

# **Helsinki Institute of Information Technology (HIIT) Review**

26-28 May 2008

Randy Katz, University of California Berkeley  
(Editor)

Alberto Apostolico, University of Padua  
Christos Faloutsos, Carnegie Mellon University  
Bengt Jonsson, Uppsala University  
Martin Kersten, CWI Amsterdam  
Kari-Jouko Rähkä, Tampere University of Technology  
Mart Saarma, University of Helsinki Institute of Biotechnology  
Angela Sasse, University College London  
Martin Vingron, Max Planck Institute for Molecular Genetics Berlin

## **Executive Summary**

The Scientific Advisory Board (SAB) finds that the Helsinki Institute of Information Technology (HIIT) has emerged as a proven and effective center-based model for scientific innovation in computer science and its applications within the modern university. We were impressed with the Institute's cross-disciplinary research activities of computer science at the interface of the scientific and societal domains. In our view, the return on the Universities' investment justifies their joint funding and the complex institutional challenges that have been necessary to make the Institution a success.

Our key summary recommendations are (1) The Institute should develop a strategic plan for the next ten years of its existence; (2) External communications and public relations should become a high priority in the near-term; and (3) the existing management structure is working, but a new organization and leadership structure will likely be needed to take the Institute to its next level of development and achievement.

## **1. Introduction and Review Process**

The Helsinki Institute of Technology (HIIT), a joint research institute in computer science founded by the Helsinki Technical University (TKK) and the University of Helsinki (UH) in 1999, held its third Scientific Advisory Board meeting 26-28 May 2008. The SAB was constituted from international experts in the underlying technical areas of research expertise within the Institute, and its membership was expanded to address some shortcomings in its expertise encountered in the prior reviews of 2003 and 2004. The SAB reviewed the Institute's progress in terms of its organization, funding, and research achievements. Areas of review were based on four unifying programmatic themes: Probabilistic Adaptive Systems (PAS), Algorithmic Data Analysis (ADA), Network Society (NS), and Future Internet (FI). The individual program leaders covered each program in an overview presentation, with supplementary short presentations highlighting group activities, and a poster session of individual researcher projects. The SAB met for three hours to draft its observations, findings, and recommendations, which

were reported to the Vice Rectors of the sponsoring universities on the morning of the third day. The SAB report was briefed in an open session to interested members of the Institute for feedback, clarifications, and corrections.

## **2. Institute Assessment**

Our summary assessment is that HIIT is a strong and rapidly growing research institute, with internationally recognized researchers. It enjoys an excellent level of funding, and is highly productive in its research output. The Institute's activities can be broadly grouped into two nicely balanced categories: (1) analytical techniques applied to a variety of scientific domains and (2) technology-developments coupled to empirical "in vivo" evaluations to assess the societal context. The high degree of cross-disciplinary research in many areas is impressive.

The level and integration of research activities since the last review is impressive. Clearly the Institute has left its adolescence, and is now developing into adulthood. The Institute's activities have reached a critical mass, and its cross-fertilization across disciplines yields increased visibility and attracts technical talent.

The amount and quality of the Institute's activities are comparable to those of the key information technology institutes throughout Europe, such as those that are the founding members of the European Consortium for Informatics and Mathematics (ERCIM). In reality, HIIT is a more appropriate partner for ERCIM than VTT. This as a positive development, and indicates that HIIT stands among the first rank of such institutes within Europe.

HIIT offers a successful model for inter-university joint ventures in Finland and represents an effective organizational collaboration between TKK and UH. The Institute offers an intriguing model for how UH can collaborate with the newly forming Innovation University that will incorporate TKK in the near future.

## **3. Institute-wide Observations**

### **3.1. Strategic Planning**

We believe that the Institute must undertake a strategic planning exercise, to address the challenges and new opportunities it will face as it moves to its next stage of development. HIIT's success implies that it is likely to persist for many years to come, and now is the time to consider how it should develop from its current solid foundation. In formulating this plan, the Institute should identify its process for how to evolve its current portfolio of research projects, striking an appropriate balance between opportunistic research driven by funding consideration and new longer-term strategic directions driven by a clear scientific agenda.

A positive attribute of the Institute that has contributed to its success is its strong technical leaders for its programs. However, this makes the organization dangerously

dependent on a small number of individuals. We are concerned about a lack of succession planning for senior program leaders. It is unclear how young researchers are being groomed to become leaders in their own right, leaving the organization fragile to the loss of key individuals. Succession planning is part of the process of become a maturing research enterprise, planning for a long term future.

