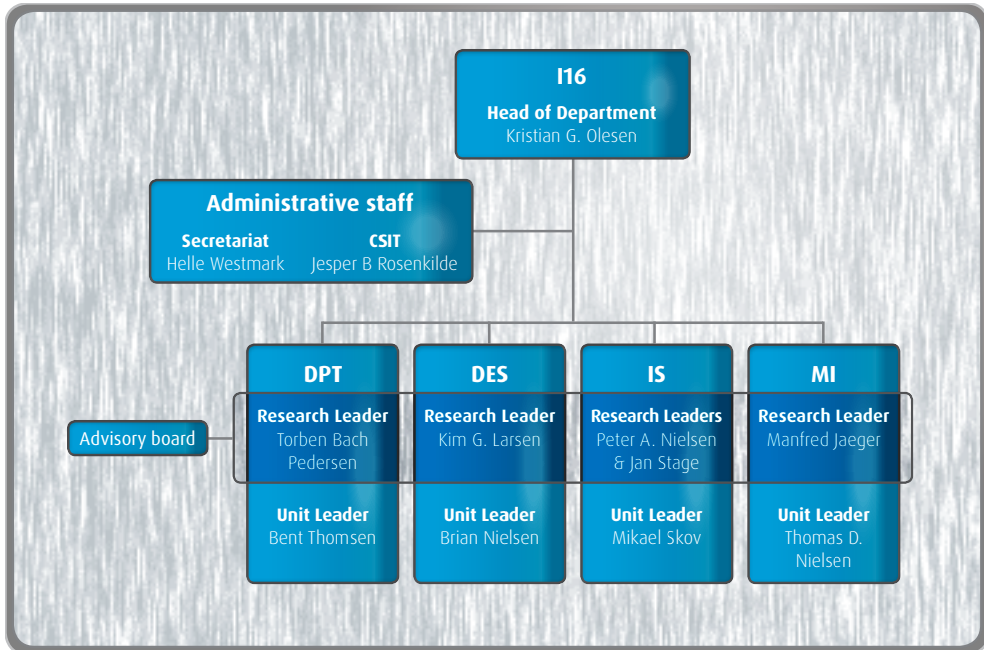
Figure 3.7: The organization of the Department of Computer Science

In 2007, the department moved to a new campus site. Prior to the move staff and students were situated in several buildings on campus, but by relocating to the new residence, which we named Cassiopeia, all staff and students in the department's programs are under the same roof.

The move supported the department's division into four research units and thereby the research unit forms the basic social environment for the individual researcher.

Cassiopeia is divided into five sections where the department occupies four of these sections. Each research unit is situated in their own section that also contains a spacious area for gathering and more social events.

Moving to Cassiopeia thus made it possible for the research units to be brought together and thereby to strengthen and develop their profiles.

3.4.1 Research driven development

The development of the department is primarily based on decentralized organic growth in the four research units. The research leaders are responsible for the evolving of the research areas and they assist the junior staff in the identification and development of their individual interests and profiles. New subjects and focus areas evolve from the interests of staff members.

The long term strategy of the department is a duality between this bottom-up strategy and a top-down strategy.

The top-down strategy is based on management decisions to pursue a specific subject and direct resources towards that area. It can both be on the unit level, e.g. agent technology within the Machine Intelligence unit and on the departmental level, where human-computer interaction was established years ago.

Some modifications have occurred in the naming and organization of the research units. Distributed Systems and Semantics changed its name to Distributed and Embedded Systems to emphasize this important area in the name. A new research unit focusing on the intelligent web grew up in the Information System research unit in 2006. The unit has collaborated with all units, and in August 2010, it was decided to relocate the unit in the Machine Intelligence research unit to reflect the direction of research and to seek synergy between the units. In continuation of the recommendations of the previous research evaluation the department has been granted new full professorships such that we now have at least one full professor in each research unit.

3.4.2 Internationally recognized research

For several years the research conducted at the department has had a very good reputation. The former research evaluation even characterized some of the research conducted at the department as "world leading". This is a reputation that the department, of course, takes great pride in and one that the researchers are constantly working on maintaining and evolving.

The department strategy on recruitment then is to announce all positions within a relatively broad spectrum of areas rather than to focus on specific subjects. All four research units are normally represented and involved in the advertisement of new positions in the department. Furthermore, all positions are announced internationally in order to recruit the most qualified applicants.

Announcements are usually based in the existing research areas, but with enough latitude to consider new and emergent areas that fit into the overall picture. In this way the department gains good insight into current trends and areas under development and organic growth is facilitated.

The department has succeeded in attracting highly qualified applicants from throughout the world and, as shown in Figure 3.8, has experienced an increase in international staff.
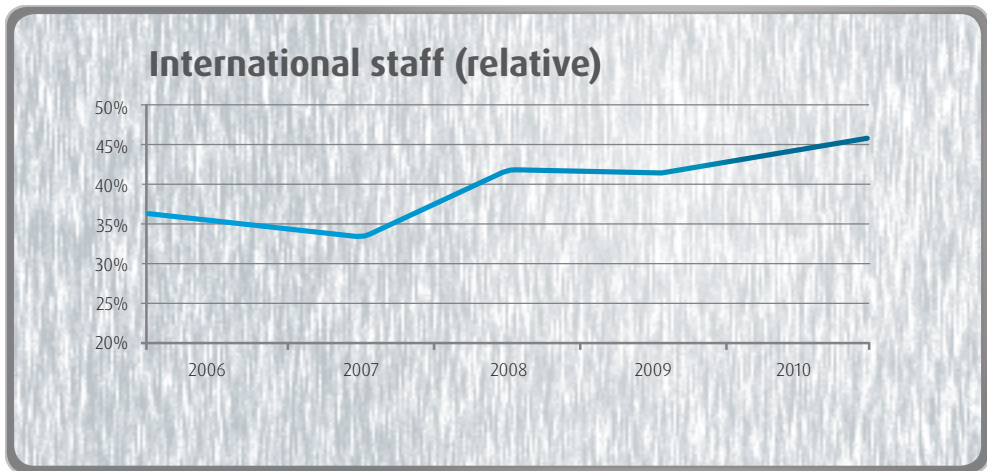
Figure 3.8: The relative development in international staff

The impact of the research conducted at the department has been recognized several times during the period, both inside and outside the researchers' circles. The h-number is a widespread indicator of scientific impact, and Jens Palsberg at UCLA maintains a list of computer science researchers who have an index of 40 or higher. Two Danish computer scientists are on that list and both, Christian S. Jensen, Professor, and Kim G.

Larsen, Professor, are affiliated with the department. The h-number profile of tenured staff is shown in Figure 3.9.
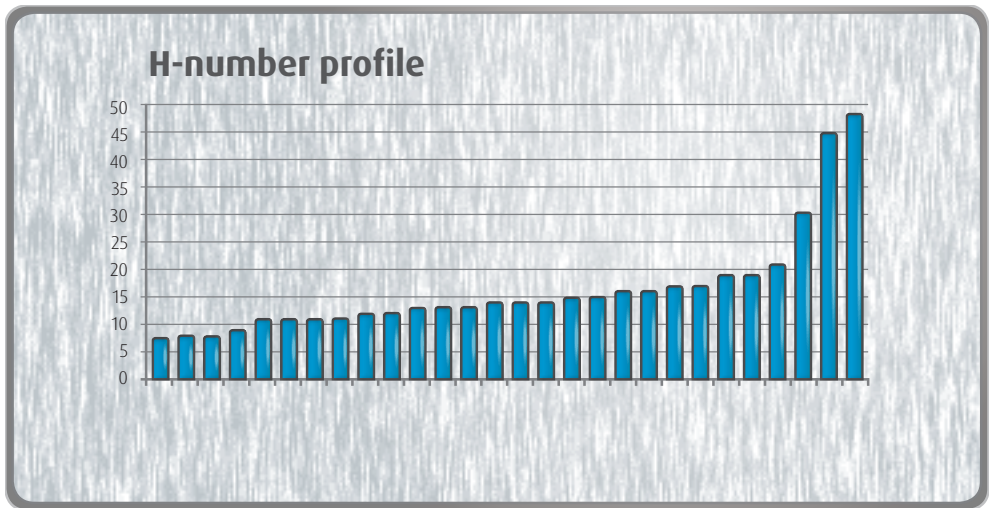
Figure 3.9: The h-number profile of tenured staff

In 2007, Kim G. Larsen was appointed as a member of the prestigious Danish Royal Order of Dannebrog, and later in 2007, he also became an honorary doctor – Doctor Honoris Causa – at Ecole Normale Superieure Cachan (ENS Cachan) in France. The honor was given for outstanding achievements in research and as recognition of his exceptional effort to foster collaboration and international breadth in the field of embedded systems.

In 2008, Christian S. Jensen became IEEE Fellow for his outstanding contributions to temporal, spatio-temporal, and mobile data management.

In 2010, Finn Verner Jensen, Professor, joined the members of the Danish Royal Order of Dannebrog and he received Nordjysk Universitetsfonds Innovation Award in recognition of his significant role in establishing Probabilistic Graphical Models as an independent research area.

3.4.3 Research interaction

Researchers from the department contribute to the international community both through publications and through extensive collaboration. All units have activities internationally, nationally and locally.

International enterprises range from individual co-authorships to larger projects including several funded by the EU to a Danish-Chinese basic research center.

At the national level examples include a joint Ph.D. school within the Information Systems area incorporating most Danish universities, and the MT-LAB, funded by the Villum Kann Rasmussen foundation, that was established in collaboration with the Technical University of Denmark and the IT University in Copenhagen.

Locally there are several contacts with other departments in particular the Department of Electronic Systems, but also with e.g. the Department of Development and Planning. As will be detailed in the following chapters, there are also extensive cross-unit collaborations ranging from joint publications to larger activities such as the dissemination project InfinIT.

3.4.4 Research-based elite education

In the spring of 2009, the Danish Ministry of Science, Technology and Innovation designated a few areas at Aalborg University as nationally and internationally outstanding in terms of research. One of the areas appointed was Embedded Software at the Department of Computer Science allowing the department to offer an elite program for particular skilled students. The elite program is a 2-year full- time research-based Master's program. The elite program at the Department is rooted in the research unit, Distributed and Embedded Software and the students admitted to the program work with semantic modeling theories, design, implementation and modeling for distributed systems and network. Furthermore, they work with algorithms and methods regarding verification, validation and performance analysis of programs and systems. The students are part of a

highly active research environment with regard to regional and national companies along with international research units. At present 17 students are enrolled at the elite program in Embedded Software.

The Department of Computer Science has just had one more elite program within Data Intensive Systems approved by the faculty. The students will be given the opportunity to enroll in this program in autumn 2011.

3.4.5 Research collaboration with industry

A strong collaboration with the ICT industry has traditionally been one of the trademarks of the Department of Computer Science.

NouHauz was the first large initiative in the department to direct attention towards industry collaboration. In 2003, the unit of Distributed Systems and Semantics initiated the Centre for Embedded Software Systems (CISS) and became strongly involved with a large number of industrial partners with focus on collaboration within the area of software construction and embedded systems.

In 2006, the unit of Database and Programming Technologies established the Center for Data-intensive Systems (Daisy). Like CISS, Daisy became strongly involved in industry collaboration and the department now has a powerful platform for industry collaboration.

In order to consolidate and strengthen the profile on collaboration with the ICT industry, Cassiopeia Innovation (CaIn) was established in 2008 as the department's joint third mission section. The early initiative NouHauz had proven viable and Cassiopeia Innovation was established as an extension of the original ideas. The purpose of CaIn was to stimulate the application and exploitation of ICT knowledge for social and economic development in collaboration with external partners. CaIn serves as a general access point to researchers in the department and as a portal to more specialized third mission activities in the department.

CaIn supports industrial product and process development, innovation, and competence building based on research. In return, the department and the researchers gain a broader and stronger platform for obtaining external funding, opportunities to work on real world challenges and complexities, and more ways to utilize research.

3.5 Future challenges

The department is well-situated to meet future challenges. We have been able to double student intake through a dedicated effort over the last three years, and we have been able to increase external funding in terms of both total amount and the number of grants. After a period of consolidation, we are now prepared for further expansion and ready to engage in new enterprises.

The current focus areas are continued development of the research profile, adaption to the new economic regime and the Ph.D. area. These issues are interconnected and dependent on a solid economic base. The continued emphasis on student recruitment is fundamental to ensuring basic research funds, but it has to be supplemented with other research funding. A particular challenge is to maintain the high input from external sources, and continuously attract new funding as existing grants expire in order to provide room for renewal and expansion.

New senior staff supports the organic growth strategy, and an increased number of Ph.D. students contribute not only to exploring existing research directions, but also leaves space for new areas to be developed by particularly talented students.

Particular emphasis is directed towards the Ph.D. area. There is focus on the intake of new students, and an increase in the number of Ph.D.'s implies challenges in funding and recruitment. Even if the funding issue is resolved, there is still the issue of attracting good students that are actually able to complete their theses on time.

An increasing part of the public funding of universities is going to be based on production, measured in terms of publications and Ph.D. degrees. A national bibliography has been

established for the former and a metric based on a ranking of journals and publishers has been constructed. It is deemed essential that this bibliometric system will include conference articles, as this area is of particular importance for computer science. The new system must recognize the computer science tradition for peer reviewed full articles in conference proceedings in order to secure a true and accurate picture.

3.6 Committee evaluation

It is inappropriate to make one generalised statement of research quality across the department, because there are significant differences between groups and individuals. There are some international islands of excellence while other work is of a good national standard. The focus in reporting was on excellence of science, but there is also clearly a (less explicitly measured) focus on knowledge transfer (KT, also referred to as dissemination and economic impact). The panel believe that there should additionally be explicit consideration of the development of human capital, through training and staff development.

3.6.1 Strengths, weaknesses and vulnerabilities

The department has many commendable strengths overall, although many of these have corresponding weaknesses or vulnerabilities, which are most easily expressed in tabular form.

Theme	Strengths	Weaknesses and vulnerabilities
Quality of science	The department has focused attention on the quality of publication outlets and bibliometrics.	There are other measures (both qualitative and quantitative) that need to be taken into account to improve resilience and build even stronger foundations for further excellent research.
Developing human capital	The department appears to have a good strategy for developing staff within the constraints of the funding model. The problem-based learning approach involves students in research.	While attention is paid to developing junior staff, less attention is paid to developing middle-ranking staff (e.g. in writing research proposals and managing research), and there needs to be more explicit succession planning to cover departure or sabbaticals for the research leaders. There is the possibility to take more advantage of student projects to support research.
Knowledge transfer	There are excellent links with local, national and international industry. There have been many spin-out companies.	In some cases, a focus on applicable work might conflict with the delivery of excellent (novel) science of international significance. There is a need to measure and articulate the impact of KT and spin-out activities to better assess achievements.
International esteem	There are some (2.2 FTE) leading research stars in the department, and several other members with significant international research presence.	The department is currently vulnerable to the loss of just a few key individuals.
Collaboration	The panel were impressed by the efforts which have been made over the current review period to increase collaboration within and between research groups, as well as with other departments and outside organisations.	There are further opportunities for fruitful collaboration, though collaboration should be regarded as a means to an end (better science, staff development, improved KT) rather than an end in itself.
Space	The department is endowed with excellent research and teaching space, which it uses well.	In terms of space, there may be a challenge in overcoming physical distance between research groups and ensuring ongoing informal communications at all levels.

3.6.2 Recommendations

To further strengthen the research quality and impact of the department, the panel has the following recommendations:

Measuring the quality of the science: the approach with focus on bibliometrics is one of several measures. Consideration should also be given more explicitly to other quantitative measures such as citation counts and H-indices. There should also be a greater focus on articulating the research contributions. For example, in the next quinquennial report, each research groups should identify its 8-10 leading research outputs and explain for each why it is significant; this might be because of the scientific contribution, or because of the KT impact, or for another reason, corresponding to the aims of the department.

Measuring KT / economic impact: there needs to be a clearer account of the actual impact of engagement with industry, beyond a list of the companies that are collaborating or spun out.

Facilitating collaboration: The department needs to remain alert to the risks of silos and continue to create mechanisms to break them down (e.g. away days and joint PhD activities).

Developing human capital: The approach of reporting achievements at group level can hide wide variances in contributions and achievements, and hence hide problems that ought to be addressed. There should be greater transparency about the achievements and roles of individuals so as to more readily celebrate success, motivate individuals and identify and address problems. The department should develop a stronger mentoring approach for mid-ranking staff to develop research skills, particularly proposal writing, to improve succession planning. It should also develop a cross-group review procedure to ensure that the proposals that are submitted are as competitive as possible, and to share best practice.

Group staffing: It will be better to have clear staffing resources per group based on need, to allow groups to plan more strategically.

Publication outlets: Continue to aim for selectivity while also recognising that lower quality outlets can have great value for learning (e.g. feedback to PhD students) and dissemination (e.g. through specialist workshops).

Beyond publications...: Within Computer Science, software products are outputs of research. There are four types of software tools that can be developed:

1. Prototype. Typically, a PhD student will develop during his studies a prototype tool to implement the algorithms he develops and show some academic small examples of problems he can solve with this software. No special considerations have to be developed for this software, as it probably will not be used by others, and the thesis and publication will document the software sufficiently.
2. Professional tool development. These tools require a lot more effort, not just by researchers, and will require continued attention. There are three user groups:
   a. Colleague researchers at other universities.
   b. Students within the department.
   c. Applications in an industrial setting.
This type of software requires professional development and maintenance.

3. Commercial tool development. When companies want to use tools under 2, they require good support for such a tool, and guarantees for permanence and maintenance, and a good help function. It takes non-academic manpower to provide this. A spin-off or spin-out company may take this up.
4. Experimental tool development. Students of software engineering want to experiment with existing mature tools, to see if they can improve things, or change things.
If a CS department wants to have tools like 2 or 3, then this should be managed well. Such an infrastructure is missing here, and it is left to the groups to manage this. Things that should change:

1. Software should count as output of a CS group, just as publications and patents do. If a staff member contributes to such software development, this should be taken into account, by for instance lower teaching obligations, or lower publication requirements.

2. CS is an engineering department just like other engineering departments, and should have a full-fledged lab function. Technicians can be employed that do software development as their main task. Function development, adequate salary scales, should exist.

