Centre of Excellence in Computer Science

cs.ioc.ee/excs

Tarmo Uustalu, Institute of Cybernetics at TUT

Seminar of Estonian CoE programme, Tartu, 16 April 2010

EXCS in one slide

- A CoE on computer science and computational sciences (bioinformatics, human language technology)
- Formed of four target financed themes (TFTs) at three institutions:
  - Trustworthy software and human language technology (T Uustalu) — Institute of Cybernetics at TUT
  - Security of information systems (A Buldas) — Cybernetica
  - Methods, environments and applications for solving large computational problems (J Vilo) — University of Tartu
  - Natural language processing for Estonian (M Koit) — University of Tartu
- Approx 60 senior staff (incl postdocs), 60 junior staff (mostly PhD students)
- SF funding over 7 years 4.25 MEUR + an additional Estonian state grant of 0.25 MEUR and reimbursement of VAT

Objectives (according to proposal)

- General objective:
  to consolidate and advance the Estonian computer science in six areas of recognized strength.
- Specific objectives:
  to boost the research potential of the groups involved by facilitating collaboration and safeguarding their sustainability and growth,
  to increase the impact of their research results in academia, industry and society as well as to popularize them.
- To be achieved by:
  carefully planned coordination and joint actions, to create a thriving, highly reputed research environment, attractive for young researchers, in particular from abroad.

Activities

- The activities fall into two groups:
  - coordinated research:
    regular research activity, aiming at the highest quality and emphasizing, in particular, collaboration across institutions and topic areas
  - coordinated support actions:
    specific actions targeted at developing the research potential of the groups involved and increasing the impact

Working groups (WGs)

- Research is centered around 6 thematic working groups (WGs) (for the 6 areas of strength):
  - programming languages and systems (PLS),
  - information security and cryptology (Sec),
  - software engineering (SE),
  - scientific and engineering computing (Comp),
  - bioinformatics (BI),
  - human language technology (LT)
- These cross the TFT and institution boundaries wherever appropriate.

Research activities

- The research activities comprise:
  - research within the WGs,
  - cross-WG research efforts,
  - dissemination,
    via high-level scientific publications, tutorials, intensive courses, seminar talks at foreign universities, the centre's web portal
- The MC monitors the quality of the centre's publications and other dissemination.

Support actions

- The support actions go into strengthening the centre and increasing its impact.
- Strengthening the centre: making it a sustainable thriving research environment capable of attracting and keeping talent.
  - Human resources:
    - positions for postdocs PhD students, technical personnel,
    - training.

    Personnel development is conducted in adherence to the European charter for researchers and code of conduct for the recruitment of researchers.
  - Equipment.

  Most of the centre's funds are put into personnel and equipment.

- Increasing the impact: enhancing the centre's visibility, i.e., raising the awareness of the target groups of the centre's research results.
  - International cooperation:
    - organization of high-level scientific events,
    - international cooperation projects.
  - Technology transfer:
    - contact days for industry,
    - industrial cooperation projects.
  - Contribution to policy-making:
    - contributions to shaping of policies in R&D, higher education and IT related areas, technology roadmapping and foresighting, standardization etc
  - Popularization:
    - media coverage,
    - popular books,
    - open-door events for the general public

Management

- Management committee (MC):
  operative management, consists of the leaders of the 4 TFTs and the leaders of the WGs
- General assembly (GA):
  institutional strategy, consists of institution administration representatives (one from each) + the leaders of the 4 TFTs
- International advisory board (IAB):
  scientific advice, consists of internationally renowned researchers from abroad
  - Ivan Damgård (U. of Aarhus),
  - Reino Kurki-Suonio (Tampere U. of Techn.),
  - Kim G. Larsen (Aalborg U.),
  - Heikki Mannila (Helsinki Inst. of IT),
  - José Nuno Oliveira (U. do Minho),
  - Martin Volk (U. Zürich),
  - Reinhard Wilhelm (U. des Saarlandes).

