GEOGRAPHIC CONTEXT ANALYSIS OF VOLUNTEERED INFORMATION
(GEOCONAVI)

Frank O. Ostermann
COST Energetic Meeting 26.05.2014, Zürich
GEOGRAPHIC CONTEXT ANALYSIS OF VOLUNTEERED INFORMATION

PRESENTATION OVERVIEW

- What questions were we trying to answer?
- What data have we been using?
- What methods have we been using?
- Where are needs for further research/ collaboration?
RESEARCH OBJECTIVES

Deploy a system for using VI in crisis decision support on forest fires

• Integration
• Quality control
• Communication of risk/uncertainty, alerts

Assess the added value of using VI for forest fire response.
# GEO-SOCIAL MEDIA AND CRISIS MANAGEMENT

<table>
<thead>
<tr>
<th>Social media offers…</th>
<th>Crisis management needs…</th>
</tr>
</thead>
<tbody>
<tr>
<td>rich up-to-date information</td>
<td>up-to-date information</td>
</tr>
<tr>
<td>new paths of communication</td>
<td>redundant paths of communication</td>
</tr>
<tr>
<td>noise, uncertain lineage and accuracy</td>
<td>high-quality and reliable information</td>
</tr>
</tbody>
</table>

Crowd-sourced data curation faces limits of

- Sustainability
- Scalability
FACTORS OF DATA QUALITY

Credibility

Source

Content

Context

Location

Relevance
PROCESSING FRAMEWORK

**Retrieval** (Web 2.0 APIs)

**New VGI**

**Sensor**

**Assessment & Enrichment** (Classification & spatial queries)

**Relevant & Credible VGI**

**GeoCONAVI Analyzer**

**Clustering** (Spatio-temporal statistics)

**VGI-Clusters**

**GeoCONAVI Clusterer**

**Dissemination & Alerting** (WFS, WMS, SES, SMS, …)

**Fully Assessed VGI & VGI-Clusters**

**Disseminator**
GEOCONAVI

1.1 Retrieval
Scheduled Java code accessing APIs

1.2 Storage
Scheduled Java code writing to DBMS

2.1 Topicality
Scheduled PLSQL job

2.2 Geo-Coding
a) Scheduled PLSQL job
b) Scheduled Java code

2.3 Geographic context
Scheduled PLSQL job

2.4 Quality Assessment
Scheduled PLSQL job

3.1 Spatio-temporal clustering
Scheduled Python script calling SatScan job

3.2 Quality Re-Assessment
Scheduled PLSQL job

Twitter Streaming API
Flickr Search API
EFFIS Hotspot Data
European Media Monitor
Geo-coding API
Dissemination
SMS, WFS, WMS, RSS, SES

Oracle DBMS
FOREST FIRES IN FRANCE 2011
FOREST FIRES IN FRANCE BY GEOCONAVI
FRENCH FOREST FIRE SOCIAL MEDIA

(1) Containing French keywords: 659,676 Tweets and 39,016 Flickr images

(2) Machine-learned relevance filter: 25,684 items left

(3) Geocoded and context enriched: 5,770 items left

(4) Clustered in space and time: 129 clusters with 2,682 items

(5) Second relevance filter: 11 clusters left with 469 items
SOME PUBLICATIONS


RESEARCH THEMES “CLOUD & CROWD”

1) **Societal implications** of new production and consumption modes of geographic information.

2) Methods to assess and improve crowd-sourced and volunteered geographic **information quality**.

3) Methods to increase **interoperability** through geographic data integration and linking.

4) Applications to empower citizens, increase participation, and improve **quality of life**.
ISSUES I

New modes of GI-Production

- Crowd-sourced vs. volunteered
- Opportunistic vs. participatory sensing
- Ethical and legal (liability, privacy)

New methods for GI-Processing

- Data quality (reliability, relevance, accuracy)
- Rapid decision-making
- Geographic information streams
ISSUES II

Interoperability of GI
- Syntactic heterogeneity
- Semantic heterogeneity
- Place-based vs. space-based perspective

Applications of crowd and cloud approaches
- Digital divide and uneven information geographies
- Lack of technical infrastructure and expertise
- Entrenched power structures
IMMEDIATE OBJECTIVE

Crowdsourcing …
• Supervised Learning
• Parameter Choice

… for Processing …
• Classification and Labeling
  • Context Analysis
• Spatio-Temporal Clustering

… of Geo-Social Media Streams
• Twitter
• Flickr
• …