4 Database and Programming Technologies
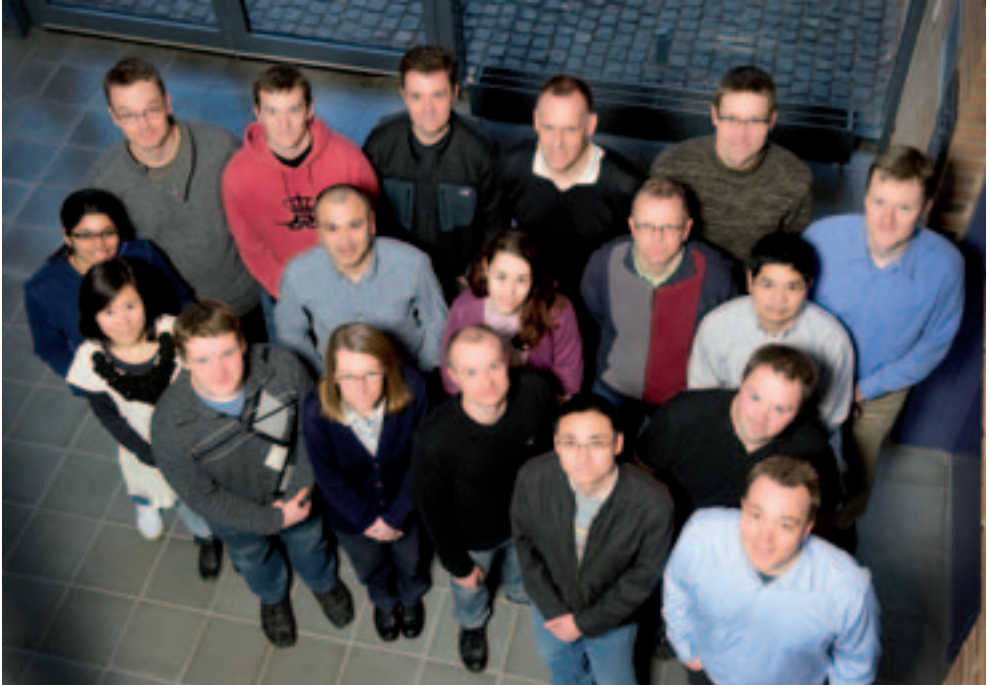
The Database and Programming Technologies Unit.

4.1 Executive Summary

The Database and Programming Technologies unit, also known as the *Center for Data-Intensive Systems (Daisy),* performs research within spatio-temporal data management, mobile services and indoor positioning, data warehousing and business intelligence, multimedia data management, data management infrastructure, and programming languages, concepts, and environments.

The goal of Daisy is to be an internationally leading research center within its area. In the period, Daisy believes to have performed at a high level on the key performance indicators of publications, external funding, and Ph.D. education.

Daisy researchers have published 235 peer-reviewed papers in the period, including 82 in the top conferences and journals within Daisy's area.

The unit has been involved in 27 externally funded projects during the period. These projects correspond to around 37 million DKK in grants from a wide range of sources. Most projects have substantial industrial participation, and many involve other units at the department or the university.

The unit has also been very active in educating new researchers, with 12 Ph.D.s graduated in the period. Currently, another 17 Ph.D. projects are in progress.

Additionally, Daisy has been successful in attracting top candidates for its positions. Daisy researchers have performed numerous invited talks and services to the research community. Furthermore, Daisy has recently been granted an elite masters education.

In conclusion, the unit believes it has been very successful research-wise in the reporting period. In particular, the unit believes that its publication record compares favorably with that of similar units across Europe.

4.2 Profile

The research in the database and programming technologies unit covers a wide range of areas within data management and programming.

The area of programming technology includes general-purpose programming languages as well as special-purpose languages, e.g., for embedded, mobile, data intensive, web based and high-performance applications. Programming paradigms and specific language facilities and concepts, such as integration of functional and object oriented pro-

gramming, are covered. Also studied are environments that offer integrated tool support for application and program development, covering the activities of design, implementation, documentation, test, and maintenance/operation.

Within data warehousing and business intelligence, the research covers data warehousing of spatio-temporal, multi-granular, and web data, near-real-time data warehousing and data streams, scalable and distributed DWs, open source business intelligence tools, and mining of data cubes and spatio-temporal data.

The research in multimedia data management covers data management and information retrieval for music data, similarity search and subspace clustering for high-dimensional multimedia data (time series, images), text mining (question-answer systems, categorization, influence), and social network mining.

In the area of spatio-temporal data management, the research covers a wide variety of topics including different aspects of spatio-temporal indexing, query processing algorithms, and location-privacy. Data modeling, indexing, and query processing is explored in the context of spatial-networks and in indoor settings.

The unit has built infrastructure enabling location-based mobile services. These include indoor Wi-Fi-based positioning; indoor moving object management based on RFID-type positioning; seamless indoor-outdoor positioning; location privacy techniques; and techniques for outdoor location-related spatio-textual queries.

Finally, in the area of data-management infrastructure, research is done in query optimization for moving objects as well as performance and correctness test of database applications.

4.3 Staff

4.3.1 Current Staff

Professors: Torben Bach Pedersen, Christian S. Jensen (part time)

Associate professors: Hua Lu, Kurt Nørmark, Simonas Šaltenis, Bent Thomsen, Lone Leth Thomsen, Kristian Torp.

Assistant professors: Ralf Rantzau, Ira Assent (on leave).

Post-docs: Christian Thomsen.

Ph.d-students and research assistants: Ove Andersen, Laurynas Biveinis, Rene Hansen, Nadeem Iftikhar, Dalia Kaulakienė, Simon Kongshøj, Saulius Samulevičius, Darius Šidlauskas, Laurynas Šikšnys, Kostas Tzoumas, Carmen Ruiz Vincente, Dinming Wu, Xiufeng Liu.

Research assistants: Ove Andersen, Rene Hansen.

4.3.2 Staff Development

With the core of the unit's staff remaining stable, the staffing of the unit evolved substantially during the period.

Among the permanent staff, Torben Bach Pedersen became a full professor in 2008. Professor Christian S. Jensen has taken up a professorship at Aarhus University in 2010, and will only be part-time affiliated with the unit from now on. Finally, Hua Lu who started in the unit as an assistant professor in 2007 was hired as an associate professor in 2010.

A number of assistant professors who were employed at the beginning of the period left the unit when their positions expired. They include Linas Bukauskas who left in 2006 to Vilnius University, Albrecht R. Schmidt who left in 2007 to European Space Agency, Janne Skyt who left in 2007 and Xuegang Huang who left in 2008, the latter two taking up jobs in Danish companies.

Additionally, a number of assistant professors and post-docs joined the unit and left it during the period. Man Lung Yiu was employed in 2006 and left in 2009 to Hong Kong Polytechnic University. Gao Cong joined the unit in 2008 and left in 2010 to Nanyang Technological University, Singapore. Bee Yong Chua was employed from 2007 to 2008. Ira Assent joined in 2008 and is currently on leave.

A large number of Ph.D. students and research assistants pursuing Ph.D. degrees were employed during the period, but are no longer with the unit: Agnė Brilingaitė, Xin Cao, Kim Christensen, Per Madsen, Stardas Pakalnis, Maria Magdalena Ruxanda, Laurynas Speičys, Dalia Tiešytė, Rico Wind, Bin Yang, and Xuepeng Yin. In addition, a number of individuals were employed for short periods of time as research assistants.

A number of Ph.D. students were supervised by the unit while being employed elsewhere, including Thomas Bøgholm, Gyözö Gidófalvi, Anders Henriksen, Stephan Korsholm, Morten Goodwin Olsen, Morten Middelfart, and Nerius Tradišauskas.

4.4 Goals (2006-2010)

In the following, we summarize the research goals for the evaluation period. The detailed plan may be found in the report on the previous evaluation period.

Focusing on quality and significance, the unit planned to continue following the publication strategy of publishing the best research results in the best general journals and conferences and other results in the best specialized journals and conferences. To further this goal unit ranking of top general and specialized publication outlets was devised.

To strengthen the focus of external industrial collaborations, the unit planned to aim for fewer, more substantial, and more focused collaborations, as well as to prioritize larger and longer funded projects with funding for one or more Ph.D. students on topics that are within the core competences of the unit.

Overall, the unit planned to engage actively Ph.D. and M.Sc. students in its research activities. To guide research it was considered important to perform *use inspired* research, through contacts with practitioners. At the same time the goal was to achieve results that have *general applicability*. Finally, the research was planned to be *constructive* and *experimental*. Patenting and commercialization was to be explored.

Balancing between the need for focus on core research competences of the unit and the need for diversity within the unit, three focus areas were defined.

The **P2025 (**″**How would you like to program in 2025**″**)** project was formulated with the goal to influence the next generation of mainstream programming languages. Ease of using and learning as well as the richness of a programming language were defined as key guiding goals of the project. The project was planned to involve analytical and empirical studies involving design and implementation of languages, compilers and experimental systems.

Data Management and Programming Foundations for Mobile Services focus area was defined with the goal to contribute substantially to the invention of new data management and programming foundations for software systems that will enable the delivery of mobile services in an emerging infrastructure of vast quantities of interconnected computing and sensory devices. Special focus was to be given to geo-location aspects in such services. This focus area had the goal to serve as a framework for a range of more specific activities.

Querying the World focus area was motivated by the increasing amounts of different kinds of sensors populating the world and emitting large amounts of data, along with ever more data accumulated in traditional databases. The vision was formulated that the data from the sensors (present state), databases (past state), and prediction models and simulations (future state) was made available through a unified infrastructure and could be queried seamlessly.

4.5 Activities and Results

In this section, first, main research focus areas of the unit are described, followed by descriptions of the key research projects during the evaluation period.

Programming Technology is an essential part of any software development, and during the evaluation period the focus area of programming technology has been rather active with more than 30 papers published and involvement in a number of projects, thus raising the profile of the area in accordance with the recommendations of the previous research evaluation panel. The research has both been long term, medium term and applied. The focus on long term research has been on integration of object oriented programming and declarative programming and looking into the future of closer links between dynamic programming and statically typed languages. The focus on medium term has been on enabling real-time programming in Java, as well as interactive collection of tests during the programming process and the focus on applied research has been on mobile applications and indoor positioning, as well as on educational issues.

Spatio-temporal data management has been a key research topic throughout the evaluation period, with some 80-100 papers having been written on the topic. The research conducted within this area is generally motivated by the increasing abundance of spatio-temporal data, often in the form of GPS data obtained from a variety of so-called moving objects, as well as the increasing mobile use of services and the Internet.

Specific topics include: (i) support for workloads with very frequent updates, (ii) tracking of moving objects, (iii) spatio-textual search, (iv) ranking of so-called spatial web objects, (v) query processing in indoor settings, (vi) modeling, indexing, and query processing proposals for spatial-network constrained objects, and (vii) location privacy. Particular attention has been given to indexing techniques. Thus, several new indexing techniques have been proposed, and benchmarking of indexes has been pursued.

Mobile services/indoor. The unit has engaged in a range of research activities that target technologies for outdoor as well as indoor mobile services. The Streamspin project that focuses on enabling user-generated location-based services has served as prototyp-

ing infrastructure for several of these activities, including outdoor, GPS-based tracking of mobile objects, Wi-Fi-based indoor positioning, as well as seamless indoor-outdoor positioning.

In addition the unit has conducted studies in an indoor setting with RFID-type positioning. Proposals encompass the following: indoor space modeling infrastructure, moving object tracking, indexing techniques, and query processing.

Further, the unit has been active in the area of location-related privacy. Relevant contributions apply to a variety of location-based services and geo-social networks.

Finally, a range of contributions target so-called spatial web objects. The unit has developed techniques that identify and rank such objects from GPS data collections, as well as those that enable the effective and efficient spatio-textual querying of them.

In relation to the focus area of "Data Management and Programming Foundations for Mobile Services" formulated in the plans for this evaluation period, the unit has been quite successful in pursuing use-inspired research, particularly with new substantial research initiatives in the area of mobile services, as well as work on user-location privacy and work on moving object management for indoor settings.

Data warehousing and business intelligence was another major research topic in the period, resulting in some 60-70 papers. The research is motivated by advanced real-world business intelligence applications. Specific topics include (i) powerful code-based ETL frameworks; (ii) near-real-time data warehousing and analysis of data streams; (iii) warehousing multi-granular data; (iv) warehousing of web-related data like XML, RDF/OWL/Semantic Web data, and text-rich web documents; (v) open source business intelligence tools; (vi) highly scalable and distributed data warehousing; (vii) bitmap indexing; (viii) spatio-temporal data warehousing; (ix) spatio-temporal data mining; and (x) mining of so-called sentinels that capture cause-effect relationship between cube measures.

Multimedia data management was also a major topic, resulting in some 60-70 papers. The research is focused on providing effective and scalable techniques for different types

of multimedia data. Specific topics include (i) modeling of music databases; (ii) querying frameworks and playlist generation for music databases; (iii) effective indexing techniques for large music databases; (iv) effective query processing techniques for large music databases; (v) similarity search for high-dimensional multimedia data such as time series and images; (vi) data mining, especially subspace clustering, for high-dimensional multimedia data; (vii) text mining including categorization, and influence; (viii) question-answering systems; and (ix) social network mining.

Looking back at the "Querying the World" plans, the research concerned sensor data management less than expected, mainly due to the rejection of several large grant applications. The past-present-future integration is pursued in the MIRABEL project. Fundraising success considerably exceeds the plan. A patent application was submitted, a startup company is on the way, and key results such as sentinel mining is commercialized through collaborator companies such as TARGIT.

Data management infrastructure was a smaller research topic within the unit during the evaluation period, with 5-8 papers being written on the topic. The research is motivated by the constant increase in the data sizes stored and the end-users expectations of fast and correct answers to queries. This topic is not directly mentioned in the research plans for 2006-2010. However, it has been a natural consequence of the unit's focus on experimental validation of research results and the extensive collaboration with the IT industry, most recently in the DaIn project.

In the following we highlight a number of *key projects* that the unit has undertaken during the evaluation period.

Key project: P2025 - How would you like to program in 2025? The project was initiated in December 2004 as an umbrella project to coordinate activities otherwise carried out on an individual basis, bringing focus to our research, to the supervision of master and Ph.D. students, and to funding applications. The unit has also hosted a number of visitors from academia and industry, both nationally and internationally. Our main contributions have been three research prototypes; an F#-based prototype of the initial parts of our Abstraction Step Language (ASL), a full Scheme implementation of ASL, and a framework

for performing interactive unit testing and harvesting examples for API documentation in Scheme - all documented in a number of publications. The participants of P2025 have supervised 50 master students and 5 Ph.D. students, and organized two PhD. courses. We applied for external funding in five applications. Firstly UNIK, a large departmental research project proposal, which did not succeed. Secondly, SmartCampusAAU funded with 2M Dkr. from Væsktforum, Region Nordjylland. The programming technology part of the UNIK application was extracted and submitted to the Danish Research Council (FNU) about End User Programming. Despite receiving very good comments, it was rejected. A repositioned application to the Research Council (FTP) was also rejected. Finally, an application 'Smart Work - The Way Ahead?' (about educational aids for unit work and program development) is pending. The unit has also participated in several funded activities, such as DaIn, CaIn and InfinIT. The P2025 project ends with the expiration of this research evaluation period.

Key project: Intelligent Sound - Search and Retrieval in Dynamic Sound Databases. Intelligent Sound project was a large interdisciplinary project sponsored by the Danish Research Council for Technology and Production. It involved collaboration between signal processing researchers at the Technical University of Denmark and AAU, and Daisy data management researchers. The key project aim was to support advanced and efficient queries on music data, enabled by automatically extracting music metadata (melody, rhythm, timbre, etc.) directly from the sound signal by advanced signal processing techniques. The extracted feature vectors were then used by the modeling and querying frameworks developed by Daisy, enabling advanced queries such as playlist generation, etc. Finally, a number of effective indexing and query processing techniques were developed, enabling efficient querying on even very large music databases.

Key project: Daisy Innovation. The Daisy Innovation (DaIn) project is a large on-going project funded by the European Regional Development Fund, aiming to make Daisy's leading competences within data-intensive systems available to companies in the North Jutland Region. This is done through a number of education activities, technical advice for companies, networks for knowledge exchange, and smaller, focused sub-projects on specific topics. The project has four key partners: TARGIT, Bektra, Aveva DK, and Dafolo, with more to come. The collaboration topics cover the full spectrum of Daisy, including

object-oriented programming, application development, business intelligence, data mining (where results are already marketed by TARGIT), database testing, mobile services, and intelligent transport systems. The focus of the project is as much on innovation as it is on research, yet several papers have already resulted from the collaboration, and many more are expected.

Key project: Streamspin. **Enabling user-generated mobile services.** A shift in how we access information is happening: the Internet is increasingly being accessed from smartphones. This led Google to its Mobile First rule: ``When we announce new services for desktop computers, [...] we will debut an equally powerful mobile version. We will take advantage of this new class of smartphone, which is more sensory, acts as an extension of you, is aware of location, [...] .'' (*Vic Gundotra. Barcelona: Mobile First. Google Mobile Blog. February 25, 2010*)

The Streamspin project invents technologies that enable websites that may serve the same role for mobile services as Youtube serves for video. Focus is on inventing and prototyping technologies that enable websites that support the creation, sharing, and deployment of user-generated mobile services, with emphasis on geospatial services.

For example, a service could alert its users when they are near a friend. Another could enable a user to leave virtual Post-its, photos, or videos for friends to discover as they get near them. Yet another service could let users publish their vacation travel experiences in real time for family, friends, and the general public to react to, by continually uploading georeferenced content. The possibilities are numerous.

The Streamspin system: (i) enables user-generated services by enabling programmers to create service templates from which non-programmers can create services, (ii) offers support for basic aspects of services such as authentication, security, and privacy as well as the ability to flexibly push content to users, (iii) enables tracking of users with varying accuracy, (iv) enables service sharing, and (v) enables the scalable delivery of services.

4.6 Own Evaluation

Internal collaboration

Within the unit, but across the subgroups, the Streamspin platform has worked as a cata-lyst that have enabled people from different subgroups to both contribute to, and take advantage of, the wide range of functionality in the platform. A prominent example of this collaboration is the joint work on indoor positioning. This collaboration has resulted in externally funded projects such as SmartCampusAAU and (to some extent) Daisy In-novation. Very recently a collaboration on using the power of GPUs for speeding up com-putationally intensive queries has been initiated. The collaboration involves a start-up company, ConfiCore, initiated in DaIn and now hosted by greenHauz/CaIn.

There has been collaboration with other research units in the department. Examples include work on real-time programming in Java with DES/CISS, web transactions and compensations in high performance computing with MI, web engineering and Problem Based Learning with IS/MI, IT in farming and Intelligent buildings with DES/CISS. The context was typically externally funded projects such as Java Objekt, InfinIT and LandIT.

At university level, there has been increasing collaboration with units at other depart-ments, including signal processing, traffic research, statistics, and industrial production. The context has been externally funded projects, e.g. SmartCampusAAU and Intelligent Sound. There has also been involvement in writing large scale cross university research proposals, such as UNIK and SPIR.

The unit finds these collaborations very satisfactory and plans to retain this. Especially the internal collaboration between the subgroups, collaboration within the department, and at the university level has increased significantly.

External collaboration

Over the period, the unit has collaborated with more than 30 companies and organiza-tions. This level of collaboration is very satisfactory, and perhaps even on the high side volume-wise compared to the unit's resources. The unit should aim to have larger, more targeted, and more research-focused collaborations in the future.

The unit maintains close research collaboration with top universities and industrial/government research labs. Key examples include Google, National University of Singapore, and Cambridge University. Members of the unit actively take part in the EU COST actions and the EU Erasmus exchange programme. The unit finds this level very satisfactory and plans to retain it.

