TOWARDS THE LAST ANNOTATION TOOL

K. B. COHEN, D. DEMNER-FUSHMAN, K. FORT, C. GROUIN, L. E. HUNTER, U. LESER, A. NÉVÉOL, M. NEVES, P. ZWEIGENBAUM

1. Problem to be solved

We propose to take advantage of the diverse mix of areas of expertise that will be represented at BLAHmuc to design analyses, surveys, and usability studies of biomedical text annotation tools, with the goal of determining a set of functionalities and requirements specifications such that if an annotation tool, or a framework for assembling annotation tools, had those functionalities, it would not be necessary to continually develop new ones. A significant amount of the funding available for research in natural language processing is spent on developing annotation tools. Just in the past five years in the biomedical text mining community alone, we are aware of 5 projects that have developed or are developing their own tools for annotating biomedical literature, at an estimated cost of $500,000 US. Neves and Leser [1] carried out a survey of annotation tools, particularly ones that had been used in the biomedical domain. One of the findings of the work was that multiple good annotation tools are available. Why, then, does our community keep building new ones? We hypothesize that we do so because none of the existing tools has a sufficiently broad set of capabilities; that it is possible to determine with some degree of accuracy and completeness what would constitute a sufficiently broad set of capabilities; and to design a set of tool or architecture requirements that would make it possible to deliver those capabilities.

We will focus this hackathon project around a small set of questions and deliverables. Is it possible to design a single annotation tool that could fit the needs of all or most textual annotation projects? If it is not: can the community design an architecture that would allow for plug-and-play annotation tool components with a standardized data format, and what functionalities should it provide? How could such a tool or architecture be evaluated? The biomedical field presents an especially interesting challenge here, since it has particular concerns that are not necessarily found in annotation tasks in other domains, such as the necessity of accommodating large ontologies as annotation schemata, and long texts (e.g. full journal articles) as the data to be annotated. Solving the problem of designing a single annotation tool or architecture for quickly and easily assembling an annotation tool from provided components could result in a significant decrease in the proportion of research funds that need to be spent on annotation tool development, and a concomitant increase in the funding available for other biomedical text mining research areas.

2. Approach proposed

BLAHmuc participants will include annotation tool developers, text mining methods developers, and builders of public biomedical databases. We propose to take advantage of this diverse mix of areas of expertise to collect data on the requirements for annotation tool design for biomedical literature annotation. We will make use of the following methods to collect data:

- Review and analysis of extant tools to develop a catalogue of functionalities provided by current tools
- Established requirements collection methodologies, as described in Wiegers and Beatty 2013, *Software requirements, 3rd edition*
- Design a user satisfaction survey to distribute to ongoing annotation projects to collect data on their experiences with their current tools
- Set up usability studies for a small set of annotation tools

3. Deliverables

The products of the work done over the course of the one-week hackathon would be:

- A draft of a review paper on requirements for annotation tools for biomedical text, based on data collected at the hackathon
- A user satisfaction survey that can be distributed and analyzed during the period after the hackathon
- A protocol for carrying out usability studies that can be executed by hackathon participants after returning to their home institutions

4. Background

Neves and Leser [1] carried out a survey of annotation tools, particularly ones that had been used in the biomedical domain. Out of around 30 tools available, they examined 13 in detail. These were analyzed with respect to 35 criteria, falling into the general areas of functionalities provided, input and output formats supported, and portability. Some of the conclusions of the study were that none of the tools satisfies all needs, but that many good solutions are already available. Fort [2] examined a number of annotation tools and approaches to manual annotation, concluding that the availability of a common platform for annotation would open up myriad opportunities for obtaining high-quality annotations at lower costs than at present.

The work of Neves and Leser demonstrates that achieving insights into a large number of annotation tools is possible. Fort's work supports the surmise that pushing Neves and Leser's work in the direction of requirements definition and of usability studies can help us design an annotation tool (or platform for assembling tool components) that would serve the needs of the biomedical text mining community for some time, potentially leading to increased financial efficiency of annotation tasks, thus freeing up resources for other areas of research in the field.

We understand the temptation to build a newer and better annotation tool—among us, the proposers of this project have funded the development of a number of new annotation tools. A hackathon project with the goal of defining the right set of functionalities and defining a solid set of requirements would be a first step—and a big one—towards making that kind of work necessary only one last time.

References

[1] Neves, Mariana, and Ulf Leser (2012) A survey on annotation tools for the biomedical literature. Briefings in Bioinformatics, Oxford University Press.
[2] Fort, Karën (2016) Collaborative annotation for reliable natural language processing: Technical and sociological aspects. Wiley-ISTE.