While our sense is that Institute's age distribution is appropriately biased towards early career researchers, it is important that the Institute should track the age distribution of its technical staff as part of its assessment of the health of the Institute, and to identify candidates for future leadership.

We observe a surprising lack of women among those who presented to the SAB, suggesting that there are few women among the Institute's senior technical leadership. This is especially surprising given Scandinavia's justifiably renowned record of gender equity in professional fields. The Institute should assess whether there is any inherent gender bias in the way it operates. Furthermore, providing enhanced technical opportunities for women should become a priority for the Institute.

As part of its planning exercise, the Institute leadership should consider the limiting factors in its size and organization. Is the Institute size and composition limited by management span of control? Or is research funding the limiting factor? What about space, or the ability to attract further technical talent to the Institute? Further, the process by which Programs, Groups, and individual Projects are assessed, old activities ended and new ones launched to evolve the Institute's research portfolio is unclear. While less pressing for a new Institute, this becomes a major issue for a mature and established research organization. It is particularly difficult to end activities that are no longer at the edge of the research frontier. Similarly, to foster new research directions, the Institute should have a process for determining new expertise and how to acquire it.

As a thought process, suppose that the Institute's strategic planning process identified parallel computing as important expertise to add to its portfolio of competences. How would this be acquired? Would it be developed from within the Institute's current staff? Or would it be attracted from activities at one of the sponsoring universities? Or perhaps a collaboration could be forged with other scientific activities within Finland, e.g., the national supercomputer center? Any strategic plan should also consider how new areas of research are to be funded. We can imagine two approaches: (1) by investing internal opportunity funds to start work in the new area, or (2) seeking external funding, and once secured, attract new researchers to work in that area.

### **3.2. Institute Visibility and Communications**

The SAB notes that while HIIT's individual researchers are well known internationally, the Institute is not as internationally visible as it *deserves* to be. HIIT should consider how to craft a "corporate identity," and undertake a campaign to make HIIT better known internationally. This could take the form of an external communications program, generating press releases describing the Institute's research successes and placing general

interest articles describing the Institute's projects in venues spanning from *IEEE Spectrum* or *Communications of the ACM*, to international newspapers like the *Financial Times* or *The Economist*. The Institute can join with any efforts by the Finnish government to publicize high technology in Finland to an international audience, to insure that the Institute is mentioned in any such campaign.

We understand that the Institute has recently hired a media communications professional. We strongly support the Institute's investment in this kind of professional expertise that can help raise the international recognition of the Institute.

Just as important is a program of internal communications, to enhance a sense among the staff of belonging to a world-class Institute. Institute management should carefully consider how all aspects of the Institute's activities can present a consistent image, even down to such subtle ideas of having a common poster and presentation format that prominently displays the HIIT logo.

The concept of the program appears to provide a useful tool for structuring the Institute's activities and communicating them to the outside world. The group level view is too complicated and fragmented for this purpose. However, in presenting the programmatic structure of the Institute, we continue to see the need for a clear, non-generic mission statement for the Institute and its constituent programs.

The Institute should put the posters of its individual investigator and small group research projects on line as part of its web site. It is a good idea to make these available before a review, and use events like external reviews to force posters to be updated. Putting this material on-line is an excellent way to make visible the broad and deep range of the Institute's research activities.

### **3.3. Assessing Impact**

Research has "impact" if it is so important that others depend on it as the foundation to build on for their own success. This definition applies equally to conceptual and theoretical work as it does to engineering developments. While the Institute is justly proud of its record of high quality publications in selective venues, it should also report on its significant technical contributions in other forms.

For example, contributions to technical standards are an important kind of high impact activity for a technology-oriented institute. The development of useful software codes and tools, especially when these gain wide-acceptance outside of the Institute, is also important. Modern web technology makes it easy to track the number of downloads and their geographic spread. These should be tracked and reported to assess the influence of the Institute's software artifacts.

Since a major activity of the Institute is the training of young researchers, their careers should be tracked as well, with a particular focus on graduate students, postdoctoral researchers, and junior researchers. Leaving the Institute and moving to other positions in

Finland and internationally should be seen as marks of success. Similarly, as the Institute attracts increasing numbers of international students, postdoctoral researchers, and research visitors, this will be a sign of its growing international reputation.

### **3.4. A Better Accounting**

To insure an adequate in depth assessment of the Institute, it is important that clear and precise information be provided to the SAB. Since we consider the training of students to be an important metric of the Institute's success, it is particularly important to distinguish between research staff and actual graduate students working in the Institute towards their advanced degrees. We understand that this is not a simple request—a high percentage of the Institute's staff believes they are working towards an advanced degree. However, it is important to distinguish among them those making true progress towards their degree goals, and to report these numbers.