Scientific output

Similar to the previous evaluation period, the publication strategy of the unit continued to be focused on promoting quality and visibility of research. Thus the unit strives to publish its best research results in the best general journals and conferences within the general scientific communities covered by the unit (mainly data management and programming languages). Other results are published in the best specialized journals and conferences within specialized subareas, e.g., spatio-temporal data management, mobile services, data warehousing, data mining, and embedded or functional programming. The lowest priority is given to publication in general outlets with little quality control and consequent relatively low quality. Such a publication strategy increases impact and visibility of research and promotes a high peer reputation for the unit and its staff. In the unit's own ranking of publication outlets, tier A general journals include TODS, TOPLAS, VLDBJ, TKDE, and PVLDB. Tier B general journals include IS, DKE, CLS&S, SP&E, IEEE SW, and IEEE Computer. Tier A specialized journals include GeoInformatica, TGIS, and DSS. Tier B specialized journals include ITS, JOT, HOSC, JDWM, and JMC. Tier A general conferences include SIGMOD, VLDB, ICDE, EDBT, CIKM, SIGKDD, SIGIR, ICDM, WWW, OOPSLA, HOPL, POPL, PLDI, and ECOOP. Tier B general conferences include DASFAA, DEXA, ADBIS, and IDEAS. Tier A specialized conferences include SSTD, ACM SIGSPATIAL, MDM, SSDBM, TIME, SOCC, ICDCS, ER, SDM, ECML/PKDD, PAKDD, DaWaK, ISMIR, MMM, ICME, OOPSLA/SPLASH, ICFP, and SAS. Tier B specialized conferences include MobiQuitous, MobiDE, W2GIS, STDBM, MobiLight, PerTrans, ITS, STDM, DOLAP, StreamKDD, BIRTE, BIMA, ICMB, JTRES, ISORC, ELW, Scheme&FP, LCTES, BYTECODE, and WCET.

We observe that the top-tier outlets are the most selective. In particular, publication in tier A general conferences is highly competitive. Papers are typically 10 to 12 pages long (2-column, 9 or 10 point type) and most accepted papers report on experimental studies that require substantial software development. The acceptance rates for the first tiers of general conferences and for some of the specialized conferences are below

those of the best journals. In particular, the acceptance rates at tier A general conferences are below 20%. A special note is in order regarding the PVLDB journal, which publishes papers written in conference style with a rapid turnaround review process. The papers are presented in the associated VLDB conference. We include VLDB among the conferences to account for our publishing there prior to the establishment of PVLDB in 2008.

There is a very good match between the unit's top-ranked journals and the top level in the Danish national bibliometric ranking. In particular all tier A journals are at the top level. However, it is essential that top conferences are treated equally to journals in the national ranking. The national ranking committee for computer science, headed by Torben Bach Pedersen, is currently lobbying for this to happen.

During the evaluation period 235 peer-reviewed publications were co-authored by the staff of the unit. A significant portion of them, 68, are journal publications, of which 34 are in tier A general journals. Of the remaining conference and workshop papers, 48 are in tier A general conferences and 40 in tier A specialized conferences. The unit was also very active in the encyclopedia initiatives of the database community, co-authoring 50 encyclopedia entries.

We find this level of productivity very satisfactory as well as exceeding the unit's own expectations and the unit's performance in the previous evaluation period.

Funding

During the evaluation period the unit was active in the total of 27 externally funded projects. The funding for these projects totaled to DKK ~ 37M counting the funding received by the unit from national, European, and other international funding sources. In summary, the unit has been very successful in attracting funding. The attracted funding resulted in new research activities and helped to improve the relevance of the unit's research. On the negative side, the large number of projects resulted in increased administration overheads. The units resources could be utilized better by focusing on larger, less administration heavy projects.

Scientific service and recognition

The members of the unit served on more than 135 program committees for conferences and workshops, including committees for the most prestigious conferences in the area. In addition the members of the unit served in leadership roles (such as chairs or co-chairs) in the program committees of 20 conferences and workshops. In 2009 the unit successfully organized the 11th *International Symposium on Spatial and Temporal Databases* in Aalborg. During the evaluation period, the members of the unit served on editorial boards of 5 journals including *ACM Transactions on Database Systems*. Christian S. Jensen is the editor-in-chief of the *VLDB Journal*. In addition, members of the unit are performing services as project and program reviewers for a number of funding agencies. In summary, we believe that the unit has performed very substantial service to the scientific community. This demonstrates the visibility and high regard of the unit in the research community.

Hiring

The 2001-05 period was quite turbulent staff-wise, with only two out of seven permanent faculty members being employed throughout. The 2006-10 period has been less turbulent, but still dynamic. The unit had one new permanent faculty member and lost another. The unit has hired six new assistant professors, including several from top universities such as National University of Singapore, RWTH Aachen, and University of Hong Kong. Most of the assistant professors have left for positions at other universities or in industry. In addition, the unit has hired 2 post docs, around 20 Ph.D. students, and about a dozen research assistants. Most of the Ph.D. students and many of the research assistants come from international universities. In summary, the unit has been successful in attracting good candidates at the Ph.D. student and assistant professor levels. However, it is vital for the unit's continued success to be able to recruit more internal Ph.D. candidates and to be able to offer permanent positions to more of the top assistant professors.

Other aspects

The period has had a lot of reorganization at most levels of the university, regarding both research and studies, which has taken a heavy toll especially on permanent faculty members. The unit members perform many heavy administrative tasks, and should perhaps try to reduce these in the next years. While the teaching load has been mostly stable in

the period, the unit fears that it will increase in the coming years. To be able to cope with the situation, the unit needs to recruit more staff in some areas, e.g., programming. TAP support for project application writing and project administration has improved, but the unit feels that support for other administrative tasks and the IT infrastructure for trivial tasks such as course and exam management needs to improved, and in general, a more stable IT infrastructure is needed.

4.7 Plans (2011-2015)

4.7.1 General Plans

The overall approach to conducting the research in the unit will continue along the same lines as earlier.

We thus still strive to be *use inspired*. Research decisions will often be guided by concerns for use in practice or anticipated practice—the research in this area is mostly motivated by perceived practical concerns. In line with this principle, we will try to interact with practitioners on a regular basis, e.g., via participation in advisory boards or boards of directors for technology companies, via collaboration with practitioners in funded research projects, via participation in industry associations, and via lecturing to industrial audiences. This is expected to have a positive effect on relevance and impact.

In order to gain synergy between individual activities, the unit focuses on a few advanced key application areas, namely mobile services, intelligent transport systems, energy, and logistics.

Second, the activities will aim for insights, concepts, principles, and techniques that have *general applicability*. Thus, the objective is to obtain general results that go beyond a specific application. This implies that the outcomes are more abstract than solutions to specific, practical problems. We consider it the task of industrial collaborators to apply these outcomes to solve concrete problems. This division of labor within the value chain that includes both research and innovation is quite important.

Third, the activities are generally *constructive* and *experimental.* They fundamentally involve the design of artifacts. This research process often, although not always, requires that prototypes be implemented, either as proof of concept or in order to elicit design properties that may guide the continued design process. Much of the research will concern artifacts that aim to stand out by virtue of their performance, rather than, e.g., functionality or usability, focus will be on empirical studies of performance characteristics.

The activities are thus different from basic research, which is not use inspired, and product development, which aims for specific solutions to specific problems, but not for general insight. Using a concept from the framework articulated by Donald E. Stokes, we aim for Pasteur's Quadrant and Edison's Quadrant. The general approach will be constructive and iterative: New techniques are designed and subsequently prototyped, either for proof of concept or with the purpose of eliciting design properties, typically performance properties, via empirical studies. The resulting insights are then used for re-design and improvements. Some of the techniques invented may well be applicable in the near term, while others may be applicable only indirectly and in the longer term.

With the focus on quality and impact, the unit plans to follow the general publication strategy formulated in the previous section. We plan to seek funding and industrial collaboration opportunities involving strategic partnerships in the main application areas.

4.7.2 Main and Focus Areas

To organize and focus the unit's research, we define an umbrella research area and three focus areas under it.

Main area: data-intensive services - mobile, ubiquitous, cloud, and beyond.

The main overall research theme is inspired by the increasing proliferation of IT services. These services are expected to be increasingly data-intensive, requiring processing of ever larger and more complex data sets in order to provide the most optimized service. The services will span an increasingly complex sphere, with some used on mobile devices and others used in ubiquitous settings, with some used by persons and others used by

computers, with an increasing interplay between many diverse services. The services will increasingly be based on a cloud computing environment.

As part of the overall theme, the unit plans to develop a number of building blocks for data-intensive systems in a cloud computing environment. Research will be done on developing frameworks, algorithms, data structures, and query processing techniques that fully exploit the inherent parallelism found in new hardware architectures such as multi-core CPUs, as well as in massively scalable cloud computing environments. In addition, as the data-storage hardware diversifies data management techniques will be explored that take advantage of novel data storage hierarchies involving, for example, main-memory and flash storage. In this context, the unit plans to work on indexing and query processing techniques that can achieve very high throughput of updates while supporting simultaneous efficient queries.

Focus area: cloud intelligence

Within the areas of business intelligence and data warehousing, the main focus area will be on *cloud intelligence*. Cloud intelligence emerges from the migration of BI and analytics technologies to a cloud computing environment while exploiting the associated new opportunities. A multitude of different types of data exist "in the cloud," including structured, relational data, multidimensional cube data, text data, semantic web XML/RDF/OWL data, geo-related data, and sensor data. Finally, many analytical models of data have been developed through data mining. To achieve true cloud intelligence, all these types of data/models must be integrated and analyzed in a coherent fashion, including privacy protection.

Our overall vision is to develop technologies that enables a "cloud warehouse" (CW). The basis for the CW will be a novel kind of data model that supports both multidimensional concepts and the flexibility from semi-structured data. At the same time, the model should be capable of supporting the wide range of data mentioned above.In this context, we will explore data models, query languages, query processing and optimization techniques, data integration techniques, and techniques for integrating databases, sensors, and analytical and predictive models of data.

Within the general area of cloud intelligence, several concrete activities are planned. Methods for spatio-temporal data warehousing and data mining (especially prediction and forecasting) will be developed for the specific domains of traffic services and mobile phone networks. The work on sentinel mining will continue by investigating the predictive power of sentinels and cloud deployment. Effective compressed bitmaps indices will be further developed and matured for real-world usage. Techniques for warehousing semantic web data will be developed. Real-time data warehousing techniques will be further developed. An ETL framework for cloud computing will be developed. Architectures and query processing techniques for massively distributed large-scale data warehouses with tight integration of historical, streaming, and forecast data will be developed.

Focus area: spatio-temporal data management and mobile services

The unit plans to build on its strengths in spatio-temporal data management. In particular, the unit will continue research efforts in spatial and spatio-temporal indexing and query processing. Efforts in this area will focus on exploring important, for applications, trade-offs, such us, for example, how query result accuracy or freshness depends on the updating costs. In addition to experimental studies, deeper theoretical understanding and cost modeling of workload-dependent performance will be pursued.

In general, the unit plans to explore more realistic, application-specific assumptions, such as models involving privacy of user locations and models for indoor and hybrid idoor/outdoor spaces. The unit plans to research spatio-temporal data management techniques for profile-based, context-sensitive, and adjustable privacy of user locations. Trade-offs between privacy, quality of services, and price of services will be explored.

The unit will continue its internationally leading research on the data management in indoor settings. A particular attention will be paid to relatively resource-limited, hand-held mobile devices that, as increasingly popular hardware, accommodate new inventions for indoor mobile services. In addition to the RFID-based positioning, other indoor positioning technologies will be assumed in the future research. It is thus interesting to integrate heterogeneous underlying positioning technologies in order to improve the accuracy of tracking, indexing, and querying indoor moving objects. Within the area of mobile services, it is desirable that services are offered across indoor/outdoor boundaries

in a seamless way. As the data management foundation, an integrated space model for indoor and outdoor spaces is needed. Future research will cover this crucial aspect. Particularly, such a model should be able to measure distances according to Euclidean and/or network distance metrics. It should also be able to denote both outdoor and indoor POIs (points of interest). Further, such an integrated model should support service requests that involve both outdoor space and indoor space. On top of such a model, integrated indexing and query processing is another relevant aspect to explore.

Focus area: an umbrella project in programming technology

Programming technology will remain at the core of any software development. The demand for software keeps increasing despite crises in IT and financial industries, thus research into improving programmer productivity through improved tools is paramount to the IT industry in particular and society in general.

This fact is often overlooked by funding agencies and policy makers, thus research into the core disciplines, such as new programming languages, is often ignored in favour of newer, more glamorous, and admittedly, more applied research areas where results are directly turned into new products or services. Nonetheless, we find that it is important and central to have activities in this core area, and our new project is, like its predecessor the P2025 project, intended as an umbrella activity to coordinate activities that would otherwise be carried out on an individual basis.

We expect to continue to take part in collaborative efforts, especially in mobile and embedded systems, but also anticipate new directions in data intensive applications such as business intelligence and eScience.

In the short term we will engage with (local) industry to help them understand the shift in software construction from the object oriented approach to the more declarative and dynamic approach being offered by the new language features in Java and C#, as well as new languages such as F#, Scala and Clojure. Participation in projects, such as DaIn, CaIn and InfinIT, will continue and we anticipate taking part in new collaborative efforts and even seek the opportunity to lead efforts in this direction.

We also find that it is important to raise the profile of the unit in this core discipline. To give focus to this endeavor we will initially pursue a promising new direction which we call Programmatic programming, i.e. programming constructs that allow the programmer to write programs that manipulate programs. In addition we intend to research the transitions between programming with dynamic and static types.

Initially we plan to utilize research funds provided by the department and involve M.Sc. student projects, but as the work progresses, we plan to seek external funding, in particular we plan to seek funding for at least one Ph.D. student in this area.

4.8 Committee evaluation

The DPT Research unit comprises 7 permanent academic plus significant and varying numbers of non-permanent academic staff. These numbers together with the fact that the senior staff members are well established at the international level, while the junior staff members are highly talented, depict that the unit had a good size and had the human resources necessary to compete nationally and internationally.

During the period under examination, the unit members have produced a significant number of articles (68 journal publications, 167 conference publications, plus several other publications such as book chapters, books, edited volumes, etc.). According to the unit's own categorization scheme, 57% of the published articles appear in top journals, whereas according to the national categorization, the respective number reaches 69%. This figure is exceptional. Moreover, the unit is highly selective with respect to the conference outlets for submission. The unit has succeeded to publish 50% of its conference articles in such top conferences. The unit's record of publications in top journals and conference proceedings places the unit as a leading group at the European level. The unit is among the top-5 database groups in Europe, on level with institutions like ETH, and with an analogous significant reputation at world level.

4.8.1 Challenges

The unit has managed to attract significant funds from national and international sources, totalling 37 million DKK over the evaluation period. Collaborations with many industrial partners have been established, mainly in the subareas of business intelligence, intelligent transport systems and logistics. Further collaborations along the same lines are anticipated. Also, the Streamspin system and indoor/outdoor positioning and tracking constitute a modern research line and should lead to important commercial products and services and a tight collaboration with industry. The unit has collaborated with over 30 companies and organizations in the evaluation period, an impressive volume. The unit has been the breeding ground of 5 spinoff companies, which should be increased on the way..

During the period, the Unit is characterized by some two-way mobility of personnel from/to top universities. The unit has produced 12 Ph.D. graduates with most candidates going to industry positions, e.g., in Google. To continue this way and keep up at this level is a desideratum.

The Unit comprises two subgroups with quite distant topics. Both subgroups have spent serious energy in establishing a common ground for research collaboration, applying for joint projects, etc. In its current state, the collaboration is not as desired.

4.8.2 Recommendations

The possibility of extending (as tends to happen at this stage) interests in less traditional and more modern topics should be considered. Further deepening into data-intensive research areas, such as transport systems, logistics and so on, e.g. into side sub-areas, is an equivalent direction. Thus, the potential of the unit could be multiplied.

The unit should exploit the research outcomes in the direction of creating spin-off companies to commercialize them into products and services with the unit staff participating in their organizations. At least the unit members should act as catalysts so that close

collaborators (non-permanent staff) of the unit may follow an entrepreneurial direction.

The two subunits should insist in trying for tighter collaboration between them and have concrete results in fund raising and publishing. In addition, the unit members should work further towards collaborating with other units, e.g. with MI unit as happens already with the DES unit, through the subunit that specializes in Programming Languages.

The Daisy concept should be further developed and exploited for external use since it is a nice branding idea.

Much of the present successful outcome is due to several senior or younger persons, which have left or are going to leave the unit. It is necessary that in the future, the current members of the unit maintain the throughput in terms of fund raising, publishing in top journals and conferences, Ph.D. student supervisions, service to the community, and so on. Finally, it is important that the unit is enabled to fill another full professorship as a replacement for Christian S. Jensen.

5 Distributed and Embedded Systems
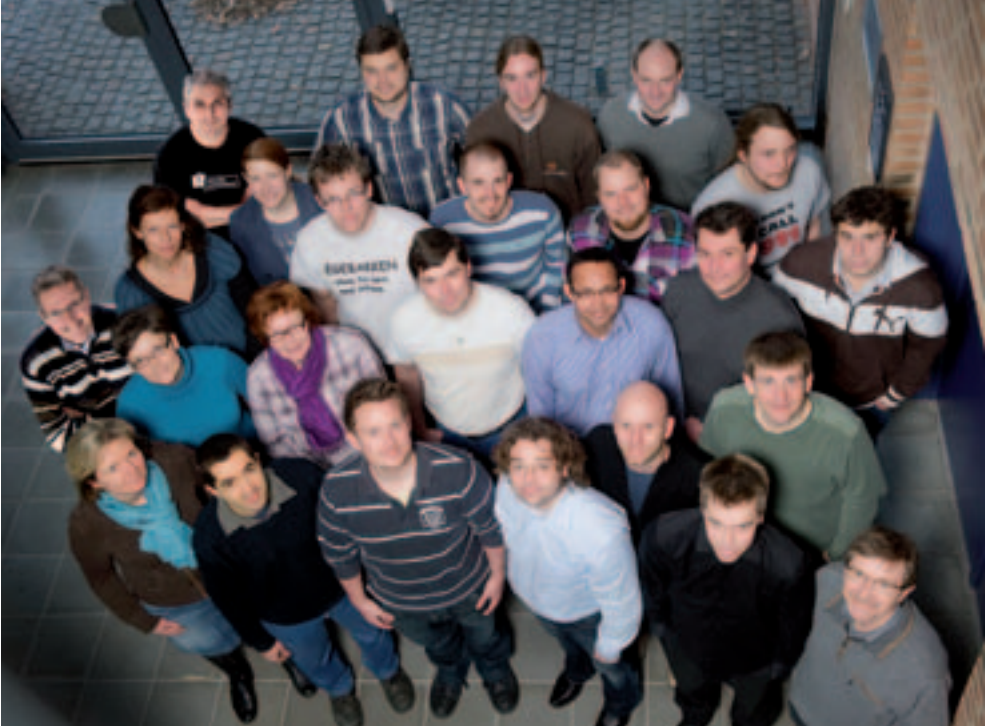
The Distributed and Embedded Systems Unit and CISS.

5.1 Executive summary

The research of the distributed and embedded systems unit concerns modelling, analysis and realization of computer programs, with emphasis on distributed and embedded systems. The contributions of the unit include work on semantic foundations, algorithms and tools for verification and validation, and design and implementation of networks and operating systems. Thus research activities are both theoretical/foundational in nature as well as practical/experimental.

The unit takes a leading international position within verification and validation, embedded systems and concurrency theory, and has a high national profile in networks and grid computing. The work centered around the verification tool UPPAAL has been a central activity of the unit with high international impact. Following the main strategic goal for the period 2006-2010, Center for Embedded Systems has established itself as a preferred and coordinating partner in national and European projects, with significant industrial impact.

