

DuOMAn

Dutch Online Media Analysis

STEVIN project: April 2008 – March 2011

Valentin Jijkoun
ISLA, Universiteit van Amsterdam

June 27, 2008

DuOMAn: goals

- Development of a set of Dutch language resources and tools for identifying and aggregating sentiments in online data sources
- Sentiment mining:
 - *Who supports/criticizes whom and on what topics?*

DuOMAn: goals

- Data:
 - edited (news)
 - user-generated content (news comments, blogs)
- End users:
 - Media-analysts (toolbox)
 - General public (demos)

DuOMAn: research team

- University of Amsterdam coordinator
- University College Ghent coreference resolution
- Rijksuniversiteit Groningen parsing
- GridLine lexical resources
- TrendLight end user, annotations

DuOMAn: key deliverables

- Data:
 - Dutch corpus of blogs and news comments
 - Sentiment lexicons and test sets
 - Syntax- and anaphora-annotated corpora of blogs/comments

DuOMAn: key deliverables

- Tools:
 - Syntactic parser for blog texts (ALPINO)
 - Entity coreference for blog texts (COREA)
 - Sentiment extraction for news and blogs
- Demos:
 - yearly; first demo: September 2008
- More: <http://ilps.science.uva.nl/DuOMAn>