Highlights of the first two years 2008–2010

- A postdoc programme launched highly successfully
- Contributions to a number of FP6/7/other European projects
- High-level training events (winter/summer schools, theory days), industry contact events
- Two absolute top international conferences attracted to Tallinn for 2011/2012
- Translation of a popular book (D Harel, *Computers limited*) to appear before summer

Postdocs

- Was a central item in the proposal
- A major international recruitment campaign in autumn 2008
- At the moment, 8 international postdocs work at EXCS:
  - PLS: J. Chapman, S. Capobianco, K. Nakata (IoC)
  - SEC: M. Gonzalez (CybAS)
  - SE: L. Garcia-Bañuelos (UT)
  - COMP: S. Srirama (UT)
  - BI: B. Rajashekar (UT)
- The LT WG is also seeking a postdoc
- The postdocs are major contributors to the CoE project (research activity, supervision of junior researchers etc)

EU R&D projects during 2008–2010

- IST coord action TYPES (proof assistants, dependently typed languages) (2004–2008, IoC, partner)
- IST integ project MOBIUS (proof-carrying code for small devices) (2005–2009, IoC, partner)
- IST integ project AEOLUS (overlay computers) (2005–2009, CybAS, partner)
- IST STREP BalticTime (timestamping and e-government) (2006–2009, CybAS, partner)
- LifeSciHealth STREP COBRED (colon and breast cancer diagnostics) (2007–2010, UT, partner)

EU R&D projects during 2008–2010 (ctd)

- ICT STREP RoboSwarm (intelligent robot swarms) (2006–2009, TUT, coordinator)
- ICT collab action VirtualLife (security in virtual life) (2008–2010, CybAS, partner)
- ICT integrated project HATS (trustworthiness and adaptability of software) (2010–2013, IoC, partner)
- Infrastructures prep phase project CLARIN (common language resources and technology infrastructure) (2008–2010, UT, partner

cost

- action IC0701 (verified object-oriented software) (2008–2012, Estonia, signatory, T Uustalu, MC member)
- action IC0702 (soft computing and statistical methods for data analysis) (2008–2012, Estonia, signatory, J Vilo, MC

Theory days

- A biannual domestic (Tallinn-Tartu) training event mostly targeted at junior researchers
- A forum for Tallinn-Tartu interactions
- 16 editions held thus far, since 2002
- Theory days at Jõulumäe, autumn 2008, in cooperation with U of Latvia
- Theory days at Mäetaguse, autumn 2009, in cooperation with St Petersburg Dept of Steklov Math Inst (PDMI)
- Next edition autumn 2010 in Latvia

EWSCS winter schools (similar: ESSCaSS summer schools)

- A series of international schools in computer science, with a theory bias, again targeted at junior researchers
- A top event with 15-year tradition, has outlived all projects and programmes that have supported it
- 5-day schools with approx 50 participants (25 from Estonia, 25 from abroad)
- A school's programme usually consists of 5 courses of 6 hrs
- The lecturer list of the schools includes many of the world's most renowned computer scientists — some of them the founding fathers of their fields
- EWSCS '09: N Courtois (UCL), P Dybjer (Göteborg U), R Gennaro (IBM TJ Watson RC), P W Goldberg (U Liverpool), M Müller-Olm (U Münster)
- EWSCS '10: R Cockett (U Calgary), J Groth (UCL), A Kiayias (U Athens), C Morgan (UNSW, Sydney), A Mycroft (U Cambridge)

Symposia on Innovative Software Technology, IST

- Local industry contact events 2008, 2009, with international academic and industrial speakers alongside local presenters
- Organized by University of Tartu (M Dumas)

EUROCRYPT 2011

- EUROCRYPT is Europe's main cryptology conference, the flagship conference of IACR in Europe (along with CRYPTO and ASIACRYPT globally)
- It is normally a 5-day event
- EUROCRYPT 2011 in Tallinn will be the 30th edition
- Organized by Cybernetica AS (H Lipmaa)