During the period the unit has been involved in extensive research collaborations. At a national level it played a principle role in two large research projects DaNES (Advanced Technology Platform Project) and MT-LAB (VKR Center of Excellence) as well as in the Innovation Network InfinIT. The external funding has risen considerably from 4 projects with 41 MDKK in the previous five-year period to 30 projects with 65 MDKK in 2006-2010. The unit is now also a member of the *Danish Center for Scientific Computing* (DCSC). Internationally, the unit has been engaged in several EU projects including co-coordination of ARTIST Design -- *the* Network of Excellence within embedded systems -- as well as coordination of the STREP project Quasimodo, both under EU's 7th Framework Programme. The international visibility of the unit is also witnessed by a high number of publications in international conferences/workshops (138), journal articles (31) and book chapters (10), research awards, invited talks, numerous services to the research community, and the high number of external visitors.

5.2 Profile

The research of the unit concerns modelling, analysis and realization of computer programs, with the emphasis on distributed and embedded systems. This includes the following areas:

- Semantic theories for modelling the behavior of computer programs and systems.
- Design, implementation and models for analysis of distributed systems and networks.
- Algorithms, methods and tools for verification and validation of programs and systems.

In recognition of the increase in research on embedded systems the name of the unit has been changed from *Distributed Systems and Semantics* to *Distributed and Embedded Systems (DES).* In 2001 the Centre for Embedded Software Systems (CISS) was established in collaboration with the Institute for Electronic Systems at Aalborg University, with the ambi-

tion to create an industry oriented research center of excellence with a distinctively visible profile within software construction and embedded systems interacting closely with the research conducted by the DES unit. During the period 2006-2010 the research on embedded systems has been even further increased with substantial effort on model-driven and component-based development, enabled by the Advanced Technology Platform Project DaNES and on establishing an adequate theoretical foundation for embedded systems by integrating discrete, stochastic and continuous aspects, enabled by the VKR Center of Excellence MT-LAB. Consequently, the research of the unit spans a wide range from development of the semantic foundations of computing and concurrency to contributions to the technological state of the art within embedded and networking systems. To conduct quality research within such a wide spectrum is a challenging task but a common emphasis on distributed/concurrent systems focuses the activities within the unit.

Each of the three research areas mentioned above constitutes a subject in its own right. Moreover, the areas are interrelated in a number of ways: semantic models offer important guidelines for development of languages and paradigms for distributed systems; semantic models are necessary prerequisites for development of verification algorithms and tools; the development of validation tools provides new insight into the underlying semantic models on one hand, and it is applied in environments for the construction and analysis of distributed systems; the evolving nature of distributed systems provides insight into the strengths and weakness of existing semantic models, and serves as an inspiration for the development of new ones; finally, distributed systems truly expose the limits of developed verification algorithms. The current research of the unit includes the following activities.

Concurrency Theory: Semantic theories and meta-theories for concurrent processes and their logical properties. Models, logics and analysis for hybrid, real-time, and probabilistic processes. Semantic theories for processes whose communication topology changes dynamically, including security protocols.

Verification and Validation: Development and implementation of data structures, algorithms and tools for model-checking, static analysis and testing of embedded systems and real-time systems. Applications to communication protocols, control programs, and planning and scheduling.

Networks & Operating Systems: Analysis and construction of services and protocols for computer networks, including grid computing and high-performance computing. Real-time and embedded operating systems.

Embedded Systems Methodology: Methodologies for specification, analysis and testing of embedded systems. Modelling and analysis of industrial case studies. Theoretical foundations of embedded and hybrid systems.

5.3 Staff

5.3.1 Current Staff

Professors: Kim Guldstrand Larsen, Anders Peter Ravn

Associate professors: Gerd Behrmann (on leave), Alexandre David, Hans Hüttel, Josva Kleist (parttime), Brian Nielsen, Arne Skou, Jiri Srba

Assistant professors: Ulrik Mathias Nyman

Post-docs: Ulrich Fahrenberg, René Rydhof Hansen, Radu Mardare

Project related researchers: Jens Alsted Hansen, Jacob Illum (parttime), Jan Jacob Jessen, Marius Mikucionis, Andreas Popp, Andrzej Wasowski (parttime)

Ph.d-students: Thomas Bøgholm, Mehdi Gholami, Arild Martin Møller Haugstad, Line Juhl, Kenneth Yrke Jørgensen, Morten Dahl Jørgensen, Morten Kühnrich, Mikael Harkjær Møller, Piotr Niemizye, Mads Christian Olesen, Petur Olsen, Mikkel Larsen Pedersen, Christoffer Eg Sloth, Claus Rørbæk Thrane, Saleem Vighio

5.3.2 Staff Development

During 2006-2010 the scientific staff has been reduced by 1 associate professor, i.e. the unit is by now formed by 2 full professors, 7 associate professors and 1 assistant professor. Moreover, due to reductions in teaching loads (to serve the activities in CISS and Nordugrid), the teaching capacity is further reduced by the equivalent of 1.5 permanent position. Overall, the teaching capacity over the last five year period has been reduced by one full position.

On the other hand, due to successful grant applications, the unit has had a substantial growth in the number of postdocs and project related researchers, and it currently employs 9 persons of which 6 are dedicated to industrial collaboration (2 only on a part time basis), and 3 contribute to PhD supervision. The calls for PhD grants have attracted a high number of applicants, and the number of PhD students employed at the end of the evaluation periods has increased from 8 to 15.

5.4 Goals (2006-2010)

The **overall strategic research goals** of the unit in the period 2006-2010 were to: become a leading European research unit within Embedded Systems; maintain and strengthen our leading position with respect to verification and validation of real-time system models; become a leading national unit in GRID research and services; maintain our high international visibility in concurrency theory.

Concrete plans for the existing four research areas were as follows. In *Concurrency Theory* the work on decidability, axiomatizations, and expressiveness for infinite state processes as well as logical specification formalisms will continue. For mobile process calculi, the work on equational logics, type inference and expressive power for process calculi with sites will be in focus. In *Verification and Validation*, the work on UPPAAL will continue: extending current verification technology with techniques from abstract interpretation; supporting model-checking with full TCTL and even with a distributed algorithm. The development of tool support for optimal scheduling and controller synthesis within UPPAAL CORA and UPPAAL TIGA will continue. UPPAAL TRON will be augmented with interfaces for test execution on various platforms. In *Embedded Systems Methodology* new case studies will be

done. Compositional methods for analysis of embedded systems will be pursued. Within hybrid systems, the main goal is to create a strong interdisciplinary research environment with researchers from control theory. In *Networks and Operating Systems,* the focus will be on solutions for data management in grid systems, adaptive firewalls and peer-to-peer information systems. A specific activity is an effort towards the realization of a European Verification Grid. Work towards secure, networked embedded systems will continue with industrial partners.

Staff and Funding: The (likely permanent) leaves of absence and reduced teaching obligations of several senior staff members affect negatively the research and teaching effort of the unit. This situation should be resolved as soon as possible. A main challenge of the period 2006-2010 has been to secure new funding of similar size as for the previous period(s), where in particular the center BRICS (Basic Research in Computer Science) was a major source of funding. As most future funding was predicted most likely to be subject to industrial collaboration, it was found important that means were identified for reserving sufficient portions of the funding for pursuing basic research goals.

5.5 Activities and Results

5.5.1 Concurrency Theory

Weighted Automata and Games: While *timed automata* are well suited for modeling of real-time systems, the extension towards *priced timed automata* have proven very useful for modeling quantitative resource consumption and for performance analysis. A number of optimization problems for priced timed automata have been shown decidable, extending the original decidability results for cost-optimal reachability from 2001. In particular, the problem of *optimal infinite runs* both in terms of long-term cost per time ratio as well as a discounted measure have been shown decidable. Also optimal reachability in the presence of multiple costs has been considered and shown decidable. Finally both model-checking for priced timed automata and the problem of synthesizing optimal winning strategies for *priced timed games* has been considered. Though known to be undecidable in general, decidability and complexity results have been obtained in the setting of one-clock models.

We are glad to report that our work on priced timed automata has been invited as a CACM review article which will appear shortly.

With colleagues at CNRS Cachan (F) we have introduced the formalism of *energy timed (game) automata* which is a syntactically simple extension of priced timed automata allowing for *positive* as well as *negative* cost-rates. From the notion of *cost* being a pure observer variable with no effect on the timed behavior, in our model it makes sense to have (lower and upper) bounds on the energy level possibly restricting the underlying timed behavior. For this new model a number of (un)decidability and complexity results have been obtained in the setting of a zero or one clock and in the setting of linearly or exponentially growing cost-functions. However, a number of interesting problems are still unsettled and the model has in the last two years attracted a considerable international research interest.

Recently we focused also on the expressiveness results related to real-time extension of finite automata and Petri nets, demonstrating close connections between such models up to different behavioural equivalences and/or TCTL model checking.

Infinite State Systems: The need for modelling of unbounded data structures, control-flow graphs of recursive programs and parameterized reasoning give natural motivation to study models of computation with infinitely many reachable (discrete) states. We continued the research on decidability and complexity questions about infinite state models, with a particular focus on systems defined via various variants of prefix rewriting, like pushdown automata and more general classes of systems from Stirling's hierarchy. In a series of papers resulting in a journal publication in JACM we provided a characterization of decidability of bisimilarity on prefix-rewriting and other systems like PA and we also studied the problems of bisimilarity on visibly pushdown automata. Weighted extensions with some positive decidability results were also established.

Metrics on Reactive Systems: For analysis and verification of embedded and/or hybrid systems, standard system equivalences and preorders such as bisimilarity or trace inclusion have to be generalized to (hemi)metrics. These metrics measure e.g. how well a certain system implements a specification, or to what extent two different implementations are to be considered equal. During the evaluation period, we have started research in this area. So far this has resulted in a number of workshop papers, three journal articles including one

Best Paper Award (MEMICS'09) which introduces several interesting system distances and studies their properties like expressiveness, computability/complexity and axiomatizations.

Mobility and Security: Variants of the pi-calculus can be used to describe cryptographic protocols, and type/effect systems can then be used to reason about properties of such protocols through type checking. Firstly, we have shown how to extend these ideas to a process calculus with localities and broadcast communication and considered type systems for properties of receptiveness. Moreover, we have found algorithms for type/effect reconstruction; the underlying idea relates the work on type/effect systems to that of computing analyses in the work on flow logics for pi-calculi. We have further studied the expressive power of process calculi with localities providing a characterization of a finite-control property in the higher-order Homer calculus.

Modal Transition Systems: Modal transition system is a formalism which extends the classical notion of labelled transition systems by introducing transitions of two types: must transitions that have to be present in any implementation of the system and may transitions that are allowed but not required. The framework has proved useful as a specification formalism of component-based systems as it supports compositional verification and stepwise refinement. During 20 years since the original concept was introduced, the complete characterizations of the key semantic notions of consistency and refinement have remained open. During a sequence of papers in collaboration with researchers at ITU (DK), Brno (CZ), Imperial College (UK) and CNRS Cachan (F), we have settled the complexities of many open problems. Restricting to deterministic modal specifications may be more manageable yet still interesting for several applications. We have given a comprehensive account of how such a restriction may improve complexity of key notions such as consistency and refinement.

In 1991 Jonsson and Larsen introduced a probabilistic variant of modal transition systems known as Interval Markov Chains (IMC) or Probabilistic Specifications. Though equipped with proper definitions of satisfaction and refinement, the formalism lacks several closure properties. As the result of an intense research collaboration between ITU (DK) and IRISA, Rennes (F), the notion of Constraint Markov Chains, being a generalization of IMC, has been introduced. This formalisms provides the first complete compositional specification theory for Markov chains that supports parallel composition, conjunction, renaming, refinement and satisfaction checking.

Contracts and Interfaces: A software product line theory based on model transformations to simplify a general model to a given environments inability to observe certain behaviour (color-blindness) was developed. An interface theory based on I/O automata with an explicit split of assumptions from guarantees was established and a formal correspondence between this theory and interface automata was given. From this basis an interface theory and an accompanying variability model for software product line development was introduced based on modal I/O automata; the novel aspect being the ability to describe liveness, thus disallowing trivial implementations. Recently the first complete specification theory for real-time components, specified as timed I/O automata, with constructs for refinement, consistency checking, logical and structural composition, and quotient was developed and implemented in the tool ECDAR.

5.5.2 Verification and Validation

Model Checking: During the period several contributions has been made towards improving the verification for timed automata with subsequent implementation in the engine of UPPAAL. The techniques include compact and efficient symbolic datastructures for state-space representation by on-the-fly merging of symbolic states and by introduction of (exact) abstraction techniques resulting in improved performance of an order of magnitude. In addition, advances have been made on the algorithms used for state-space exploration including adaptation of AI planning heuristics as well as abstraction-base heuristics in order to obtain efficient directed model checking, development of a space-efficient agent-based search engine as well as implementation of abstraction/refinement for timed automata verification.

The modeling formalism of UPPAAL has been extended with a number of features, allowing full use of broadcast channels, possible decoration of processes and/or transitions with priorities, and allowing the use of stop-watches. The specification formalism has also been enriched with quantitative queries and with the possibility of specifying properties using the graphical notion of Live Sequence Charts. The tool and its algorithmic foundation has been disseminated in a number of tutorial papers, including publications and lectures given twice at the Marktoberdorf Summer School. Further indication of the wide dissemination of the tool can be seen from the fact that the paper "UPPAAL in a Nutshell" from 1998 is the most cited paper overall in the Journal Software Tools for Technology Transfer by Springer. Utilizing

the verification engine of UPPAAL, the tool TAPAAL has been developed allowing for verification and simulation of Timed-Arc Petri Nets.

Static Analysis: The theoretical foundations for the Coccinelle program analysis and transformation tools were developed and the Coccinelle tool itself was further evolved and developed. The tool has been used, and is actively being used, for finding potential security vulnerabilities in critical infrastructure software such as the Linux kernel and the OpenSSL cryptographic library. The use of static analysis for assessing and evaluating potential insider attacks has been investigated and relevant models have been developed and analysed. Towards the end of the period significant work has been conducted on combining static analysis of low-level source code with model checking for analyzing embedded systems comprising (executable) code and execution platform using abstract timed automata. In particular a tool support for WCET (worst-case execution time) analysis of C-code taking into account architectural aspects (e.g. caches and pipelines) of the underlying the architecture (e.g. ARM7 and ARM9 processors) has been made. Also static analysis and model checking have been (in cooperation with the DPT unit) combined to enable schedulability analysis of Safety Critical Java programs executing on FPGA. Finally, we used the combination of model checking and static analysis techniques for applications in testing and schedulability analysis.

Testing: Model based testing is a promising approach for testing of complex software systems where effective test cases are automatically generated from models of system requirements or specified system behavior. The overall aim is to develop testing theories and efficient model-based tools for functional and quantitative aspects of distributed and embedded systems. The Yggdrasil tool generates edge-covering test suites of large models. By using heuristic agent-based search algorithms and the UPPAAL-CORA optimization engine, it aims at maximizing the achieved coverage while reducing the number of test cases and their individual length. The approach has been put into commercial industrial use; application to GUI testing of medical devices at Novo-Nordic has reduced testing maintenance time from 30 to 3 days.

More fundamentally, we have successfully re-interpreted the real-time test generation problem and theory in a game theoretic setting enabling a winning strategy to be used as a test case, and thereby coping with timing uncertainty and uncontrollability. We used the timed game solver UPPAAL-TIGA to generate tests and cooperative testing strategies.

Several significant usability, applicability and performance improvements have been implemented in the online testing tool UPPAAL-TRON that has brought it to a mature level. To support testing of hybrid systems, we have made a first integration of TRON and the linear hybrid automata symbolic model checker Phaver for monitoring of continuous signals, and TRON has been extended with co-simulation capabilities with Matlab-Simulink. We have recently engaged in collaborations with the MI unit on learning systems models using a combination of testing, traditional active and passive learning algorithms, as well as (probabilistic) machine learning algorithms.

Synthesis and Scheduling: We formulated the problem of controller synthesis as a timed game and developed UPPAAL-TIGA based on the original paper in 2005. This tool is now fully integrated with the UPPAAL tookit and supports its advanced syntactic constructs. Two case-studies have been reported on the successful application of the tool, namely a climate controller for stables and the Hydac oil pump case-studies. In addition, we extended both the theory and practice from games with full observability to games with partial observability. We have further developed an engine to solve turned-base games that have applications in simulation and refinement checking. Regarding scheduling, UPPAAL has been extended with stop-watches. This addition makes the (exact) model-checking problem undecidable but we have implemented a practically usable over-approximation technique. This addition has been put to good use for a range of different schedulability problems.

Tools: We understand the importance of validating the theoretical results via experiments and our unit has been involved in the development of numerous tools, among others UPPAAL, UPPAAL TRON, UPPAAL TIGA, UPPAAL CORA, OPAAL, TAPAAL, ECDAR, Coccinelle, VPlus, Yggdrasil, SARTS and Metamoc.

5.5.3 Networks and Operating Systems

High Performance Computing: The main goal w.r.t. high performance computing was to have AAU become a member of the *Danish Center for Scientific Computing* (DCSC). This goal was reached in 2007. The DCSC installation is used by all of AAU and provides the base for collaboration with other departments on parallel algorithms. The main result

has been a journal paper on optimization of windmill wings. Our unit is also involved in a recently started Nordic project on creating a platform for sharing Nordic HPC resources.

Grid: The grid computing activities have to a large extend been related to the collaboration with Nordic DataGrid Facility. The outcome has been a series of publications related to distributed storage of large amouts of data, handling of LHC computations and the construction of a common Nordic grid infrastructure.

Embedded Systems Platforms: A Java profile suitable for development of high integrity embedded systems has been developed and described in a number of papers. The main contribution is an arrangement of the class hierarchy such that the proposal is a generalization of Real-Time Specification for Java (RTSJ). Also, several papers have addressed the issue of how to use model checking and/or static analysis to analyse *Java programs for predictable timed behaviour* and to analyse *Worst Case Execution Times* for ARM processors including cache behaviours. This research is conducted in cooperation with the DPT unit.

5.5.4 Embedded Systems Methodology

Service Oriented Architectures: In collaboration with external partners within the CoSO-DIS project we have investigated application of model checking based analysis techniques for orchestration and choreography languages for service oriented applications. A deeper investigation of the main orchestration language WS-BPEL has resulted in developing an operational semantics for the full language which has uncovered several ambiguities in the informal semantics. This formal semantics has been encoded in the Maude tool in order to check its consistency. Protocols for Web Services like WS-Atomic Transaction and WS-Business activity were also investigated and formally verified. A part of this research is a collaboration with the MI unit (previously with IS unit).

Hybrid Systems: As part of CISS and of the externally funded MT-LAB project, a collaboration with people from the control department in Aalborg has been initiated, with the dual aims of applying control theory reasoning in verification of hybrid systems and of applying real-time tools like UPPAAL to control theory problems. One specific goal was to apply control

theory techniques and tools to make hybrid systems amenable to verification using priced timed automata technology. This collaboration has resulted in a number of conference papers (at HSCC and CDC).

