4.2 Tools

Nachum Dershowitz, Matthieu Exbrayat, Eyal Ofek, Micalis Panagopoulos, and Ségolène Tarte

What tools are needed to progress in this field? The assets of computers are their ability to deal with big data, using memory, distinction/identification of fine differences, and rare occurrences. The assets of humans are dealing with complex data, making sense of the data, and gestalt questions. It is vital that these two sets of assets are combined through semi-automatic and interactive tools, not through 'black boxes': we must always keep humans in the loop! This includes

- Provide training data / annotated data
- Online training / expert-in-the-loop
- Crowd-sourcing

Rather than a single product, we also need a collection of tools that contribute to each other: a toolbox to account for the different needs of different researchers.

Low-level tools:

- Binarization – segmentation – alignment / matching / registration (for later comparison) – physical feature extraction – expert feature extraction (angles, curvatures, strokes. . .) – similarity measures (for comparison between characters, words, texts, fragments, documents, corpora)

Medium-level tools:

- Clustering – classification – character recognition – word spotting – searching (text via string – text via image – image via text – image via image – characters) – image-text correspondence
- Databases: organisation of data in a way that allows fast queries of metadata, transcripts, text qualities, etc.

Higher-level tools:

- Interfaces, ergonomy (CHI) – searches of combinations of characters/words (bigrams, trigrams) – correspondence of expert vocabularies – inferences of paraphrases and synonyms for searches through metadata
- A transcription tool to make the connection between text as shape and text as meaning

Other principles for development:

- Feedback loops and cognitive triggers: drawing/touch screen technologies – simple interactive image enhancements – visualization aspects of interactions with these tools (of results, of databases) – interactive visualisations (e.g. time varying graphs) – customizable visualisations – multiple languages – rationale building support, tracking of expert hypotheses in interpretation building – statistical tools with tests of significance – information sharing – sounding the texts
- Web-services to provide access to such tools via internet?

This topic has potential links with medical imaging, cognitive sciences, CHI, and NLP, all of which should be explored in future work.