Gabmap – A Web Application for Dialectology

John Nerbonne, Rinke Colen, Charlotte Gooskens, Peter Kleiweg and Therese Leinonen

CLARIN-NL Workshop
Utrecht, Aug. 25, 2011
Gabmap

› What is Gabmap? (demo)
› Why Gabmap?
› Point of Departure
› Goals
› Team
› Problems
› Agit-Prop
› Reflections
What is Gabmap?

> Map-making facilities (based on Google Earth) available
  - Direct import into Gabmap
  - Not shown here (time)

> http://www.gabmap.nl
Why develop Gabmap?

› Stimulate investigation of exact techniques in dialectology
  - Influence especially younger dialectologists

› Encourage use, but also deeper understanding of these techniques (practitioners still a minority in dialectology)
  - Empower criticism, better-informed debate
Point of Departure & Goals

› Point of Departure
  • 14-yr research line, dozens of papers
  • A suite of UNIX facilities (L04), effective but inaccessible for most dialectologists

› Goals
  • Easy-to-use facility, suitable for classroom
  • Intelligent options for novices (difficult)
  • Feedback wrt to interpretation of results (often misunderstood)
A Team Effort

Groningen (Scientific)
- Charlotte Gooskens (Manager)
- Therese Leinonen (Tutorials)
- Rinke Colen (Film)
- John Nerbonne (Ideas)

Meertens (Infrastructure)
- Jan Pieter Kunst

Peter Kleiweg (Development)
- Very productive, involved

CLARIN-NL (Feedback)
Problems

› Disagreements:
  • Scientific development ≠ Development for unpracticed users

› Standards often incomplete
  • Multiple entries in tables (= multiple responses)
  • Most phonetic, phonological and dialectological notions
Agit-Prop!

- Recruit users, do not merely develop software
  - Features of general interest (data mining)
  - Website, film (screencast), online tutorial, manual
  - Journal paper
  - Tutorials
    - UvA (Heeringa), *Tagung des Forums Sprachvariation* (Erlangen), Universität Potsdam, Tvarminne (Finland), Helsinki, Uppsala, ICLAVE (Freiburg), Digital Humanities 2011 (Stanford), Methods in Dialectology (Ontario), LOT 2012 Winter School (Tilburg)
  - Collaborations (co-publications) underway on Catalan, Tuscan pronunciation and Estonian lexico-syntax
Reflections

› Technically successful now even if incomplete
› It will not be scientifically successful until it attracts more users (I want to see 20 papers published on the basis of Gabmap results.)
  • A textbook on dialectology using Gabmap would be great (but time consuming)

› The source code of Gabmap is open to all, but
  • Open source development practices were not followed, making it difficult for others to modify Gabmap.