Component Based Methodology: An effort has been put into the design and tool development for the refinement calculus for Component and Object Systems (rCOS) in collaboration with the unit at UNU/IIST. The rCOS formalism has done well in the component modelling contest (CoCoME) which is documented in an LNCS volume that is an outcome of a Dagstuhl seminar.

Industrial Case Studies: As part of several externally funded projects (CSDR, Quasimodo, Multiform, DaNES), the UPPAAL model analysis tool suite has been applied to contribute to specific parts of ongoing product developments at selected industrial partners including Novo Nordisk, Terma, Skov, HYDAC, CHESS. Also, new tools have been developed to form a bridge between UPPAAL models, industrial standard notations, and the software interfaces for the industrial products. The results have been published at industrial conferences and they include a UML model transformation tool, an automated test script generation tool, and a general framework for schedulability analysis. Also, the application of UPPAAL TIGA for synthesis of control programs and configuration testing, has been successfully demonstrated and compared to existing controllers.

5.5.5 Dissemination

Towards Students: We consider the dissemination of our research results and its integration into students' curricula as an important aspect of our research. We have had a great success with teaching the theory of reactive systems to undergraduate students and a number of lecture notes gathered by different lecturers over the last decade or so resulted in a book "Reactive Systems: Modelling, Specification and Verification" published in 2007 by Cambridge University Press. The book is nowadays used at more than 15 universities around the world. Very recently another book aimed at teaching syntax and semantics to undergraduate students called "Transitions and Trees: An Introduction To Structural Operational Semantics" was published by Cambridge University Press.

Within Embedded Systems numerous contributions to PhD schools have been given by members of the unit. Also, in 2010 we organized the international PhD School *Quantitative Model Checking* with more than 100 participants and featuring 12 leading international lectures. In order to further strengthen the integration of our best master students in our research activities, we offer Elite Education in Embedded Systems. Currently, we accept 5-10 Elite Students per year into the Master's program. Finally, we joined the EU 7th Framework Programme on Teaching, Research and Innovation in Computing Education (TRICE) in order to further improve our teaching competences. A research about educational aspects of teaching computer science and the theory of concurrency in particular has recently started and resulted so far in two peer-reviewed publications. Researches from the DPT unit participate in this project too.

Towards Industry: Industrial dissemination of research results serves the purpose of contributing to industrial innovation, to competitiveness, and to stimulate further industrial collaboration. The activity is funded by a number of projects, where the central projects are CISS2, CISS-ConneCT (dissemination at regional level), InfinIT (dissemination at national level), and the European NoE ARTIST Design, which serves as a strategic platform for international collaboration and dissemination. In order to support and attract dissemination, a number of so-called 'demonstrators' have been developed and presented at industrial exhibitions together with regular industrial collaboration results. The demonstrators include home automation scenarios, production plants and various sorting equipment. Also, dissemination has been pursued through handbook material and through the development of industrial courses that have been taught both as stand-alone courses and as parts of accredited life-long learning activities. Researchers from the IS unit have participated in several of the national projects on dissemination and applied research and also the DPT unit has been involved in one such national project.

5.5.6 Funding

The beginning of the evaluation period (2006) marked the ending of a substantial research and dissemination funding from the Danish Council for Technology and Innovation and from the Region of North Jutland (in total 37 MDKK, 2002-2006). Also, the funding of the Basic

Research Centre BRICS ended in 2006 (in total 15 MDKK 1994-2006). During the evaluation period, the possible funding sources have changed into smaller portions of money per call, and in order to maintain resources for research and dissemination (including PhD scholarships), the unit has been involved in a somewhat high number of research grants -- in total 30 projects with a total funding (with the unit as main applicant) of more than 100 MDKK of which 65 MDKK were dedicated to be spent by the unit. The figure below illustrates the distribution of projects and funding on their degree of basic and applied research and also dissemination.

Project type	No. of projects	Total grant	Av. grant/project
Dissemination	17 (2 international)	27.220.000 DKK	1.601.176 DKK
Applied research	4 (1 international)	16.400.000 DKK	4.100.000 DKK
Basic research	9 (4 international)	23.211.000 DKK	2.579.000 DKK

The key projects are QUASIMODO, GASICS, MULTIFORM, ARTIST and MT-LAB (funded by FP7, ESF and the VKR Foundation for basic research, dissemination and networking), DaNES and InfinIT (funded by HTF and RTI for industrial dissemination and collaboration), and finally CISS2 and CISS Connect (funded by RTI and Region North).

5.6 Own Evaluation

In our opinion the research activities carried out by the DES unit during the period 2006-2010 have been very satisfactory both in terms of quantity and quality. The number of publications by members of the unit has remained high despite the loss of three productive associate professors at the beginning of the period. The high quality of publications is evidenced by their appearance in well-respected journals, as invited contributions to special issues and book chapters and in proceedings of competitive conferences. The total number of publications (181) in the period 2006-2010 matches that of the previous period (183) and includes 2 books and 10 invited book chapters. Of the remaining 169 reviewed publications, 30% of the publications in (refereed) conference proceedings are in the top rank, and 47% of the journal publications are in the top rank, according to the ranking compiled by the unit. As it was the case for the previous periods, most publications appear in conference proceedings,

with the proportion of journal publications and book chapters accounting for about 23%. We consider these figures as an indication of the high quality and impact of the unit, which is further witnessed by the Honary Doctorate (Honoris causa) awarded by ENS Cachan, France, in 2007 to Professor Kim G. Larsen; the same year as he became Knight of the Order of the Dannebrog.

Overall Strategic Research Goals

We see our strong involvement in the ARTIST Design NoE with SC membership and coordinating role for the activities on Modeling and Verification as witnessing our leading European role within Embedded Systems. At the Danish national level, our unit has maintained and strengthened its role in Embedded Systems as coordinator of the projects DaNES, MoDES and the ICT-innovation network InfinIT. Together with Danish Technical University we have played a leading role in the successful Danish involvement in the ARTEMIS Joint Technology Initiative (1 project with Danish participation got funded out of 27 applications in ARTEMIS call 2008 and 4 projects with Danish participation got funded out of 42 applications in AR-TEMIS call 2009). Regionally, CISS is funded as an ICT-competence center with the obligation of introducing embedded software as vehicle for innovation to regional companies. At the university level, embedded systems is a strategic area, selected as the topic for elite eduction at the Master's level having been accredited by the Ministry of Education. Also, the strong interaction with research units from the Department of Electronic Systems has been maintained, witnessed by the joint Honorary Doctorate awarded to Professor Alberto Sangiovanni-Vintientelli as well as the continuation of the collaboration within the center MT-LAB.

Undoubtedly, the continuous improvement and support of the verification tool UPPAAL is the key factor in our internationally leading position with respect to verification and validation. During the period, various branches of UPPAAL have had a significant academic as well as industrial impact in areas such as planning and scheduling, testing as well as controller synthesis. The membership of the *Danish Center for Scientific Computing* is a clear indication of our leading national role in grid computing. Within concurrency theory our international visibility has shifted from fundamentals of process algebra to theories for infinite-state, mobile, timed, weighted and stochastic systems. Also, there has been a significant international interest in the formalism of modal transition systems where we contributed with closing a few long-standing open problems and provided new applications to interface theories.

The external funding obtained during the period 2006-2010 (65 MDKK) exceeds that of the previous period by more than 40%. In particular, the newly started center MT-LAB has secured significant funding to basic research activities. After the loss of three permanent staff members in the beginning of the period, the number of permanent staff is now almost back at the same level. Moreover, the unit has had a substantial growth in the number of assistant professors and post-docs, vital for the continuation and renewal of our unit.

5.7 Plans (2011-2015)

5.7.1 General Plans

- We want to strengthen our collaborations and current leading role within Embedded Systems both nationally and at the European level, e.g. involving creation and offering of a national set of PhD courses on Embedded Systems.
- We want to maintain strong collaboration with research units from the Department of Electronic Systems within Control Theory and MPSoC, overcoming the distribution in location.
- We want to maintain high internationally visibility within the areas of model checking and concurrency theory.
- We want to expand our leading position with respect to verification and validation of distributed systems by a clear strategy on our effort on tool development. This includes application of our existing mature tool engines to specialized tools (e.g. RT Java Workbench), establishment of a lightweight model checking tool box for easy prototyping, and a full exploitation of its potential by combining existing expertises (model checking, testing and static analysis) as well as bringing in new scalable techniques (constraint solving, SMT solving, statistical model checking and learning).
- The effort of the unit on GRID computing, currently active on the Nordic DataGrid Facilities, should focus more on the unit both research- and teaching-wise. This will allow the unit to engage in the emerging area of Cloud Computing and link to the existing Embedded Systems Platform.
- We will extend the domain of distributed systems considered in our formalisms, tools and methodologies from Embedded Systems into the direction of Multi-Core and Cypher-Physical Systems.

5.7.2 Concrete Plans

Concurrency Theory: We plan to further investigate the introduced formalism of energy (timed) automata. Currently, we extend the framework with multiple costs. We will also work on extending the existing techniques for finite state systems to infinite state ones and the work on visibly pushdown automata and timed extensions of finite automata and Petri nets will continue. We will develop a general theory of system metrics as a generalization of many equivalences and preorders with applications especially in real-time and hybrid systems. The theory of modal transition systems will be extended with the combination of time, cost and probabilities as well as with explicitly treating concrete data (by predicates). Larger industrial case studies and applications of timed I/O automata and the ECDAR tool will be considered. We plan to also develop a more abstract theory for time abstracted modal transition systems together with a tool support. In mobility and security we will study properties of type systems for process calculi as instances of general type systems for the psi-calculus.

Verification and Validation: In the model checking area the work is planned along several directions. We want to investigate further the benefits from combining model checking, static analysis as well as testing in order to improve the scope and scalability of verification. Concretely we are working on a new concept termed model checking with lattices. We are in the process of implementing statistical model checking for the full range of UPPAAL timed automata model. We will work on utilization of PC-clusters for the (parallel) analysis of parametric UPPAAL models. In static analysis the primary goal is the integration of the Coccinelle tool with other state-of-the-art program analysis tools such as Clang, the C frontend of the Low Level Virtual Machine (LLVM) project. Our long term vision in testing is an integrated testing theory and industrially applicable tool for testing quantitative system aspects (time, hybrid, probability and cost). Work will progress both on the tool integration side but also on more fundamental developments. The work on combining TRON and Phaver will continue and the capabilities resulting from the integration with Simulink will be fully explored. We will also study probablistic system models for the testing scenarios. In synthesis and scheduling direction we will work on probabilistic scheduling by applying duration probabilistic automata.

Networks and Operating Systems: We expect to increase the activities related to HPC. The DCSC installation at AAU is now well established and attracting more and more users. This provides a natural platform for collaboration with other departments on parallel algorithms, also algorithms running on GPGPUs. The collaboration with NDGF on grid computing will continue. The Nordic infrastructure provides a natural platform for experiments at large. The focus will be on how to handle computations on large distributed amounts of data. In Embedded Systems Platforms the work on defining resource-aware environments for high level programming programing languages (like e.g. Java) that support the development of embedded systems will be certainly continued. Also, as it is foreseen that it will become increasingly difficult to distinguish between embedded and service based systems, it is likely that the work on resource aware language platforms will include relevant aspects of support for service oriented architectures.

Embedded Systems Methodology: The work here is planned along two lines. Recent developments on methods and tools will be published as contributions to the 'Quasimodo Handbook'. Furthermore, the results from Quasimodo and DaNES will be further evaluated and published as part of the newly initiated EU ARTEMIS project RECOMP. The work on formal analysis and modelling of web services protocols as well as component based methodology will continue. We expect to conduct a similar number of case studies as in the previous period.

Staff and Funding: Given the large number of assistant professors and post-docs it is outmost important that adequate opportunities for tenure positions are made available at the departmental level. Also, there is a need to compensate for the drop of one full position in the teaching capacity. It is expected that the main source for PhD grants also in the future will have to come from external grants, and the unit will therefore continue the funding efforts. The previous period has been somewhat unbalanced with the high number (17) of dissemination projects, and the unit will in the future improve the balance.

5.8 Committee Evaluation

5.8.1 Strengths and Weaknesses

The research is internationally leading. The tool UPPAAL is world leader as a real time model checker, and its extensions and additions are gaining prominence.

The group has high international visibility and a very high level of external funding. The group has industry interaction, regionally, nationally and internationally.

People in the group have received many prestigious honours and awards.

The group has a very inspiring group leader. As a consequence, the group is also depending to a large extent on its leader. The group should think about a succession strategy.

We commend the strategy of the group. It has articulated policies concerning collaborations (within the department, and with other departments) and industry relations.

The group has too few permanent staff to handle all projects in which it participates, and supervise all PhD students.

5.8.2 Challenges and Recommendations

The group has a very high level of external funding, of three types: the Edison type (applied research), Pasteur (use-inspired basic research) and Bohr (basic research). The group needs the Edison type funding to keep up industrial relations, especially regionally, but wants to focus preferably on the Pasteur and Bohr types, and preferably larger grants without too much overhead. Overhead can be outsourced to some extent to an organisation like CISS.

Software tools are very important for the visibility of the group, and have a wide user base. Tool development of staff members should be valued just as publication and patent output.

Lab technicians as other engineering departments have are needed, who work on tool development.

The return of Josva Kleist to a full-time position will enable more research in the Networks and Grids area. In time, this could even turn into a new group.

The ABC division of journals and conferences in the report is compiled only from opinions within the group. It should be validated by external data, e.g. by Danish bibliometric data.

6 Information Systems

The Information Systems Unit.

6.1 Executive Summary

The Information Systems Unit conducts research on the development and use of information systems with focus on software-based systems. The research is based on and makes contributions to the intersection between software engineering, information systems and human-computer interaction. The results fall within three areas: methods, systems and practice.

The unit has strong international positions in research on usability evaluation methods, intelligent web information systems, interaction design and software development practice. The usability evaluation methods research has dealt with development of new methods and the interplay with software development; it has been based on two large, externally funded and several smaller projects. The intelligent web information systems research has dealt with methods, tools and techniques for intelligent information systems and web technologies; it has been based on several externally funded projects. The interaction design research has dealt with understanding use domains, design and implementation of prototype systems, and studies of these systems deployed into use; it has been based on a variety of projects. The software development practice research has dealt with systems development, software process improvement and e-government; it has been based on several large, externally funded and other projects. Research on interaction design methods, innovation and agile development are emerging as new and promising research areas.

The unit has been involved in about 20 externally funded research projects. About half of these have a significant budget, with funding mainly from national and EU sources, and the main applicant and project leader in most of these is a senior researcher from the unit. There are also about 10 internally funded research projects that have contributed considerably to the research in the unit. The unit has produced a total of 274 publications, including 177 peer-reviewed papers in journals and conferences. This research has had good impact in the relevant research communities as well as on software development practice in the companies that we have collaborated with.

6.2 Profile

The Information Systems Unit conducts research on the development and use of information systems with focus on software-based systems. The research field is development and use of computerised systems at two levels: humans and organisations.

The research in the IS Unit is based on and makes contributions to the intersection between software engineering (as traditionally known within computer science), information systems (as traditionally known within business administration) and human-computer interaction. Intellectual support is sought in adjacent disciplines: engineering, systems science, organisation theory, sociology, media science, theory of research, and philosophy. The shared research approach in the unit is empirical and experimental as well as theoretical, and it encompasses action research, longitudinal process research, case research, laboratory experiments, and field experiments. The research is based on qualitative as well as quantitative methods. The results fall within three areas: methods, systems and practice.

Methods: Research in this area deals with development of methods to support key processes in information system development. The basic research question is: how should professionals develop or evaluate particular systems or features? There are four streams of work in this area:

- Usability Evaluation
- Innovation
- Interaction Design
- Agile Development

The activities in this area focus on development of and work with relevant methods for each stream. There are often practical use of methods involved but with the aim of assessing or developing the methods.

Systems: Research in this area deals with development and evaluation of specific systems. Key questions are what the key features of a particular system are and how these may be measured and explained? The systems are developed and evaluated to inquire into specific aspects, e.g., the specific interaction between a user and an interactive computerised sys-

tem or the functionality of an intelligent web system. These activities usually involves an evaluation, e.g., a usability evaluation of a specific interactive systems in order to provide a basis for improving the design of these systems. There are two streams of work in this area:

- Intelligent web systems
- Interaction Design for Ubiquitous Computing

The activities in this area focus on specific systems and applications. They are developed and tested, often by using methods and in practice, but the heart of the research is the systems themselves.

Practice: Research in this area deals with the practice of developers and evaluators who are working in or related to software development organizations. The key question is how software is developed and evaluated in practice. This involves efforts to understand the field, e.g., a system in use in a context of activity (including organised activity), or professionals developing software.

- Software Development
- Software Process Improvement
- E-Government

The research in software development seeks to improve professional practitioners' ability to engineer systems as well as their ability to plan and manage effective social and organisational interventions. It deals with design and evaluation of whole software systems as well as the specific interaction between a user and an interactive computerised system. The research on software process improvement is specifically addressed towards the key topics in this area; including collaboration with software companies to understand and improve their software processes. Research on e-government tconcerns understanding and improving the design, implementation and management of the information systems that governments and public agencies use both for its internal administration and to communicate with citizens.

Across the three research areas, it is also within the unit's research interest to improve the research methods applied.

6.3 Staff

6.3.1 Current Staff

Professors: Peter Axel Nielsen, Jan Stage
Associate professors: Ivan Aaen, Jesper Kjeldskov, Jeremy Rose, Mikael B. Skov
Assistant professor: Jeni Paay
Post Docs: John Stouby Persson, Janne Jul Jensen
Ph.d-students: Anders Bruun, Lise Tordrup Heeager, Karsten Jahn, Kenneth Møller Nielsen
Adjunct professor: Steve Howard

6.3.2 Staff Development

Both Peter Axel Nielsen and Jan Stage were associate professors until February 2009 where they became full professors.

A significant part of the staff development during 2006-2010 has been first the recruitment of Peter Dolog as assistant professor (2006) and then as associate professor (2008) and later his transfer to Machine Intelligence (2010). After Peter Dolog joined the unit, participation in a number of EU FP7 projects were secured leading first to the recruitment of three new PhD students (2008) and later three research assistants (2009). Except for one PhD student these new recruits transferred to Machine Intelligence in 2010.

In 2006 Jeni Paay was recruited as amanuensis. She left the unit in early 2008 to work for CSIRO in Australia and rejoined as assistant professor in 2010. From 2008-09 Jesper Kjeldskov also worked for CSIRO in Australia while on leave from his position as associate professor.

In 2007 Steve Howard was made Adjunct Professor. During 2006-2008 the EU FP6 project on e-government employed Clive Sandford (2007-2008) as a senior researcher and Jesus Rios (2008) as a research assistant. Andreas Munk-Madsen was a temporary and part-

time associate professor until 2009. Gitte Tjørnehøj was a PhD student in an assistant professor-like 5-year position until 2010. In 2009 Janne Jul Jensen and John Stoubye Persson took up positions as Post Docs after finishing their PhDs in the unit.

6.4 Goals (2006-2010)

A main goal for the IS unit was to maintain the promising directions of the Systems Development and Human-Computer Interaction groups established in the previous period. Specifically this involved: (1) increased quality and impact of publications, (2) establishing a stronger research focus, (3) more international collaboration, (4) recruitment of more PhD students, (5) recruitment of more staff, and (6) stronger involvement with the research community.