The increasing numbers of Ph.D.s granted to students affiliated with the Institute since the last review is a positive development. Nevertheless, given the facilities and concentration of excellent researchers within the Institute, we expect even higher levels of production of advanced degree holders, and will expect evidence of this in future reviews.

The introduction of a program-oriented structure providing greater internal visibility to budget issues is also a positive development. Nevertheless, we understand that most of the Institute's funding is associated with specific projects and is allocated to groups. It would have been useful for the SAB to see how the University's funding—presumably the most flexible—is being allocated as opportunity funds to support new activities.

### **3.5. Comments on the Review Process**

The organization of the review of the Institute's research programs could be improved. During this review, the focus on the high-level programmatic view, juxtaposed with the detailed individual researcher poster prospective, misses the important group and project-oriented view. The latter level of review is particularly important, because research groups are where the detailed technical work is performed.

We understand that it is difficult to review all of the Institute's research in a relatively small amount of time. One alternative is to reserve the first day of the review for overview and highlight talks, and to split up the committee to permit parallel detailed reviews at the group-level on the second day. Alternatively, a review schedule can be developed so that programs are evaluated, down to the group level, once every other year. This would allow, for example, two programs to be reviewed in detail every other year.

### **3.6. Comments on the Poster Sessions**

We were impressed with the enthusiasm of the poster presenters, their ability to describe their technical work, and their diversity, particularly in terms of the number of women

and foreign researchers associated with the Institute. The marked increase in the number of foreign graduate students compared with the 2004 review reflects the growing international reputation of the Institute and its senior researchers. The project demonstrations were also very impressive.

### **3.7. Institute Funding Profile**

The increase in and diversity of the Institute's funding sources is an excellent development. The Institute is doing well in competitive research programs, and its core expertise makes it a highly sought partner for proposals in response to funding opportunities. This success in winning external funding is another healthy sign of the rising reputation of the Institute and its senior researchers.

The diversity of funding sources—University, Finnish Academy, Tekes, Industrial, and European Union—makes the Institute reasonably resilient to shifts in funding. Less clear is the implications of this funding profile for the mix of basic versus directed research conducted within the Institute. We suspect this is not uniform over all programs and groups within the Institute. A focus on too many short term funding opportunities could compromise the longer term scientific vision of the Institute. In particular, more information should be provided on how University funding used to seed new opportunities.

## **4. Program Observations**

### **4.1. Probabilistic Adaptive Systems**

In this program, we assess the Institute's activities as excellent. The groups have a strong publication record, and its researchers have assumed important leadership and organizational positions within their technical fields. The Neuroinformatics Group is very strong, with an established track record in Independent Component Analysis (ICA). Its new direction of using ICA for causality is very promising. The Statistical Machine Learning and Bioinformatics group provides an excellent linkage between the PAS and ADA programs.

Despite this obvious excellence, we are motivated to ask the question, "What is the Helsinki school of PAS?" This is to encourage the Institute to articulate its particular strengths and unique approaches to research in this area. The mission statement for this program is too generic, and needs better focus. There are a large number of groups working under this program, and the overarching research themes and how they inter-relate was not adequately clarified.

In our own view, the Institute's strength is in its approaches to modeling sources. The use and extension of MDL is a critical underlying thread being leveraged by the research groups. MDL is a critical enabler and is particularly appropriate as a distinguisher, as it solves many of the general problems faced by machine learning efforts.

Given the strength of the program, we were surprised that it has produced so few Ph.D.s over the last six years. However, we were pleased to learn that there are several students likely to emerge from the Ph.D. pipeline in the near future.

We are concerned about the detachment of Henry Tirri and the departure of Wray Buntine, who were formerly leading figures in this program. What new technical leaders are emerging? How are the topic areas evolving under their leadership? For example, we see the winding down of the information retrieval activities is clear and appropriate, with the build up of other activities. The strategic direction of this program was left unclear.

#### **4.2. Algorithmic Data Analysis**

This program has one of the clearer mission statements: “Useful data analysis methods for other sciences and industry.” Nevertheless, we ask the question “What is the Helsinki school of ADA?” in an effort to understand how the Institute perceives its unique strengths and intellectual approach. In our view, the program’s key focus is its work in algorithmic pattern extraction and combinatorial matching. The program’s methodological approach is to formulate computational concepts in strong collaboration with domain experts. This is an excellent example of deeply mathematically-based yet “use-inspired” research.