ETAPS 2012

- ETAPS is Europe's main software technology event, run by three associations, EAPLS, EATCS and EASST
- A confederation of 5 highly competitive conferences, running partly in sequence, partly in parallel during 5 days ...
- ... together with around 20 satellite workshops that take place during 2+2 additional days before and after the main conferences
- ETAPS in Tallinn will be the 15th edition
- Organized by Institute of Cybernetics (T Uustalu)
- Projected number of participants: 600

EXCS values

- High-quality research has priority over any other activity. Nonsense cannot be afforded
- People matter most
- Quality of research is defined by recognition by true experts (the international research community) rather than spreadsheet software
- Indicators to assess research must be meaningful and fair

Value of ISI WoS indexing or citation counts as indicators

- Why is publicly funded research measured by a commercial database with closed access?
- Coverage extremely uneven over different disciplines
- Citation matching with cited publications very poor
- More often than not 2nd authors do not get their citations counted
- In math and comput sci, traditionally, alphabetical order of authors is used (a different order is frowned upon!)
- In some cases, correlation between ISI publication/citation counts between real numbers of quality publications/citations is practically nil
- Several leading comput sci journal editorial boards, professional fora (Informatics Europe, ACM) worldwide have protested against the ever-increasing reliance on ISI WoS alone

More on ISI WoS

- Nonetheless, we monitor our ISI WoS counts, but also counts in Scopus, Compendex etc
- EXCS webpage displays detailed record identifiers of all ISI WoS indexed EXCS publications

EXCS publication quality policy

- Indexing $\neq$ quality
- Funders' reliance on counts rather than community recognition has led to a proliferation of fake publication venues (J'accuse!)
- EXCS MC recommends choosing Elsevier, Springer, ACM, IEEE and other professional society publications (in particular electronic and/or open access publications), irrespective of their indexing status
- EXCS MC discourages publishing with Bentham Open, GESTS, IASTED, NAUN, SRP, WASET, WSEAS and similar spam publishers, irrespective of their indexing status
- EXCS does not pay for conference/publication costs with the above publishers.

What we would see as meaningful indicators of success?

- Advance of the field and take-up by economy/society
  - publications, patents
  - citations, use
- Recognition of the scientific community
  - grant income
  - prestigious duties (PC chair, editor, opponent/expert, leadership)

Questions from Archimedes

Has the CoE project been successful so far (impact etc)?

- Undoubtedly yes
- I personally put the highest value on the postdoc programme and human resource development in general ...
- and also on internationalization of research and international visibility
- ... in particular the two absolute top conferences in 2011/2012

Has CoE status eased access to other funding?

- Not really, but with an exception:
- The TransFICS proposal for an EIT KIC (knowledge and innovation community) in ICT was built on STACC/ELIKO/EXCS/CEBE in the Estonian colocation node
- Only 1 KIC was awarded (EIT ICT Labs), ours, led by Scotland/U Edinburgh was ranked 2nd!

Do cross-institutional WGs work?

- Yes, but...
- ... mostly where good collaboration was in place before
- Depends really on the individual WG leader

Does the management structure work?

- Yes, on paper perfectly!
- (Minutes of meetings, documentation in one place etc...)

Does admin of CoE differ from admin of other research projects?

- Is it similar to it in any way?

Interplay with other SF instruments of funding?

- (Comp centers, infrastructure projects, MOBILITAS, DORA)
- More money, but ...
- .... the multiplicity of these instruments complicates admin
  - need to decide which activity and expenses belong in an often unclear situation
  - in particular, need to guess the rules that have not yet been invented/communicated, but will be enforced retroactively
- Some measures are outright contradictory (MOBILITAS vs DORA n vs DORA doctoral schools)
  - their objectives duplicate each other, but your activities within them must not