It was also a goal to increase the collaboration between the two groups focusing on Systems Development and Human-Computer Interaction across projects. We planned to achieve this through joint IS unit seminars with open exchange of ideas and results, joint work on the topic "Mobile Systems", and implementation of the newly established research leadership roles in the unit.

Increase quality and impact. The aim was to maintain productivity but increase quality and impact by publishing a higher proportion of papers in A-level outlets and also more in journals rather than conference proceedings. However, the unit was also aware that exclusively pursuing top-tier journals (e.g., the AIS Basket of Six) would not necessarily lead to the desired results. Hence alternative journals with more appropriate audiences would also be prioritized. We also intended to increase practical impact through direct collaboration with a limited and carefully selected companies.

Stronger research focus. The unit has previously had a broad focus leading to the exploration of several promising directions, but the goal for 2006-10 was to focus more exclusively on the following areas:

- Information systems development
- Software process improvement
- E-government
- The interplay between usability evaluation and interaction design
- Mobile and context-aware systems
- Systems for the domestic domain

Extend international collaboration. We wanted to develop our existing international relations further. This should be achieved by involving ourselves more deeply in international research projects and by closer collaboration with a few selected international researchers.

Increase efforts with research education. We planned to increase the PhD production by recruiting more PhD students and completing four to six PhD degrees in the period.

Recruit more staff. The unit expressed a particular concern for its recruitment of more research staff. Based on an estimate of teaching as well as administrative tasks, the unit estimated that we would have to double its number of research staff. We wanted to do this gradually and in a long-term effort, and with a clear priority for assistant professors or postdocs in order to also lower the age profile of the unit. As a short-term strategy we wanted to make use of temporary research staff for filling the gap until long-term recruiting could show results.

Stronger involvement with the research community. We wanted to involve ourselves more in the research community by organizing conferences and workshops and by editing journals as well as specials issues of journals.

6.5 Activities and Results

6.5.1 Methods

The research in this area deals with development of methods to support key processes in information system development. The development and test of methods and techniques has been a main area in the research activities of the unit.

Usability Evaluation

This activity focuses on how to conduct useful usability evaluation and how to integrate such evaluations into the software development process. In particular, the activity has focused on the interplay between interaction design and usability evaluation by investigating tools and techniques for feedback and through studying early evaluations in software development projects. Furthermore, the activity has focused on how to reduce efforts in evaluations through new ways of analysing usability evaluation sessions and it has focused on different contexts for evaluation. Finally, the activity has focused on the particular challenges of conducting usability evaluations for web development.

The activity highly involves close collaboration with industry on adoption of methods and techniques for usability evaluation. The activity utilises a combination of field and laboratory experiments as well as action research and it applies both qualitative and quantitative measures.

Two major externally funded projects have contributed to this activity. The USE project focused on bridging the gap between evaluation and design in software development through development of feedback techniques and tools. The WPU project focuses on the particular challenges of creating usable web portals through the creation of new methods for usability engineering. Finally, a additional number of externally and internally funded projects have contributed to activity, e.g. the MAUSE project, the TwinTide project, the CUE-8 project, with different objectives within usability evaluation, e.g. training evaluators, reducing efforts in evaluations.

In the period, the usability evaluation activity has resulted in two completed PhD dissertations, 7 journal publications, 16 peer reviewed conference publications including two articles presented at ACM CHI, and 2 book chapters. We have co-edited a special issue of the International Journal of Human-Computer Interaction on the Interplay Between Usability Evaluation and User Interaction Design. Furthermore, we have co-chaired the Eighth Danish Human-Computer Interaction Research Symposium.

Innovation

Globalization and outsourcing calls for software development projects in high-cost countries to produce high-value solutions. This observation forms the point of departure for the Innovation activity. This activity focus on innovation and creativity at the level of the software developing team. The purpose is to develop methods, tools, and techniques as well as infrastructures and conceptual models to support such teams in producing valuable solutions throughout the span of a systems development project.

The activity involves one research project (SIRL/Essence) and one knowledge dissemination project (IKT Agil). A textbook for the Software Innovation course was produced as part of this activity.

The *IKT Agil Project* (February 2009 - January 2011) is funded by the European Social Fund. The project is part of CISS (Centre for Embedded Software Systems) and Cassiopeia Innovation. The project seeks to support companies in using agile development principles, organizational patterns and methods for innovation and creativity. The interactions included workshops in individual companies as well as 9 seminars for all participating companies. More than 20 companies and institutions participates in the project, and the industries involved include health, energy, building, games, administrative software, communication and more. The project is summarized in a report targeting a practitioner audience.

The *SIRL/Essence* project is internally funded. The project experiments with physical infrastructures and concepts aiming to develop methods, tools, and techniques for software team. The research approach is grounded primarily in Design Science with experiments on complex design challenges and evaluations based on artefacts produced and feedback

from participants. Concepts and infrastructures are developed in SIRL - *Software Innovation Research Lab* - and methods, tools, and techniques are developed to form a method framework named *Essence*. Central to SIRL and Essence is the use of physical views and team member roles. The approach taken is inspired by ideas on agile software development and aims to enhance and utilize the flexibilities known from agile development methods.

In the period, a total of 10 master thesis students have worked on Essence and SIRL, and versions of Essence have been applied in the Software Innovation course on several educations. The project has resulted in 6 peer reviewed conference publications, 1 journal publication, and a short book outlining the research area.

Interaction Design

These activities deal with development and assessment of methods for interaction design. This is still a limited area for the unit, with two main activities.

The first activity relate to the Object-oriented analysis and design method OOA&D that was developed in the unit between 1995 and 2005. In this period, we have developed an interaction design method that starts from the results of an object-oriented analysis and supports the process of developing an interaction design. The method is called ADRIA (Abstract Design of Rich Internet Applications), and it introduces interaction spaces as a general abstraction to detach interaction design from specific interface elements such as windows, widgets, etc.

The second activity has emerged from the Web Portal Usability project (WPU). This project is primarily focused on methods for usability evaluation. However, we have gradually realized that it is difficult to conduct a usability evaluation and subsequent collaboration on redesign without a repertoire of interaction design patterns for web portals. In the WPU project, we have developed and used one such pattern so far.

One externally funded project, WPU, one internally funded project, ADRIA, and one PhD project has contributed partially to this activity.

In the period, this research has resulted in 6 peer reviewed publications of which 1 is a journal article.

Agile Development

These research activities focus on methods for software development and in particular on how professionals can perform software development with agility. The activities aim to overcome main challenges with traditional development approaches, and the term 'agile' is interpreted in a broad sense.

One part of this activity has been concerned with extending agile development approaches. For example by extending the agile approach XP with a practice called developer stories as a supplement to user stories. Another example is the development of an agile approach to designing mobile applications. Along a slightly different path there has been the design and validation of a risk management approach for agile and distributed development teams.

These research activities also include work on usability evaluation, and in particular how it can contribute to agile development. The costs involved in conducting usability evaluations is a main obstacle against increased use of these methods in any software development practice; and it is even more of a hindrance in an agile development process. Two of our research activities address this obstacle: (1) development of methods that reduce the time and cost for specific activities in usability evaluation, and (2) development of programmes for training software developers to conduct usability evaluations in an agile manner, that is quickly and at low cost.

The research in this area activities in this area have been quite limited. It has involved part of two PhD project, and the internally funded the Simplified Usability Evaluation Project.

In the period, this research has resulted in 8 peer-reviewed publications including 2 journal publications.

6.5.2 Systems

Research in this area deals with development and evaluation of specific systems. Key questions are what the key features of a particular system are and how these may be measured and explained?

Intelligent Web Information Systems

This activity is influenced by world wide web research. It studies and contribute with methods, tools and techniques for *intelligent information systems and web technologies*. In particular, the focus is put on infrastructures for intelligent information systems on the web, web personalization strategies and algorithms, and engineering adaptive web. Infrastructures allow intelligent and personalized access to information and services, information exchange and coordination of conversation between distributed services. Strategies and algorithms take a knowledge about a user and his context into account when recommending information and services. Engineering methods concern how the experiences from web personalization strategies and algorithms as well as infrastructures can be transferred into guidelines for engineering adaptive web applications. The research is applied in e-learning, e-health, and e-business application areas so far.

This activity involved several externally funded projects. As an associate partner, the activity members have been involved in the EU IST/FP6 network of excellence Prolearn. The personalization strategies with focus on tag based recommenders have been studied in EU ICT/FP7 KIWI project where personalized recommendations have been implemented and experimented with in semantic wiki environments as well as in the EU ICT/FP7 MEco project where personalized recommendation are studied in the context of medical epidemic surveillance. Further, infrastructures and engineering methods have been studied besides EU ICT/FP7 KIWI project also in EU ICT/FP7 idSpace project which featured an infrastructure for creativity support for distributed teams. In non funded project FROGS this activity has been involved in strategies and an infrastructure for long running web service transactions with human involvement where different compensation strategies have been researched and infrastructure developed which is hosted as an open source project.

In the period, this activity has produced 7 journal articles including IEEE Transactions on Service Computing, ACM Transactions on Internet Technologies and ACM Transactions on Web articles, 25 refereed conference and workshop papers, 2 book chapter, 2 books, 1 edited special issue in journal of web engineering, and 4 edited proceedings from workshops. Besides, the activity resulted in number of algorithms and components, as well as contributed to several infrastructures which are now hosted as open source projects.

Interaction Design for Ubiquitous Computing

This activity focuses on the design of interactive ubiquitous computing systems to support the way people communicate and interact in their everyday lives. The activity involves empirical studies for understanding use domains, design and implementation of prototype systems, and studies of these systems deployed into use. A common characteristic for the projects within this activity is that they explore non-desktop technologies, including mobile devices, situated displays, in-vehicle systems, and large interactive multi-touch surfaces. Another common characteristic is the exploration of use situations outside the work domain.

The approach taken is largely exploratory and creative, grounded in ethnographic inspired studies of current practice and coupled with both qualitative and quantitative evaluations in real-world settings.

The activity has involved several externally and internally funded projects. The *indexical interaction design* and *e-Spective* projects focused on interaction design for mobile systems and devices, mostly within urban settings. The *Loca-Play,* and *interaction design for children* projects focused on supporting communication and interaction within families and across generational and geographical boundaries, mostly within the home. The *in-vehicle interaction design* project focused on the challenge of interacting while driving, the *pervasive shopping in supermarkets* project focused on exploring emerging technologies in supermarkets for creating new shopping experiences, and *dual-sided user interfaces and blended spaces* projects focused on exploring new emerging display technologies for supporting co-located and distributed multi-user interaction.

In the period, the Interaction Design for Ubiquitous Computing activity has resulted in 14 journal publications, including two articles in ACM TOCHI, 34 peer reviewed conference publications including four articles presented at ACM CHI. We have also guest-edited a special issue of Springer's Journal of Personal and Ubiquitous Computing on the topic of Pervasive Computing in the Domestic Space. We have chaired an international conference on Interaction Design and Children (IDC 2006), and edited the proceeding of two international HCI conferences with interaction design themes (OzCHI 2006 and 2009).

6.5.3 Practice

Research in this area deals with the practice of developers and evaluators who are working in or related to software development organizations. The key question is how software is developed and evaluated in practice.

Systems Development

This activity focuses on studying empirically how software systems are developed and implemented in practice as it occurs in organizations, whether these are software companies or use organizations. In this activity we take a broad approach and study issues and topics which are likely to concern the professional systems developer or manager and the scope is often larger than developing the software. It includes the study of:

1. Development teams that are geographically dispersed because of outsourcing measures.
2. How systems developers' competence influences their use of methods and hence also their practice.
3. The adaptation and implementation of enterprise systems and how different user groups influence the adaptation and implementation over time.
4. The interaction between an agile software project and a large organizational development project.
5. Development of mobile applications and how it differs from traditional systems development.
6. How agile approaches like Scrum and XP are used in practice.

7. The use of an agile development method in developing clinical software-based devices under the constraints of a traditional quality management system.
8. The deployment of usability evaluation methods into a software company.

The theories applied as theoretical lenses in the analysis of the empirical data include: competence theory, control theory, coordination theory, organizational influence theory, and actor-network theory.

In the period, the activity has produced 26 publications of which 5 are A-level journal articles and 1 is an A-level conference publication. Two PhD students have graduated from this activity -- one of which are through the collaboration with Agder University.

Software Process Improvement

This activity is a continuation of a long-term research activity with a particular focus on how software development is improved in software companies. The research is linked to what internationally is called software process improvement. Software process improvement is generally based on both a dominating idea of what a software process is as a prescription for how software engineers should work and is is based on equally dominating models for assessing the degree to which a company's software practice is compliant with a set of process requirements, e.g., Capability Maturity Model. The unit's research has taken a critical stance to both the dominating assumptions.

One part of this activity has brought closure to a large national research project, SPK, during 2001-2003 where software process improvement and its relationship with knowledge management was studied in three Danish software companies. The purpose has been to publish based on the empirical data collected during this project. There has been a strong emphasis on understanding and explaining how software process improvement is performed in software companies and how different theoretical lenses can be utilized to explain what promote and what demote improvement initiatives. The applied theories include: technology adoption, social network analysis, sensemaking, knowledge management, competing values framework. The SPK project was fundamentally an action research project and a few of the published results are hence methods explaining how to practice a part of software process improvement, e.g., how to use social network analysis.

A second part of this activity has been seeking to exploit previous acquired understanding of the linkages between software process improvement and knowledge management in the KiWi project. The unit's responsibility was to adapt a generic technology for a semantic wiki (the KiWi platform) to be useful for process managers and for project managers in a large software company. This has resulted in documented requirements, a knowledge model and two software systems in addition to indirect contributions to the KiWi platform. While the KiWi project will close by the end of 2010 the more substantial results will first be published after 2010.

In the period, the activity has resulted in 19 publications of which 5 are published in A-level journals and 2 in A-level conferences. One PhD student have graduated from this activity.

E-Government

The activity responds to a strategic need for the university to develop in this area and consists both of networking and project-related research work. The university's Centre for Digital Governance was a cross-disciplinary network set up with university seed funding to develop the area in co-operation with the region's government institutions and agencies. It produced a book outlining the university's contributions in the field and a major European co-operation: DEMO-net (the European network of excellence in eParticipation) involving 21 partners in 19 countries. eParticipation involves the engagement of citizens in government through the use of web and mobile-based technologies. The objective of DEMO-net was to strengthen scientific, technological and social research excellence in eParticipation by integrating the research capacities of individuals and organisations spread across Europe in accordance with strategic goals set by the European Council. The IS unit worked both on establishing the theoretical foundations and research direction of the emerging research field, and on the design and management of web-based eParticipation systems. A spin off of the unit's work in these areas is the DISIMIT project. DISIMIT (digital service integration through effective management of IT in Danish municipalities) aims to bring Danish municipalities into the forefront of digital service integration through effective management of IT. The goal is to promote effective and efficient electronic local government through the introduction of improved IT management models and principles. DISIMIT is a cross-disciplinary

collaboration between researchers at AAU and 10 Danish municipalities, funded by the Danish Research Council. It has two more years to run.

In the period, this activity has produced 9 journal articles (5 'A'), 8 conference articles, 1 edited book, 2 book chapters and 2 scientific reports. We chaired the European Workshop on eParticipation in Venice (2006) and co-chaired for the DGO workshop 'Understanding eParticipation.' Three PhD students have graduated and a fourth has submitted. The IS unit has in this area built an international network of collaborators.

6.5.4 Research Education

This activity has cut across all of the above as it has addressed the improvement of research education in the whole unit. The activities has been part of the national project DaRSIS for IS research education and the activities have included: PhD summer schools, short courses, workshops for supervisors, international guests, and travel grants.

6.6 Own Evaluation

In this section, we compare our activities and results to the plans we made five years ago and provide an overall evaluation of our research efforts in the period. The comparison is structured in accordance with the original goals (see section 4).

Increase quality and impact. The basic aim was to prioritize quality and impact over productivity. We have produced a total of 274 publications that are classified according to the scheme in the appendix as follows. The total for the previous period was 208.

	2001-2005	2006-2010
A journals	20 (14%)	33 (18%)
B journals	19 (13%)	24 (14%)
A conferences	34 (24%)	21 (12%)
B conferences	33 (23%)	49 (28%)
C conferences	36 (25%)	50 (28%)
Sum	142	177

The distinction the unit has made between A journals and other journals corresponds by and large to the Danish national bibliometrics. The A conferences are distinguished from the rest by a similar criterion, yet at the same time conferences are generally not as difficult to publish in as in journals. Three out of five journal publications are in A journals. One out of five conference papers are in A conferences. Every third peer reviewed publication is in a journal and not a conference. Under the assumption that A journals have higher quality and impact that other journals and that A conferences have higher quality and impact than other conferences the unit has published well and has fulfilled its goal. Specifically, the unit has increased its production (142 to 177); but more interestingly has moved from conferences to journals and in particular to A journals.

So far we can measure an increase in quality. It is too early to measure impact on research with an increase in citations. Several of the unit's publications are well-cites (above 50), but these are typically published some years ago. A few publications from 2006-10 have a citation rate above 10 cites/year. There are signs of an increasing impact on research, but it is still too early to see the full impact of the changed publication strategy.

Impact on practice has been pursued by the unit for many years. That effort has been maintained at a high level by collaborating in research projects with software companies and user organisations, by actively disseminating to practitioners, by taking leadership roles in innovation platforms (Mobile Systems, InfinIT, and Cassiopeia Innovation), and by publishing in journals which also have practitioner audience, e.g., IEEE Software.

Stronger research focus. The aim was to focus on fewer selected research areas. As key areas, we identified the interplay between usability evaluation and interaction design, mobile and context-aware systems, systems for the domestic domain, information systems development, software process improvement and e-government. The description of activities and results (see section 5) corresponds mostly to these areas. There are two areas, mobile systems and domestic computing, where we have changed direction towards the more general areas of interaction design and ubiquitous computing. Thus the unit has fulfilled this goal.

The description of the research profile of the unit (see section 2) has been developed in discussions across the unit over the last couple of years. It represents a more coherent view on the unit compared to the original goals, and it also reflects success in terms of strengthening the research focus.

Extend (international) collaboration. We have during 2006-10 extended the unit's international reach and collaboration. There are strong links leading to co-authored papers with Agder University, Georgia State University and University of Melbourne. There have been visits of 3-6 months at the two latter universities and involving several PhD students as well as senior staff. The appointment of Steve Howard as adjunct professor is also a result of this closer collaboration.

Collaboration through EU projects has been increasing and has led to co-authoring within web intelligence systems while in other activities it has not yet led to published research.

Collaboration with other units in the department and with research units in other Danish universities has been expanded. This has led to significant co-authoring of papers.

Increase efforts with research education. We have by far reached the stated goal of educating four to six PhD students during 2006-2010. Eight have completed during the period, one has dropped out, five are ongoing and on schedule.

During this period the university has changed its strategy on PhD education and dramatically increased the goals for how many PhD students the department should en-

roll and complete. In response to this strategic goal-setting the unit has increased its effort in fundraising for PhD scholarships. The unit has been successful in fundraising both as the main applicant and as co-applicant (i.e., another unit or institution manages the project). The implication is that more PhD students have been employed; but it has implied also that it has not always been possible to recruit qualified applicants to scholarships despite considerable search effort. The positive side of this is that it has as a consequence been possible to transfer research funding from scholarships to postdoc positions. This has served the purpose of providing more visible career opportunities for PhD candidates hence easing the recruitment efforts at the lower steps of an academic career with the unit.

