CITIZEN MINDSCAPES

Digital Humanities Kick Off 22.3.2016, Turku
80% OF FINNS USE MONTHLY
OVER 70 MILLION
COMMENTS
FROM 15 YEARS

Owned by Aller Oy
Social Statistics, Uni Helsinki:
Statistical modeling of discussion dynamics, Tools

Pol Econ & Consumer Society, Uni Helsinki
Social shaping of Suomi24, Rhythms, Waves of emotion, Connect Lab

DigiCult, Uni Turku:
Research ethics, digital culture of social media

Welfare research, Uni Eastern Finland:
Users, their demographics & wellbeing

NatLang, Uni Turku:
Language technology & machine learning methods
Citizen Mindscapes
Academy consortium project
2016-2018
UH, UTU, UEF
Pol Econ & Consumer Research
Uni Helsinki

• Prof Jussi Pakkasvirta, area and cultural studies
• University researcher Krista Lagus, data analytics & cognition
• University researcher Minna Ruckenstein, anthropologist
• PhD student Marjoriikka Ylisiurua, social sciences
• Research Director, Mika Pantzar, economist
BRIDGING RESEARCH DIVIDES

Engaging the research collective:

- Contextualization of Data
- Practices of Collaboration
- Research Ethics
- Sharing your own data

Data analysis tools

- Visualizations
- Data analysis tools

Social sciences

Language technology

- SUOMI24 DATA
- LANGUAGE BANK FIN-CLARIN & CSC
- Linguistic searches

KORP

CONNECT LAB

Data analysis
CONNECT LAB:
CREATING COLLABORATIVE SPACES

Meetings of the Research Collective
Every 2-3 months

Format: Initial lectures & group discussions, Pecha Kucha

Atmosphere: Ask & Share

Citizen Mindscapes

Research blog:
http://blogs.helsinki.fi/citizenmindscapes/
What can people’s discussions tell us about the society?

- Everyday rhythms: daily, weekly, yearly
- Waves of emotion: “worrymeter”
- Spearhead topics: Money, debt, “sossupummi”, Latin America
- The dynamics of human interaction in discussions
- Some – Media interaction

RESEARCH ON CONTENT
Modelling the intensity of discussions
- For instance so called ‘bursts of posts’ caused by the earlier messages
- Statistical point process theory and survival analysis
Toolbox basics

- Makes the big data **accessible** to non-technically oriented researchers
- Makes the data processing and analysis process **reproducible**
Mindscapes24: UniTurku, DigiCult Team

- Prof Jaakko Suominen
- University lecturer Petri Saarikoski
- University lecturer Riikka Turtiainen
- Researcher Dr. Anna Haverinen
- Researcher Dr. Sari Östman
DigiCult teams main tasks in the project

- Contextualization of Suomi24
  - Book on cultural history of online discussions in Finland
  - Participation to research on Suomi24 user cultures, moderating practices etc.
- Research on hate speech in Suomi24
- Research on questions of research ethics related to discussion forum materials and usage
- Creation of bigger picture, contextualization of Suomi24 from the perspective of the cultural appropriation of the online discussion forums
- How the trolling, flaming, supporting, liking and moderating practices have evolved since the 1980s?
Mindscapes24: UniTurku, Language Technology bionlp.utu.fi

- Turku NLP group: 15 years history in the development of language technology, Finnish one of the focal points
  - Numerous ext. funded projects (AKA, Tekes, Kone Foundation, E. Aaltonen Foundation, industry...)
  - Involved in three of the DigiHum projects
- Dr. Filip Ginter (PI), Dr. Veronika Laippala, Jenna Kanerva, Juhani Luotolahti, Kai Hakala
Suomi24 is 2B+ words long – that is a lot of text
We need language technology
  - Essential analysis: syntactic and lightweight semantic parse of the data with the tools developed in Finnish
  - Use of modern language technology methods to classify and cluster the posts / threads, etc..
  - Help characterize the clusters
Kirjoitettu katollaisuus -palistalle Suomi24
Suomen suurin verkkoyleisö -foorumissa 11.9.2007

edes kuvitella mitään sellaista uskonnollista suuntaa joka oisii voinut vakuuttaa minut samalla tavalla kuin
Mindscapes24: UniTurku, Language Technology

- Our primary role is to apply language technology methods
- **BUT** Suomi24 is very specific data (as you can imagine)
  - We need to develop / improve language technology for such data
  - We need to develop new methods to address the research questions asked during the project
- I.E. → Language technology geeks who help but also do their own methodology research as we go
WP 1.3: Users of Suomi24: Health and well-being, inequality, poverty, and peer support

Researchers: Kari Saari, Juho Saari, Krista Lagus

CITIZEN MINDSCAPES
Users of SUOMI24/ Survey

- Two quantitative surveys during the project:
  - 1. survey (autumn 2016)
  - 2. survey (spring 2018)

- Key themes in questionnaire:
  - Who are they?
  - Why, how and to what they participate in?
  - What kinds of experiences they have on participation?
  - Is SUOMI24 a window to ill-being in Finnish society (e.g. in terms of health and well-being, inequality, poverty and peer support)?
### Approaches on SUOMI24 users

<table>
<thead>
<tr>
<th>Backgrounds and types of users</th>
<th>Sociodemographic data (e.g. age, gender, class, profession, social relations etc.)</th>
<th>&quot;Types of users&quot; (e.g. writers/readers random/regular rookies/veterans registered/non-registered names/pseudonyms etc.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents and motives</td>
<td>Contents (e.g. one-issue people/multiple users/entertainment etc.)</td>
<td>Motives (e.g. information/peer support emotions/social relations/lifestyle/&quot;addiction&quot;/curiosity etc.)</td>
</tr>
<tr>
<td>Special topics (experiences and attitudes on)</td>
<td>Experienced well-being and health Inequality Poverty Peer support</td>
<td></td>
</tr>
</tbody>
</table>