This world-class program is under the leadership of a widely recognized scientific leader. It is successful in attracting large number of postdocs, which helps it scale the research effort. Its diverse sources of external funding indicate the quality of its scientific activities. The collaboration with domain experts is impressive, and provides further evidence of the impact and usefulness of the algorithms being developed.

We applaud this outward focus, seeking inspiration from application domains to derive and study important computational primitives. This is all too uncommon in the computing field. Nevertheless, we feel that there are some potential opportunities in computer science that are being overlooked. One research direction is to develop a methodology for evaluating whether algorithms are the “best possible” for the job, perhaps identifying unifying underlying principles that could drive new algorithm development. Another research opportunity is to investigate how the algorithms can be parallelized to address the issues of scaling to much larger data sets.

#### **4.3. Network Society**

We find that the work in this area has improved. It is distinguished by a considerably more disciplined approach than that which we encountered in our 2004 review. This program has a good mission statement, and its focus on “end-to-end research” is appropriate. It has developed a solid track record of success. The number and quality of publications is excellent, and its work is appearing in the best venues for user interface and computer-human interaction research. The collaboration with industry and its sponsorship is excellent, and indicates the strong attraction of the research program. The

involvement of legal and economic expertise in the program's projects is a positive development.

We were impressed by the demonstrations, but the presenters were generally unable to clarify how their efforts contributed to the larger scientific vision of the program. New researchers, who were not intimately familiar with the work they were presenting, presented several of the demonstrations. Sometimes they were unaware of the related literature. We suspect that this is a program undergoing a transition, with some turnover of the research staff. Furthermore, the demonstration-driven nature of the work gives the appearance of being driven either by the sponsor who provides funding or by the desire to develop and apply a new technology. Research of this kind should also make use of a user-centered design methodology.

#### **4.4. Future Internet**

The Future Internet program has a strong and well-established reputation in network transport, host mobility, and security based on host identity. It contributes strongly to the international reputation of the Institute.

The presentation of research within the program by lines (or themes) was not particularly effective. Specific projects are associated with funding, and groups are collections of researchers under the direction of a technical leader. Lines are essentially functional crosscuts; in the case of the Future Internet program, these are described as transport, mobility, energy, and security/privacy. To us, it was unclear how lines are mapped onto the groups that will do the actual research work. A great strength of the Institute is its ability to deploy testbeds for experimentation. A missed opportunity is to exploit the "living laboratories" of the Network Society Program for testbeds in which to pursue Future Internet research. This is an essential way for the Institute to distinguish itself from other Future Internet efforts.

The criterion for clustering groups into programs was never clearly articulated. We could well imagine that the security work presented under Future Internet could be comfortably performed within the context of the Networked Society program. Many of the security/privacy issues faced by current and future networks are more a question of usability and user-centered design, rather than focused on new technology development.

The Future Internet program strikes us as one that is in transition, with a reorganization that is still in progress, and for which new research directions are yet to be determined. We are concerned that the proposed future directions have the appearance of being evolutionary rather than revolutionary. Since "Future Internet" is such an internationally competitive activity, it is critical for the Institute that a unique and important direction be identified for the program. A large management group—perhaps too large—is trying to chart these next steps. We suggest that a smaller, more focused group might be more prudent.

## **5. Responsiveness to Last Review**

The Summary Recommendation from the 2004 Review was as follows:

“Move the Institute’s governance model towards a truly joint venture with joint vision and mission statement that spans the independent units of the Institute. Establish transparent funding and accountability procedures under the leadership of a single director with an institute-wide budget that s/he can use to successfully direct the research in order to achieve the goals of the joint vision.”

In May 2008, we find that this Institute is more integrated and communication that integration much better than in the 2004 review. The mission statements have improved, but some tuning is still needed. The dual-directorship structure appears to be working effectively. Furthermore, the steering/management group structure does enhance budgetary transparency, but we would have liked to understand in more detail how opportunity funds are being directed within the Institute.

## **6. Summary and Conclusions**

The Helsinki Institute of Information Technology has emerged as a proven and effective model for organizing a center for scientific innovation within the modern university. The Institute exhibits impressive cross-disciplinary research of computer science with the other scientific as well as societal domains. In our opinion, the return on the Universities’ investments—in terms of world-class research, impact, and success in external funding—justifies the joint funding and the institutional challenges that were necessary to make the Institute a success.

Our Key Summary Recommendations are the following:

1. Develop a strategic research plan for the next ten years of Institute’s existence.
2. Improve external communications and public relations to raise the visibility and awareness of the Institute in the near-term.
3. Carefully consider changes to the existing Institute management structure, as generally it appears to be working well.