With more PhD students to educate, i.e., supervise and provide learning environment including relevant PhD courses, it was advantageous for the unit to enter into the more structured national collaboration on IS PhD education, the DaRSIS project. This has been very beneficial for those supervisors and PhD students who have participated.

Recruit more staff. Much effort has been put into recruiting more scientific staff. The unit has in particular focused on recruiting staff at the lower levels of the career ladder with an explicit goal of recruiting more PhD students and consequently also more postdocs and assistant professors.

Peter Dolog was recuited as an assistant professor and later as an associate professor. The intention was to expand the units profile in software engineering and in particular in a more technical branch hereof. Dolog also brough considerable experience with coordination of EU FP6 projects and a large European network. During 2010 it became clear for all that Dolog's research profile were close to research in the MI unit and a move was hence established. This has been beneficial for the department and for the MI unit and temporarily for the IS unit. It also shows how the department as a whole dynamically can expand into new areas. The IS unit's recruitment strategy has from an internal viewpoint been less successful. It has, however, led to a rethinking of the IS unit's profile and the description of the three themes (see above).

On the positive side is that the unit's postdocs and assistant professors are growing in numbers (currently three) and is gradually renewing the unit's research profile yet with a strong linkage to the three themes (see section 2).

Stronger involvement with the research community. The unit has been successful in terms of stronger involvement in the research community. The appendix provide lists of community services that reflects this stronger involvement.

Funding. This was not explicitly mentioned in the goals for this period. The appendix lists 21 externally funded projects, and some of them have significant budgets. For most of these projects, a senior researcher from the unit is the main applicant. However, we are also becoming involved in projects as partner/co-applicant. Several projects involve collaboration with national and international partners, which has been very useful for the development of the unit. The key sources are the Danish Strategic Research Council and the European Union. Overall, we have been reasonably successful in attracting external funding.

Several of the externally funded projects originate from projects that were internally funded in the previous period. In this period, we have had 10 internally funded projects which is promising as potential areas for external funding in the next period.

6.7 Plans (2011-2015)

Quality and impact of publications. The goal is to maintain the current publication strategy and then also pursue publishing in outlets with higher standards and with larger audiences. This will require careful selection of outlets. The unit is already experienced in publishing in A-level quality outlets and this will be maintained. Increased impact will come as a consequence of this and when in addition we begin targeting larger audiences - both among researchers and among practitioners. This strategy will be pursued even if it reduces the total number of publications from the unit.

Strategic research project. The goal for the next period is to establish a strategic research project that spans across the whole unit. The purpose is to create a setting for collaboration across the unit and to exploit the individual research competences in building a research effort that will become leading in our field. The centre of attention will be the area of digital public service. The tentative title is "Digital municipality 2020", and design and innovation are the key focus areas. This project will build on several existing projects, in particular DISIMIT and WPU.

Extend international research collaboration. The goal is to extend our collaboration with our existing international and national partners leading to closer research collaboration and joint research funding. In particular we want to capitalize on the extensive effort so far invested in networking and forming partnerships.

Research education. The goal is to maintain a high level of involvement in research education. This will be done in terms of PhD courses, supervisor workshops, and collaborate nationally to improve PhD education. It will also be done by working on the learning environments we provide for both PhD students and for postdocs.

Recruitment of more staff. The goal is to recruit more scientific staff at the lower levels of the career stages. We should specifically be able to fund and recruit a total of 6 PhD students during 2011-2012 raising to 8 PhD students by 2015. We should also be able to fund and recruit 3 postdocs/assistant professors during 2011-2012 raising to 4 by 2015. The goal is also to be able to offer positions as associate professor.

Research funding. The goal is to maintain the current level of national funding. We have been successful in setting up national projects that include collaboration with other research units in Denmark. In the next period, we aim to acquire a similar amount of national funding. The EU funding has mainly been secured by Peter Dolog, who has now moved to the MI Unit. The goal is nevertheless to gradually increase the amount of EU funding for the other senior researchers in the unit.

6.8 Committee evaluation

The Information Systems group has a strong national research track record, with significant international profile and activity. There are excellent international collaborations (e.g. with Steve Howard at Melbourne) that have resulted in valuable joint publications as well as researcher exchanges. One of the key strengths of the group is its focus on practice-focused research, with a strong motivation to improve information systems development practice. Importantly, there is strong interest from companies in working with the group.

Like other groups, IS reports on key research in "Category A" journals and conferences. The classification is closely based on the Danish National rankings.

There is a strong action research tradition in the group. Members of the group have clearly developed, reflected on and communicated research methods well. This feeds through into research training for students. Another great asset for PhD students is DaRSIS, which provides a forum for PhD students as well as supervisors to interact.

6.8.1 Challenges and opportunities

One of the challenges, and opportunities, for the group is to crystallise its identity, which is currently poorly articulated even though there is a strong theme of understanding and improving practice (from several angles including usability, efficiency, reliability, and delivering novel solutions).

A second challenge for the group is to locate themselves better on the international scene. One aspect of this is to recognise and articulate the relationships between the work of the group (e.g. on innovation) and that of other groups internationally (e.g. on creativity, team software development and argumentation in design rationale). Another aspect of this is to develop strategies for presenting work in the "Scandinavian tradition" in ways that are better accepted by international (particularly US) audiences. A third aspect is to consider which work is appropriate to an international audience and which is

not; for example, work on "Digital Municipality 2020" might be of limited international interest unless framed in ways that have relevance beyond a local audience. A final aspect is to create a clearer identity so that any researcher anywhere in the world will seek out Aalborg IS work when it is relevant.

A third challenge is to balance the desire to have more junior staff and the desire of senior group members to keep doing research rather than supervising and administering it.

6.8.2 Recommendations

Probably the most important recommendation is that the group should better articulate and market its identity. This should improve the research impact (as others better understand and value the work of the group), guide strategy (allowing the group to better identify priorities), motivate individuals, and provide new opportunities for collaboration. A memorable name and logo can be valuable ways of communicating identity.

There are further potential opportunities for collaboration with other groups, as well as externally to the department. The group needs to reflect on which collaborations are of greatest value, avoiding the risk that it could become a "service group", supporting others without conducting leading-edge IS research itself.

The growing focus on e-government presents great opportunities for high impact research, but the group needs to better articulate what are the important issues in this domain (that differ from other domains, for example). If they develop the theme of the "Digital Municipality 2020", they need to better articulate the value of the theme and what leverage it provides for the group, the individuals within it, and the broader department / university.

For a group that works with human participants (as end users of systems and as developers of systems), it is important to have a clear policy on ethics to ensure that all work is safe for both researchers and participants, is not exploitative (e.g. of student effort without recognition), is respectful of participants, and is honest. A lightweight but clear ethics policy should be developed.

One question that was raised during the meeting was whether it is indeed valuable to develop and disseminate methods, since industry do not generally use them in vanilla format. The view of the panel is that methods are the basic structures for communicating and learning that are necessarily adapted to context, and that they are of clear value as both research outputs and objects of study.

Finally, the group needs to continue its practice of reflecting on its strategy for seeking funding, ensuring and measuring impact, communicating with researchers and practitioners, educating students and developing staff.

7 Machine Intelligence
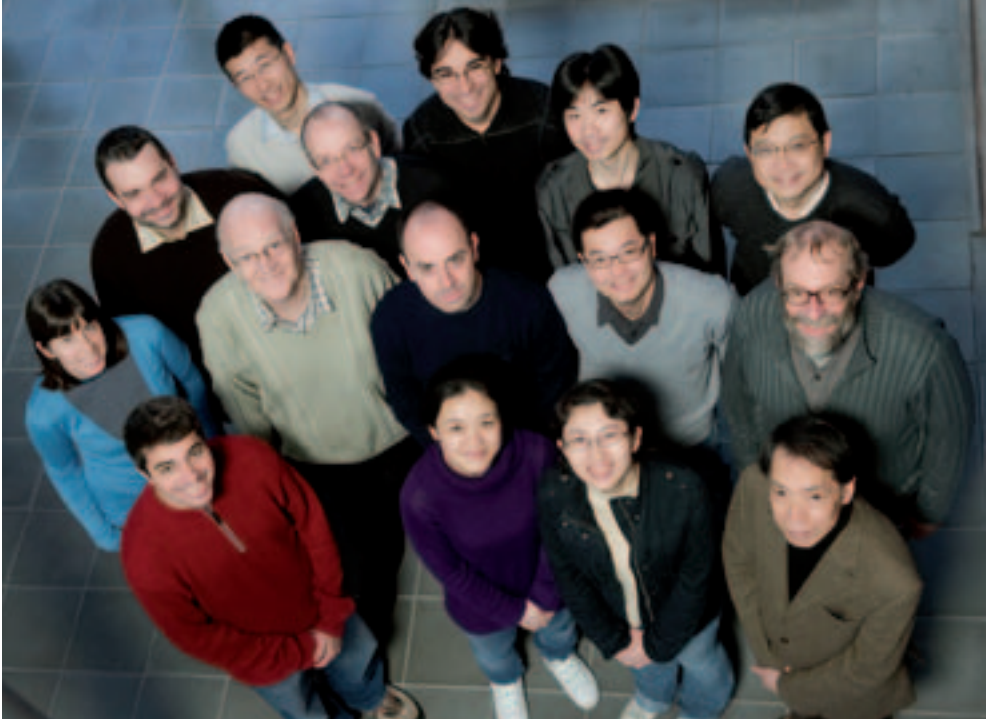
The Machine Intelligence Unit.

7.1 Executive summary

The Machine Intelligence unit is internationally recognized for its research on intelligent systems, such as decision support systems that help human operators to make optimal decisions in complex problems, systems that provide web-based services in an intelligent and personalized manner, and systems for intelligent data analysis.

Much of the unit's research centers around the question of how to handle uncertain and incomplete information in such systems. The unit develops mathematical models, computational methods, and prototype implementations and applications. In the period 2006

- 2010 the unit has published 87 scientific papers, many of them in leading international journals and conferences. The Machine Intelligence unit is very well integrated into the European research community: members of the unit collaborate closely with several European research partners; they have organized in 2008 the main European conference on probabilistic methods in Artificial Intelligence, and the unit receives many international visiting researchers. Members of the unit also are active as associate editors of scientific journals, and they receive numerous invitations for invited talks and seminars.

The unit is engaged in scientific and commercial applications of its basic research, and is actively pursuing collaborations with industry, in particular with several small software companies that have been created as spin-offs of the Machine Intelligence unit.

7.2 Profile

The Machine Intelligence research unit conducts research on intelligent learning, reasoning and decision making under uncertainty. Over a period of about 20 years the unit has been best known for its contributions to the field of probabilistic graphical models, especially Bayesian networks and influence diagrams. Members of the unit have contributed to laying the foundations of the field, and to making its techniques known to a wider audience in many different application areas. A standard textbook authored by members of the unit has accumulated (in several editions) over 4500 citations in the scientific literature. More recently, the unit has actively pursued a strategy of broadening its research, and has successfully established new research profiles in the areas of machine learning and autonomous agents. In 2010 the Intelligent Web and Information Systems unit, formerly associated with the Information Systems research unit, has joined the Machine Intelligence unit, thus adding a new dimension to the unit's research. The unit now conducts research in four main areas:

Probabilistic Graphical Models: Developing efficient design and inference methods for graphical models; in particular, frameworks for representing and solving complex decision problems under uncertainty.

Machine Learning and Data Mining: Learning from data for knowledge discovery and design of intelligent systems. The focus is on the use of probabilistic and statistical methods, as well as the use of logic-based methods for modeling complex, structured data.

Autonomous Agents: Programming intelligent behavior in autonomous agents. The focus is on the use of graphical models, and applications for autonomous agents in computer games.

Intelligent Web and Information Systems: Methods, tools, and techniques for intelligent information systems and web technologies, especially personalized and adaptive provision of information and services on the web.

7.3 Staff

7.3.1 Current Staff

Professors: Finn V. Jensen

Associate professors: Peter Dolog, Manfred Jaeger, Uffe Kjærulff (on administrative leave), Thomas D. Nielsen, Kristian G. Olesen (on administrative leave), Yifeng Zeng.

Post-docs: Guandong Xu.

Ph.d-students: Yingke Chen, Fred Durao, Jorge Cordero Hernandez, Ricardo Gomez Lage, Hua Mao, Thorsten Ottosen, Rong Pan.

Research assistant: Karunakar Reddy Bayyapu.

7.3.2 Staff Development

Within the reporting period one assistant professor has left the unit (Olav Bangsø), and one assistant professor has moved up to a position at the associate professor level (Yifeng Zeng). Three PhD students graduated (Jens D. Nielsen (10/06), Søren H. Nielsen (09/06), Nicolaj Søndberg-Jeppesen (12/09)). The most significant change came in August 2010, when the members of the Intelligent Web and Information Systems unit, formerly associated with the Information Systems research unit, joined the Machine Intelligence unit (Peter Dolog, Guandong Xu, Karunakar Reddy Bayyapu, Fred Durao, Ricardo Gomez Lage, Rong Pan).

A new assistant professor (Paolo Viappiani) will join the unit in January 2011.

7.4 Goals (2006-2010)

A central goal formulated in the 2001-2005 research report was to broaden the research activities of the Machine Intelligence unit. Building on its international reputation in the field of probabilistic graphical models the unit wanted to establish itself in the areas of data mining and autonomous agents. At the same time inner coherency of the unit was to be maintained by centering the research activities in all areas around a common reference point of probabilistic models of uncertainty.

Another important goal was to increase the number of research activities, especially those involving collaboration with European and international research units, and with industrial partners. In order to achieve this goal, external funding would have to be procured from European or industrial funding sources.

As the relevant success criteria for the unit's research the 2001-2005 report mentions standard metrics, such as publication and citation counts, as well as indirect indicators of the unit's impact and recognition, e.g. the number of foreign visitors received by the unit.

7.5 Activities and results

7.5.1 Probabilistic Graphical Models

The Machine Intelligence unit's research on probabilistic graphical models involves frameworks for representing and reasoning about domains endowed with uncertainty (e.g., the Bayesian network framework, which provides an intuitive and compact representation of a probability distribution) as well as frameworks for modeling and solving decision problems under uncertainty. Our research in this area is directed at improving the usefulness of these frameworks by lifting limitations of existing types of models in order to increase their range of applications. This requires to develop new exact or approximate algorithms for inference, learning, and decision making in these frameworks.

Much of our work in the areas of machine learning and autonomous agents also contains methodological contributions to the general theory and practice of probabilistic graphical models.

Frameworks and Solution Algorithms for Sequential Decision Problems
The influence diagram framework serves as a powerful modeling language for decision problems with a single decision maker. However, efficient use of the influence diagram framework imposes some rather severe constraints on the decision problem being modeled. For instance, the decision problem should be symmetric (e.g., future decision options may not depend on past decisions and observations), and all utilities should be specified on the same scale (e.g., be defined as monetary values).

In order to extend the class of problems that can be modeled efficiently, we have proposed more general types of graphical specification languages. This includes models for representing asymmetric decision problems as well as so-called multi-currency influence diagrams that support utilities on different scales (for example, in environmental engineering we may need to balance bio-diversity with fishing quotas, or the use of fertilizers).

Solving a decision problem formulated using the influence diagram framework, or using the more general languages mentioned above, is computationally very hard in terms of both time and space. These difficulties have motivated the development of both approximate solution algorithms as well as flexible algorithms for making trade-offs between the time and space required to solve a problem. 4 international conference and journal papers have been published on this topic.

Decision Theoretic Troubleshooting

Decision theoretic troubleshooting combines Bayesian networks with cost esti- mates to obtain a probabilistic cost optimization problem for repairing a man- made piece of equipment. A solution to such a problem is an optimal decision making procedure that takes into account all future ramifications of all decisions.

It is known that finding optimal troubleshooting strategies is a computationally intractable problem, but that for troubleshooting scenarios that meet specific requirements solutions can be computed in practice. Requirements needed by previous solution methods include the assumption that a single action can only repair a single fault, that the cost of an action is independent of previous actions, and that no diagnostic evidence is collected during the troubleshooting process. We have extended the scope of practically solvable troubleshooting problems by developing solution methods for scenarios that do not fully obey these requirements. This research is carried out in collaboration with *Dezide ApS*. This work is documented in 4 international conference and journal papers.

Learning Bayesian Networks from Sparse Data

Most conventional learning methods for Bayesian networks run into computational and statistical problems when they are applied in domains with many random variables for which only relatively little data is available. This is often the case, for example, in biological domains, social network analysis, or stock market analysis.

We addressed some of the challenging issues in learning a large Bayesian network from a small dataset, by embedding attribute clustering techniques into a pipeline-like learning framework, and by using a decomposition algorithm for the structure construction. The new learning algorithm first finds local components from the data, and then builds

the complete network by joining the learned components. We have shown the empirical performance of our proposed algorithm in several well-known benchmark networks. Our results are published in 4 international conference papers.

7.5.2 Data Mining and Machine Learning

The Machine Intelligence unit conducts research which is both directed at automated extraction of useful information from large datasets (data-mining), and at using such information to build more intelligent computer systems (machine learning). Our work is primarily methodological, and has a large number of possible application domains, including, e.g., scientific data (DNA sequence data, environmental or meteorological measurements), and data automatically generated by human-computer interactions (log files of an online shopping portal or search engine). Our focus is on the use of probabilistic and statistical methods: a dataset is seen as the product of a process involving randomness, and the main task for a data mining or machine learning procedure is to construct a probabilistic model for this process that is as accurate as possible, that provides insight into the process, and that can be used to build computer systems for more intelligent interaction with the uncertain processes or environments in the future.

Mining and Learning with Structured Data

Traditional statistical methods have been developed for the analysis of data tables that consist of a list of observations, where for each observation a small number of observed attributes is given. Much of the data that is generated today does not follow this simple structure. More frequently the data consists of a relational database in which several tables are linked via references from one table to another. Also data consisting of sequences (e.g. DNA), or spatially structured data (many environmental datasets) plays an important role. Designing and learning models for such complex structured data requires new probabilistic modeling approaches that combine elements of traditional statistical models, relational data representations, and mathematical logic. The Machine Intelligence unit has been at the forefront of these developments with the introduction of the Relational Bayesian Network modeling language, and the Primula software. The most recent version of Primula was released in July 2009. It implements a variety of sophisti-

cated probabilistic inference techniques, and a very powerful new method for estimating the numerical parameters of the probabilistic-relational model.

One of the main challenges in learning from relationally structured data is to determine the relevant features for the underlying probabilistic models. For example, when learning a model for customers' preferences from a relational database of transactions, one may need to determine that a relevant feature for predicting a user's interest in a particular item is the number of times the user has viewed other items in the same product category. We have investigated the problem of automatically constructing such features. The *Type Extension Tree* framework was developed to provide a rich representation language for relational features. A method for learning Type Extension Tree features from data was developed. Experimental results show that this learning method successfully determines features that are both interpretable (and, in fact, sometimes correspond to features which in earlier works on the same data sets were manually constructed), and lead to models with high predictive performance.

