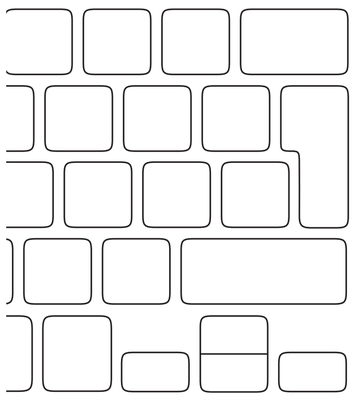
Department of Computer and Systems Sciences

Stockholm University

Using text prediction for facilitating input and improving readability of clinical text

Magnus Ahltorp, Maria Skeppstedt, Hercules Dalianis, Maria Kvist

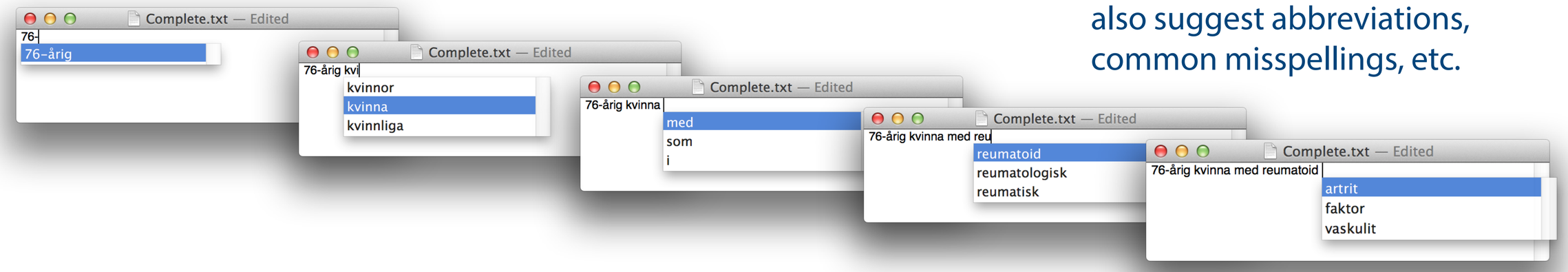


Text prediction

- has the potential for facilitating and speeding up the documentation work in health care
- offers a way to produce clinical text with fewer misspellings and abbreviations, increasing readability

However, there are obstacles:

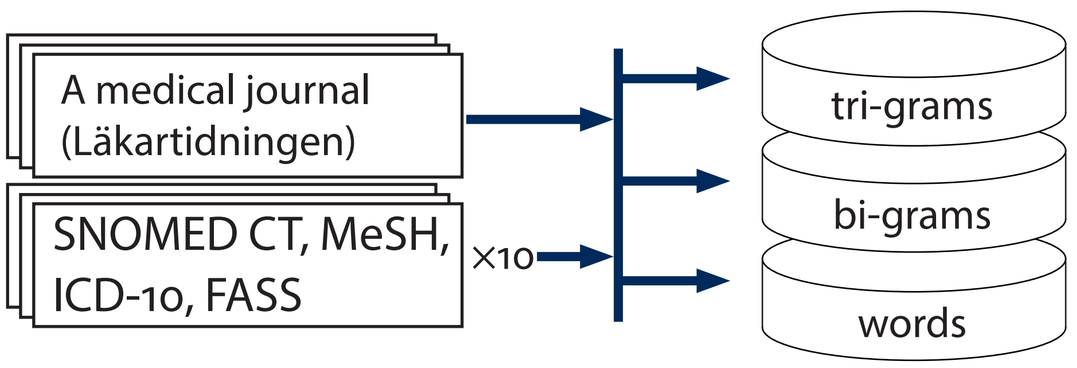
- authentic clinical text is not always available for application development and often contains sensitive information
- a text prediction tool trained on unmodified clinical text would also suggest abbreviations, common misspellings, etc.



