

Evaluating GDEX

Miloš Jakubíček



`milos.jakubicek@sketchengine.co.uk`

February 12th, 2015
ENEL meeting Vienna

GDEX as a research problem

GDEX as a research problem . . .

- given a headword and a set of concordances
- automatically produce ordering of concordances according to their suitability for inclusion as good examples for the headword

GDEX as a research problem

... is impossible to evaluate

- hard / not feasible to evaluate recall
- many in-built biases in data (text types, ordering of corpus)
- presumably very low inter-annotator agreement

GDEX as a research problem

Many NLP problems are

- practically motivated
- simply put
- impossible to evaluate
- but in need of solving

GDEX as a research problem

best / worst vs. better / worse

resource

19a) DR37700B: Resource disconnected (buffer not flushed)

vs.

Free resources are available to aid research.

GDEX as an industrial problem

GDEX as an industrial problem . . .

- given a headword
- make producing dictionary-quality examples faster

GDEX as an industrial problem

Evaluation?

GDEX as an industrial problem

Evaluation?

- some sort of A/B testing

GDEX as an industrial problem

Evaluation?

- some sort of A/B testing
- \Rightarrow which setting works better/faster?

GDEX as an industrial problem

Evaluation?

- some sort of A/B testing
- \Rightarrow which setting works better/faster?
- error driven development

GDEX as an industrial problem

Evaluation?

- some sort of A/B testing
- \Rightarrow which setting works better/faster?
- error driven development
- \Rightarrow what are the most frequent examples that are ruled out?

Conclusions

We want to support people in their own developments on GDEX.

Conclusions

We want to support people in their own developments on GDEX.

Further *technological* development is from the business point of view hard to justify unless

Conclusions

We want to support people in their own developments on GDEX.

Further *technological* development is from the business point of view hard to justify unless

§ there is evidence it helps (which is hard to provide)

Conclusions

We want to support people in their own developments on GDEX.

Further *technological* development is from the business point of view hard to justify unless

- § there is evidence it helps (which is hard to provide)

- \$ people pay (enough:)

Conclusions

We want to support people in their own developments on GDEX.

Further *technological* development is from the business point of view hard to justify unless

- § there is evidence it helps (which is hard to provide)

- \$ people pay (enough:)

Until then: problem solved (technically).

Conclusions

We want to support people in their own developments on GDEX.

Further *technological* development is from the business point of view hard to justify unless

- § there is evidence it helps (which is hard to provide)

- \$ people pay (enough:)

Until then: problem solved (technically).

⇒ **The floor is yours!**