We published a total of 9 journal articles, book chapters and conference articles on this subject. Furthermore, we co-edited a special issue on this topic of the *Annals of Mathematics and Artificial Intelligence*.

Learning from Incomplete Data
Most data sets one encounters in real-world applications are incomplete, i.e., the values of some attributes are missing in some observations. Incomplete data poses fundamental difficulties for the reconstruction of the underlying probabilistic data generating process, especially when the observed values may present a systematically biased picture of the underlying distribution (e.g., when values are missing because of a data censoring mechanism). Almost all current learning techniques are based on the assumption that values are missing due to a random mechanism that does not take the actual values into account (the *missing at random* assumption). The Machine Intelligence unit has conducted extensive research both on the foundations of learning from incomplete data, and on the development of computational techniques that enable learning from incomplete data without making the missing at random assumption.

Specifically, we introduced the *AI&M* (Adjusting Imputation and Maximization) algorithm as an alternative to the prevalent *EM* (Expectation Maximization) algorithm. Theoretical analysis shows that AI&M maximizes a profile likelihood function that in contrast to EM and the missing at random assumption is determined by the absence of any assumption on the missingness process. Experimental results obtained with an initial implementation of AI&M show the potential of this new learning paradigm for inducing more accurate models when the missing at random assumption does not hold. Furthermore, it has been shown that a comparison of the results of EM and AI&M learning can be used as the basis for a statistical likelihood ratio test for the missing at random hypothesis. Two international conference papers were published on this topic.

Learning in Hybrid Domains

Domains containing both discrete and continuous variables represent a challenge for probabilistic reasoning. The main difficulty is to find a representation of the joint probability distribution over both types of variables that supports efficient inference. Computationally, (exact) inference algorithms require that the joint distribution over the variables of the domain is from a distribution-class that is closed under summation and multiplication. The mixtures of truncated exponentials (MTE) framework represents one such distribution class. Besides supporting exact inference, the MTE framework has several other desirable properties, such as being able to approximate any distribution arbitrarily well and allowing discrete and continuous variables to be treated in a uniform fashion during inference.

Methods for learning MTE distributions from data have received only little attention. Initial attempts have mainly been directed towards parameter estimation, where the most prevalent methods look for the MTE parameters minimizing the mean squared error wrt. a kernel density estimate of the data. Unfortunately, there is no guarantee that the least square estimation procedure will find parameter values that are close to the maximum likelihood (ML) parameters. This has a significant impact when considering more general problems such as model selection and structural learning, since many well-established approaches to these learning tasks assume ML parameter estimates to be available.

We have been working on maximum likelihood-based learning algorithms for MTE distributions. The developed algorithms cover learning from complete as well as incomplete

data, and empirical experiments have shown that the proposed algorithms can offer significant improvements in terms of likelihood as well as in generalization ability. Our research in this area is published in 5 international conference and journal articles.

Classification in Complex Domains

Classification is the task of predicting the class label of an item based on a set of attributes describing that item. Example application areas include credit scoring, spam filtering, and process monitoring. When learning a classifier we seek to establish a classification model based on a database of labeled instances. A popular classification model is the naive Bayes model, which is easy to learn but based on rather strong and often unrealistic independence assumptions (i.e., that the attributes describing an item are conditionally independent given the class of that item).

We have developed a class of latent classification models (LCMs) that relaxes the independence statements of the naive Bayes classifier. The models have been specified for both discrete and continuous domains, where they extend the naive Bayes classifier by introducing latent variables to model conditional dependences between the attributes. Algorithms for learning these types of models have been developed, and experimental results show that models provide very good classification accuracy. The results of our research here are published in 3 international conference and journal articles.

7.5.3 Agents

The MI unit conducts research on decision making in multi-agent systems, such as multi-player (computer) games or teams of robots. The design of decision models and solution methods for such scenarios faces particular challenges when there is limited communication between the agents and/or the different agents have conflicting goals. The classical approach of solving a multi-agent decision problem by computing a Nash equilibrium has severe shortcomings, since it does not provide decision rules for interacting with other agents that do not act optimally, for example due to their limited computational resources. We have developed multi-agent decision models that are based on the influence diagram framework, and that incorporate explicit models of other agents into an

agent's decision model. Much of our research is devoted to solving computational complexity problems arising out of recursive agent modeling, and the modeling of sequential decision processes over many time steps.

Simple Sequential Bayesian Games

Simple sequential Bayesian games is a class of games in which two players take actions to change a shared environment. Each player has received a specific assignment which he/she must fulfill, and which is not known to the opponent.

Three problems have been addressed in our research. The first problem is how to estimate the opponent's strategy. That is, one player tries to predict the other player's behavior, who in turn is trying to predict the first player's behavior, who in turn …. Rather than calculating Nash equilibria for this situation we provide solutions which exploit shortcomings in the opponent: if the opponent does not play Nash equilibrium he/she will suffer a disadvantage. Our approach is recursive modeling based on influence diagrams, and we developed a general framework for solving simple Bayesian games. With this technique, each player may have different levels of intelligence in terms of look-ahead and depth of recursive analysis.

The second problem is the combinatorial explosion caused by a large look-ahead. The fact that all information is used only for estimating the opponent's assignment is exploited. By assuming that one will know the assignment exactly in the future, a considerable reduction in complexity is achieved (the technique is called information enhancement).

The third problem addressed is how a wrong assumption about the opponent's level of intelligence may influence the outcome of the game. Using a mixture of models make the player adapt to the actual model of the opponent. Five international conference and journal papers document our research in this area.

Interactive Dynamic Influence Diagrams

We consider a typical decision scenario where an intelligent agent shall act in an uncertain environment shared with other agents who may have either common or conflicting

objectives. The agent might make a change to some feature of the environment that in turn impacts other agents who may possibly act in similar ways. The agent needs to optimize its policy given what it believes on both environmental states and other agents' behaviors. If communication does not occur between agents, the uncertainty of other agents will accumulate over time when they act and receive private observations. Solutions to decision problems of this type require both a formal representation and efficient algorithms for solving the model.

We extended the influence diagram (ID) model to a new representation called interactive dynamic influence diagram (I-DID). We construct an I-DID from a single-agent's perspective and represent other agents' decisions as chance nodes in the diagram. Since the modeling agent does not have sufficient knowledge for predicting other agents' behavior, it reasons with a set of possible models of other agents and gets the probability distribution of their behavior by solving the assumed models. Furthermore, an I-DID explicitly models how the modeling agent updates its beliefs about other agents' models when other agents act and observe over time. We have formalized an I-DID using clear syntax and semantic features and shown its interesting applications in some multiagent settings like non-cooperative agents.

We also investigated inference algorithms for solving an I-DID. Since an I-DID is mainly built on influence diagrams, most standard algorithms for graphical models could be used to solve the model. However, predictions about other agents' future actions must be made using models that change as agents act and observe. This is reflected by the fact that the number of candidate models of other agents grows exponentially with the number of time steps. Hence we are facing the main challenge of solving all other agents' models at each time step. We have proposed several inference algorithms (including exact and approximate ones) to efficiently solve an I-DID. The basic idea is to limit the model expansion by exploiting the equivalence of other agents' behavior. For instance, we may maintain only one representative model for other agents if some models exhibit the same policy across the total time horizon. We have demonstrated the comparable performance of various I-DID algorithms on academic benchmarks and shown the achievement on the solution scalability.

An important application area of I-DID is opponent modeling in multiagent systems. Interacting with its opponent, an agent has an opportunity to update its beliefs of the opponents' models and eventually focus on the correct model. Consequently, the agent will be able to predict its opponents' behavior therefore optimizing its decisions. We have successfully shown the practical utility of I-DID on computer games, robotics, 8-card poke game and other domains.

Related to the I-DID theme, about 10 academic papers have been published in interna-tional conferences and journals. This research has greatly contributed to establishing our research profile in the field autonomous agent.

Opponent Modeling in Computer Games

In real-time strategy (RTS) games human players interact with intelligent computer-con-trolled opponents, which learn and adapt to the player's strategy. The human players exhibit certain types, such as "aggressive", "defensive", or "resource". We applied our *interactive influence diagram* (I-ID) model and the learning techniques developed for this model to enable a computer opponents to detect a player type online by updating beliefs over possible player types. We investigated opponents' adaptivity in the strategic level of RTS games and examined the learning of specific actions.

We also developed an Artificial Intelligence engine as a prototype add-on to game en-gines, and implemented an I-ID based learning component within. On this topic we pub-lished 4 research papers in international conferences and journals.

7.5.4 Intelligent Web and Information Systems

The research activities in Intelligent Web and Information Systems for the period 2006-2010 is described in section 6.5.2 Systems.

7.6 Own evaluation

The unit has made good progress with regard to its strategic goal of establishing a more diverse research profile. It is now well established in the areas of machine learning and autonomous agents. The recent integration with the Intelligent Web and Information Systems unit is another step in this direction.

The unit has maintained a high level of publication activities: in the period 24 journal papers, and 63 peer reviewed conference papers were published. A high level of quality in these publications is witnessed by the fact that 58% of the journal publications were in journals that belong to the the high quality "level 2" of the Danish national journal ranking scheme.   The papers published in the current period have already accumulated a total of over 370 citations (Google-scholar). Furthermore, the MI unit has organized a European conference on Probabilistic Graphical Models in 2008, and members of the unit are associate editors or members of the editorial board of 6 international journals. The visibility and impact of the unit is also attested by the fact that we have hosted 9 long term visiting researchers, and 8 visiting scientists staying for a shorter term. Members of the unit have accepted 22 invitations for invited talks, seminars, or Ph.D. lectures abroad. Although these numbers do not directly reflect the quality and volume of the scientific performance, we take them as evidence that the unit has maintained and extended its international reputation in the field of probabilistic graphical models and beyond.

The unit is very satisfied with the development in the number of PhD students. Whereas at the beginning of 2006 the unit had 3 PhD students, this number is 7 at the end of 2010. The increase is due to the joining with the Intelligent Web and Information Systems unit, and to a bilateral agreement with the University of Electronic Science and Technology of China. Under this agreement, the Department of Computer Science provides tuition-free PhD enrollment, whereas the students are financed by a stipend from the Chinese government. Currently, two PhD students are employed under this agreement.

The unit is not entirely satisfied with the results of its efforts to increase collaboration with, and knowledge transfer to industry, and to establish externally funded research

projects. In the report period, the unit has submitted 9 national and 6 European research project proposals, either as a partner or project coordinator. Unfortunately, none of the proposals received funding.

7.7 Plans (2011-2015)

7.7.1 General Plans

The overall plan for the next years is to consolidate the successful broadening of the unit's research activities, and to achieve additional research capacity and impact through externally funded international and national research projects. A key component in this plan is a fruitful integration of the Intelligent Web and Information Systems sub-group with the Machine Intelligence unit. Areas of common interest, notably recommender systems, will form natural focus areas of research.

7.7.2 Staff Development

The MI unit currently has 4 permanent staff members, 2 permanent staff members on administrative leave, and one member in a non-permanent position. In 2011 a new assistant professor (Paolo Viappiani) will join the unit. His research (recommender systems) will strengthen the link between the traditional research areas in the MI unit, and the new intelligent web and information systems area. Since the two members on administrative leave are not expected to return as active researchers for (most of) the period 2011-2015, the unit will consist of 6 active researchers as of 2011. This is at the lower level needed to maintain and consolidate the research activities in the current four main areas, and the unit will therefore aim for a moderate growth to about 7 or 8 active researchers at the assistant, associate and full professor level.

7.7.3 Probabilistic Graphical Models

We plan to continue our research activities related to frameworks and solution algorithms for sequential decision problems. More specifically, for the initial part of the coming period our research plans coincide with our current activities on approximate inference in decision graphs. In parallel with these activities we aim at further developing the influence diagram framework (or variations hereof) to extend the class of problems that can be modeled and solved efficiently. This includes solution algorithms for influence diagrams modeling hybrid domains as well as constrained solution algorithms that provide robustness guarantees for the solutions found. The former issue is also related to our current and planned research activities on learning in hybrid domains; in particular, the use of mixtures of truncated exponentials or mixtures of truncated polynomials, which allow discrete and continuos variables to be treated in a uniform fashion.

7.7.4 Data Mining and Machine Learning

A central goal for the next years is to highten the impact of our methodological work by application oriented research in the context of externally funded research projects.

For the research related to relational models we aim at applications in the domain of environmental modeling and management, especially in the area of water management, where based on the past involvement in the MERIT project the unit has good connections to prominent domain experts. An application for a Marie Curie Training network in this area is currently in preparation. Environmental modeling will also provide a useful context for further methodological developments, especially the integration of continuous variables with the existing discrete variable models supported by the Primula system. The Type Extension Tree framework will be developed further with the goal of developing it into a full relational modeling and learning platform that outperforms other frameworks on benchmark datasets, with a possible focus on applications in biological domains.

Concerning the research on learning in hybrid domains, we have recently become a member of the project *Mining data with probabilistic graphical models: new algorithms and applications.* The project is funded by the Spanish national research council and is

expected to start in 2011. In the project we will mainly be involved in methodological developments of the MTE framework, including structural learning as well as more foundational issues such as translating from known distribution families to the MTE framework. The methodological developments will be connected to a diverse set of application areas ranging from environmental modeling to the prediction of arrival times for public transportation.

A collaboration with the Distributed and Embedded Systems research unit on learning system models for verification has been started in 2010. We plan to maintain and extend this collaboration, with the goal of exploring a variety of new research questions arising at the intersection of verfication, embedded systems, and machine learning.

7.7.5 Intelligent Web and Recommender Systems

The main target areas in forthcoming years are adaptive web and recommender systems. The unit would like to extend the started activities on tag based recommender systems and extend them with the work from data mining, clustering, and social based trust computing.

We plan to further strengthen the multi-factor recommendation research where different aspects of user activities are taken into account in order to calculate recommendations such as user's contributions, user tags, users documents authored, user queries and so on.

The data mining and clustering will focus on discovering associations between recommendation items, users, and user groups. By doing so, we would like to improve especially performance but also traditionally precision and recall of the currently available algorithms in the area of recommender systems and in the context of specialized data sets such as for epidemic purposes.

An initial investigation has shown that part of our research on classification in complex domains can also be applied in the area of recommender systems and collaborative filter-

ing. We plan to explore this connection further, and, continuing our ongoing collaboration with the Norwegian University of Science and Technology, conduct a project aimed at developing a latent probabilistic graphical model that will incorporate user and item information into a single coherent model. This will e.g. provide support for group recommendations as well as greater robustness when dealing with new users or items.

7.7.6 Agents

Sequential decision making is a broad research area in multiagent systems. We will continue the previous work on I-DIDs and explore relevant research avenues on intelligent agents. More specifically, we are planning to work on the following three topics: More efficient solutions to I-DID, Communication strategies in multiagent systems, and Human-like non-player characters (NPC) and robots.

We will continue to address the challenges on solving I-DIDs and scale up the solution to larger horizons. Our previous algorithms are mainly built on the behavior equivalence of other agents. The resulting model space is still unbounded. We intend to further reduce the model space using the utility-equivalence concept. We will find a subset of models, which will result in the same expected utility of the modeling agent's policies as when we include all possible models of other agents. More efficient algorithms will be developed following the idea on either behavior or utility equivalence.

Another goal of our future research is to develop communication strategies for multiple agents when they interact in a common environment. This work will generalize our current solutions to multiagent decision problems. We plan to extend the I-DID framework to model agents' communication in their interaction. The new representation will become richer, but it will cause additional difficulty in solving the model. We will aim at a set of algorithms for approximating the solutions.

Our practical research will focus on designing human-like NPCs and robots. We have successful experience on the application of opponent modeling techniques in computer games and robots. We will apply the multiagent learning in both areas, and integrate

the learning techniques with opponent modeling methods in the real applications. We plan to design an adaptive NPC that would interact with human players in an appropriate manner in a real-time strategy computer game. We also plan to construct a mobile robot that would show social behaviors in a complex navigation task.

7.7.7 Intelligent Information Systems Engineering

Central goal for intelligent web and information systems area in the next 5 years is to deepen and strengthen ongoing activities.

The unit plan to further study infrastructures on the web. This includes infrastructures where the aforementioned recommendation approaches are implemented and applied as well as infrastructures for web service based integration.

The unit would like to maintain its contributions to the web information systems engineering areas which specifically includes different methods for realizing and designing web applications.

7.8 Committee evaluation

During most of the evaluation period, the Machine Intelligence unit was clearly the smallest unit of the four, with two permanent staff members on administrative leave, and was about half of the size of the other units (with respect to permanent staff). The unit is internationally well known for its central role in establishing probabilistic graphical modelling as one of the key research areas in Artificial Intelligence, and for its seminal work in Bayesian Networks and Influence Diagrams. The panel was pleased to learn about progress in new areas like multi-agent modelling and relational models, which build on existing expertise and extend the earlier focus in a logical manner. However, this expansion of the research focus now calls for more resources in the research personnel: although the research output per individual researcher is quite good, the overall volume in each of the research areas is currently not very high, and the visibility and impact of the work is not at the level it could be.

7.8.1 Challenges and Opportunities

The unit has tried to compensate for their lack of resources by collaborating more with foreign colleagues, with clear success, and they report several joint publications, some co-authored with leading experts in the world. The unit has also made several attempts to acquire external funding, but with no success during the reporting period. The reasons for this are not known. The lack of external funding has also led to a lack of Ph.D. student positions, and the output of the unit in this aspect is not very good.

The unit and the CS department have identified the problems mentioned above, and as an attempt to improve the situation, the IWIS team from the IS unit was transferred to the MI unit at the end of year 2010. This move sounds promising as the IWIS team not only brings in new researchers, Ph.D. students and already running externally funded projects, but it also has a portfolio of interesting application-oriented research topics that fit well together with the modelling techniques developed in the MI unit. The challenge of course is how well the new team integrates with the old unit members, and at this point it is too early to predict how well this organizational move works.

Another challenge the unit will obviously face in the not too distant future is how they will manage when the only professor in the unit, who is the most senior and most promi-nent researcher of the group, retires. Although the operational management is already renewed and the current group leader is also a well-known researcher in the same field, it is not clear how the forthcoming retirement will affect the group, especially concerning its internationally recognized status and visibility.

7.8.2 Recommendations

Another way to compensate for lack of resources would be to join forces with the other units of the department. There have not been many activities of this type, but the panel was pleased to hear about emerging collaboration with the DES and IS units. However, we would also like to point out that there are also many other obvious collaboration op-portunities with these units, as well as with the DPT unit, and we recommend the units

to explore more possibilities for collaborative work that would benefit all the parties.

The MI unit is determined to continue their efforts in acquiring external funding. This is advisable, and the panel wonders whether the senior members of the group who are currently administrative leave, could help the unit in analyzing the reason for their past failures, and help in finding ways for improving the application writing processes? The IWIS team may also improve the situation in the future as they have a good track record in writing successful project applications. At the same time the panel would like to re-mind the MI unit to plan carefully in which areas the project applications should be aimed at. Application-oriented projects can be most useful tools in motivating and stimulating the more basic-research type of methodological work, if the research topics are well chosen. For example, although robotics offers an obvious application domain for agent-based modelling techniques, this area is clearly outside the core of the CS department, and the panel suggests the intelligent web and software agents as a more promising application area for this type of work.