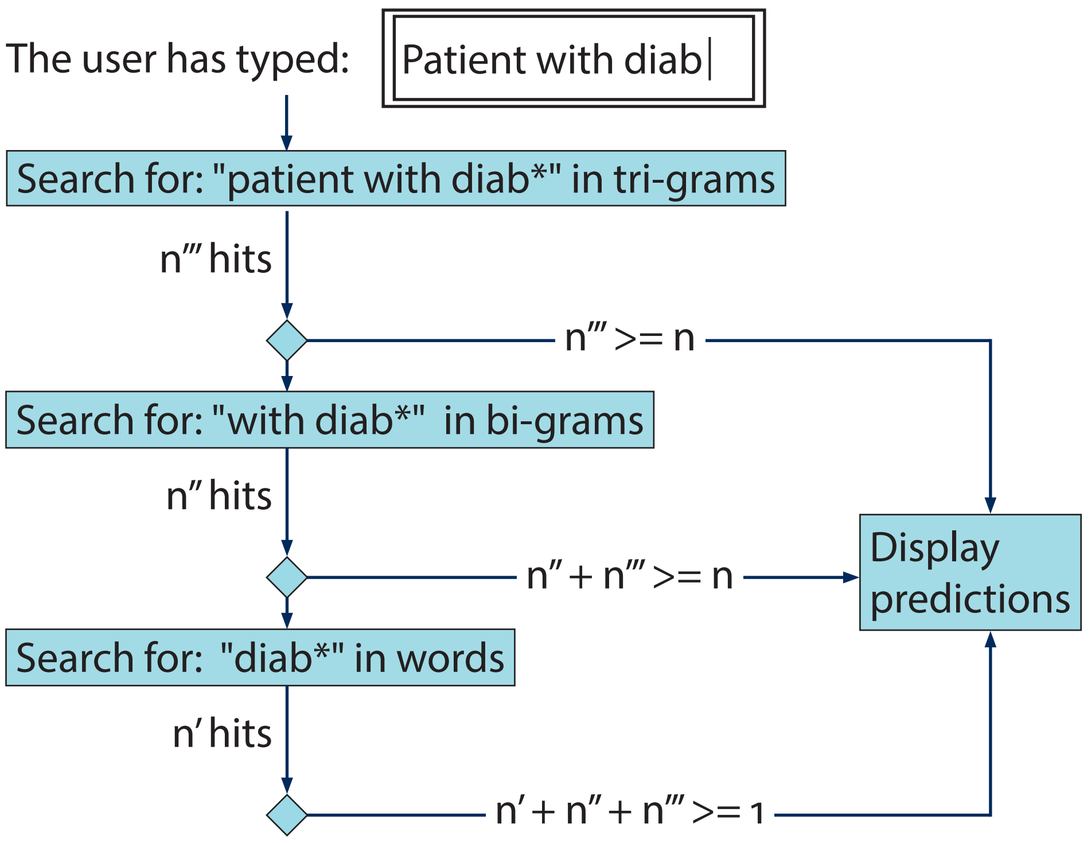
Method

Text prediction on a word level predicts what the user intends to type, given the previous context. Statistics from existing texts are used for the prediction, usually from the same domain. In this study, we evaluated the possibility of predicting clinical text using other medical text resources.

Building a database of word and n-gram frequencies



Applying frequencies for retrieving n predictions



Results

Saved keystrokes were evaluated for n = 1.

Evaluation data	Number of letters	Saved key-strokes
Upper ceiling: Läkartidningen (600 sentences)	64 560	32%
Frequent domain words, ICU notes (100 words)	675	20%
Frequent bigrams, ICU notes (100 bigrams)	870	35%
Frequent trigrams, ICU notes (100 trigrams)	1 453	33%
Medical exam questions mimicking clinical text (240 sentences)	13 535	26%

Future work

- Evaluate on a large corpus of authentic clinical text.
- Use more advanced text prediction techniques.
- Incorporate syntactic and semantic categories or frequent words or n-grams from authentic clinical text.
- Adapt/develop evaluation techniques to a fast typing user.

Correspondence to: mariask@dsv.su.se