

IAP MidTerm Review: Comments by the PC

1. Introduction

The midterm review as a whole was very positive.

The IAP pointed out some areas for improvement. The PC identified the ones listed in section 2.

The PC added its comments and envisaged actions/policy for the remainder of the STEVIN programme in section 3.

2. Areas of improvement in the IAP MidTerm Review as identified by the PC

- IPR
 - “IPR issues [...] deserve continuous attention”:
 - (summary, par. 3)
 - See also (Q1 par 5, p. 3)
 - IAP “formulates some specific advices on dealing with IPR”
 - (summary par 3)
 - Esp. Q3 par 4 p. 5
- Areas not (sufficiently) covered
 - “Multimodal / multimedia projects” are not covered (summary par 4)
 - Semantics
 - “Some Semantic aspects [...] are scarcely represented” (summary par 4)
 - Semantic discourse annotation (p. 4 Q2 par 5)
 - Semantic web aspects (p. 4 Q2 par 5)
- Visibility
 - Top level publications
 - “Publications in top level conferences and journals should be feasible” (summary, par 5)
 - “ambitious publication target [should be] a central aim for STEVIN in the next few years” (summary par 5)
 - “Publications in highly-rated peer-reviewed journals” should be encouraged (summary par 6)
 - Q3 par 2 p. 5
 - Q6 par 1 p. 7
 - “Participation in (international) evaluation exercises is encouraged”
 - (summary par 6)
 - See also Q3 par 1 p. 5
 - English language website needed
 - STEVIN should also be visible at the programme level e.g. at LangTech and in publications to be submitted to Springer “
- Promoting HLT among students and in educational academic programmes
 - Importance of stimulating (PhD) students interests in HLT research and promoting HLT also in the educational academic programmes (Q1 par 6 p.3)
- Cooperation of LT and ST
 - “Cooperation between language and speech technology needs to be encouraged further” (Q4 par 1 p.5)
- No collaboration of larger industrial enterprises
 - Probably because of the difficult situation around IPR (Q4 par 2 p 5)

3. Response PC to these items

- IPR
 - The PC agrees that continuous attention is required. The IPR committee, though currently sleeping, will be reactivated when needed. In the meantime, with the variety of projects covered, and especially projects such as D-COI, DPC and SONAR, many IPR aspects are getting better settled and increasingly more standardized.
 - The PC will study the recommendations made by the IAP and formulate a recommendation in the matter to the Board.
- Areas insufficiently covered.
 - The PC agrees that multimodality / multimedia are not represented in STEVIN. In part this can be explained because of the existence of the partially concurrently running IMIX programme, which focused exactly on these areas. However, the IMIX review committee observed that multimodality / multimedia could not be seriously addressed in IMIX given the very limited budget. So, since it is not possible to address this matter in the STEVIN programme itself anymore, we agree that this should remain an item for the research agenda for the near future, and it will be included in the plans for developing STEVIN and IMIX follow-up programmes.
 - As to semantic discourse and semantic web: The PC acknowledges that these aspects are underrepresented. The interest, knowledge and expertise in the area of discourse are certainly available since several projects in this area have been submitted (but not honoured). These matters should indeed also be part of the research agenda of the near future and be taken into account in the plans for developing STEVIN follow-up programmes. It is also possible that other, related programmes currently in preparation may offer opportunities for these aspects, e.g. CLARIN-NL and especially IIP/Create.
- Visibility
 - The PC appreciated the IAP's opinion that STEVIN as a programme should be widely known to the outside world.
 - A subset of the PC and possibly the board will write an article on the STEVIN programme intended for publication e.g. with Springer, as suggested. The LRE journal might be an appropriate journal for an article on the STEVIN programme.
 - The PC will make sure that STEVIN as a programme will be visible at Langtech and other relevant conferences in the form of presentations and other ways (e.g. demonstrations). We also consider organizing special sessions at Interspeech-2009 and LREC 2010.
 - The PC will stimulate all project participants to submit papers to high-level conferences and journals. It already started doing this at the 2008 STEVIN programme meeting (sep. 2008). However, one must realize that realistic chances for acceptance are primarily available for research projects (including application-oriented research projects), except when journals specifically dedicated to language resources (such as LRE) are targeted.
 - Participation in external evaluation exercises already happens in relation to some of the projects. The PC will further stimulate participation in external evaluation exercises, but it realizes that the project plans and budgets have already been fixed and allow only for limited modification.
 - English website has been created.
- Promoting HLT among students and in educational academic programmes
 - Promoting HLT among high school and young academic students is already being carried out in the Educational calls, two of which have been honoured so far.
 - Promoting HLT in educational academic programmes is something to be addressed, but we must be realistic since the STEVIN programme can have only a very limited influence here. We will inventory what is available, identify successful approaches and propose areas that should be strengthened.

- Co-operation of language technology and speech technology
 - There certainly is already some limited cooperation, especially in projects such as D-COI and SONAR that will yield results that will have to be useful both for language technology and for speech technology, as well as in NEON that focuses on semi-automatic subtitling.
 - We may expect a further stimulus in cooperation when SPRAAK becomes fully available. In fact, we already see such an effect in the projects DISCO that focuses on the integration of speech technology in a language learning environment and AAP that integrates speech technology in an alphabetisation course for foreign language students
 - The PC would like to understand better from the IAP what the purpose(s) would be for more intense cooperation of language and speech technology. An answer to that question will make it easier to determine for the PC whether, and if so, which additional actions are required.
- No collaboration of larger industrial enterprises
 - The PC denies the validity of the statement, and its suggested cause.
 - First, there are no “large industrial enterprises” with a major focus on LST, and a focus on the Dutch language.
 - The only possible exception is Nuance (major focus on LST, in Flanders, but of course, its focus is not on the Dutch language). This company, however, participates in several projects and until recently provided (in part) the PC chair.
 - Large enterprises as application /service providers
 - The Demonstrator projects were focused on MKB/KMO (“small businesses”) .
 - This focus explains the absence of e.g. Logica CMG in the demonstration projects
 - Large enterprises as customers
 - NPO (Netherlands Public Broadcasting)
 - VRT (Flemish Public Broadcasting)
 If these can be considered “large” .
 - Other large enterprises
 - Big financial companies:
 - Rabobank provides on PC member (Ruud Smeulders).
 - Q-Go, itself a small business, provides a PC member (Leonoor van der Beek), participates in projects and has many big banks as customers.
 - Big legal companies. These are lacking but they all work internationally and do not have their focus on the Netherlands or Flanders
 - Big medical companies / hospitals are indeed largely lacking, but they can probably only be reached via companies that provide language and speech technologies to them, and these are themselves small businesses again, such as e.g. G2Speech (which is, admittedly, absent in STEVIN)
 - Other potentially interested “large industrial enterprises” are involved at least as data providers and sometimes also active in projects:
 - PCM
 - TeleAtlas
 - Van Dale