

To the memory of Adam Kilgarriff



1960-2015

Adam Kilgarriff died on Saturday May 16 (cancer of stomach, diagnosed in November 2014)

His work

Adam Kilgarriff started his computational linguistics research at University of Sussex. During his PhD he worked on two research topics. The first one –word senses– was the topic of his thesis. Later he published the main arguments in well known journal article I don't believe in word senses (AK 1997). The second topic was corpora. He always thought that linguistics could be an objective science, it should be based on samples of data – corpora.

In the mid 1990s he investigated various statistics that were used with corpora. It became clear to him that a scheme of measuring distances between text types, as represented by corpora, was needed. Without one, the foundations of corpus methods for linguistics would remain shaky. He made a proposal in journal-article Comparing Corpora (AK 2001).

He tried to automate discovering word senses in corpora with mixed results. It turned out he could do the first part of the task – identifying grammatical patterns and collocations for each word automatically from corpora – very well. He named the summaries word sketches and they were first used to support the lexicography of a brand new dictionary, the Macmillan English Dictionary for Advanced Learners (edited by Michael Rundell, 2002). The process was presented at the EURALEX lexicography conference, in Copenhagen in 2002. The work was enthusiastically received, and afterwards a number of people came up to him and asked "can I have them for my language, please". That was the starting point of forming his company Lexical Computing Ltd. and producing the main product of the company: Sketch Engine.

Relations to Brno

Karel's contact with Adam started in late 1990s, then in 2002 the workshop Lexicom took place in Brighton where Pavel met Adam first time. Pavel presented there his corpus query system, later he extended and developed his system to incorporate word sketches, and launched the new system, Sketch Engine, as language and corpus independent tool at the next EURALEX conference, in Lorient, France, in 2004.

Since that time, most of Adam's research was around the Sketch Engine and most of the research results found its way to new Sketch Engine features or corpora. Pavel and Adam have been working together with a growing team who have largely been Karel's PhD students at FI MU Brno.

Thanks to the positive relations between Adam and Pavel the cooperation with NLP Centre at FI MU started to grow successfully. The cooperation with Brno in time become closer and more intensive in several research areas, particularly, in building large web corpora for a number of languages.

The affiliation of the Czech Branch of Lexical Computing Ltd in Brno was established in 2011.

It has to be remarked that in cooperation with Adam three Lexicom workshops were organized in 2005, 2007, 2009 in Brno.

Adam's importance

AK was one of the outstanding lexicographers and linguists in the world and he was also dealing with dictionaries and their structure. He was also deeply interested in techniques of determining meanings of the headwords with respect to their disambiguation and the evaluation of the disambiguating procedures/algorithms. Further, he was systematically extending his interest in text corpora for a number of languages (recently about 80), obtaining lexicographical data from corpora and exploitation of the Web for building large web corpora and interconnection of dictionaries and corpora. In general, he paid an intensive attention to the lexical semantics and to formalisms for representation of the lexical meaning.

His company Lexical Computing (established in 2003) in close cooperation with the NLP Centre FI MU in Brno (P. Rychlý, M. Jakubíček, K. Pala) has in the course of time become one of the leading world groups in the area of computer lexicography. He, as a person cannot be replaced but, fortunately, we can learn from the large amount of his publications and videos of his lectures accessible on the Web.

The Future of the LCL

Adam was a researcher, his LCL was a research company. After Adam's death, the LCL goes on, it still is and will be a research company. The LCL internal documents include huge amount of Adam's thoughts in form of specifications, comments and proposals. LCL developers continue in Adam's research and we can see several new papers with Adam as a co-author in near future.

Karel Pala and Pavel Rychlý
NLP Centre Faculty of Informatics MU Brno
{pala, pary}@fi.muni.cz