PROMAP: developing a research tool for protest mapping

Theoni Stathopoulou
Research Director
EKKE- National Centre for Social Research
The National Centre for Social Research - EKKE in Clarin:el

• Prospective member
• Memorandum of collaboration between EKKE and RC Athena already signed
• Some technical issues still need to be solved (EKKE IdP)
• EKKE is planned to join the Greek network during the next phase (Operational Phase, 2017-2019)
Use of Clarin:el services

1. PROMAP: a use case
2. Radicalism: verbal aggressiveness in twitter data
   • During the PROMAP project ILSP services were not yet integrated in the infrastructure
   • Goal: to replace the manual workflow used until then with automated services, regarding:
     • Collection of newspaper material to be used as corpus
     • Processing of textual material to identify useful excerpts
     • Analysis of excerpts to feed the codebook frames
     • Statistical processing of codified material
     • Visualization of results
Radicalism

Radicalized Orthodoxy?
Civic discontent and religious identification in times of crisis

Theoni Stathopoulou (EKKE) & Haris Papageorgiou (Athena RC)

RESEARCH QUESTIONS

- Verbal Aggressiveness (Twitter)
  - Muslims
  - Pakistani

- Physical Attacks (News Data)
  - Muslims
  - Pakistani

- What are the correlations?
- Does Verbal Aggressiveness lead to Physical Violence?
- Are calls for action turned into events of violence?
PALOMAR ARCHITECTURE
July 15th, 2016, 23:19:15.000
Η προφητεία του Καντάφι που τον διέκυψε: 'Δεν πέθανε, η Ευρώπη θα πέσει στα νύχτια του' https://t.co/PfIaSk9GuS

June 14th, 2016, 14:32:37.000
RT Fanma_news: ΕΚΤΑΚΤΟ: Συριακοί τολμηροί σταθμοί μεταβίβασης όπως ξεκίνησε ο πετός του Ισλαμικού Κράτος Αμφοτέρως και περικοπή έως εξα

June 13th, 2016, 17:06:58.000
RT Fadma_cnn: Δεν άφησε μεταρρυθμίσεις Ισλαμικό Βίκες δεν άφησε μεταρρυθμίσεις Βίκες

June 13th, 2016, 02:14:41.000
To Ισλαμικό Κράτος ανέβασε την ειδήσεις για τα μακελεια στην Καλαμάτα - Αναγνωρίστηκαν οι δραπέτες (photos) https://t.co/MojjD54AqY

June 12th, 2016, 19:08:09.000
RT Mandarinos: Ταξιδιωτική ομήρυξη του Ισλαμικού Κράτους προς τα μέλη του έσπασε τις πόλεις που γίνεται το γιορτάζοντας εκείνον...

June 10th, 2016, 22:18:16.000
RT tribunews: Ινδία από το κέντρο της Ελλάδας ο Ισλαμικός Κράτος https://t.co/aCi5k4w10p https://t.co/BS6QZa37Q

June 9th, 2016, 18:07:32.000
RT ilanros69: Το Ισλαμικό Κράτος ανέβασε την ειδήσεις για την τρομοκρατική επίθεση στον Φαντά, μεσάζοντας και χτυπώντας λάθος ακρόπολη

June 8th, 2016, 16:30:24.000
"Ανθιζομαι και "μπαλάλης" γιατί γεννήθηκε στον ποιντ πόκερ! http://islamico.com https://t.co/XN67pdP5TM
Rationale of PROMAP

• Eurozone and global economic crisis
• New waves of mobilization and protest, new forms of collective action
• New paradigms, repertoires and actors
• Emerging “new” forms of protest and grassroots mobilization on economic hardship have incorporated attributes of “older” forms of protest (mid 1990s)
Aim and Key objectives

*Record and map*

- varying forms of contention and mobilization in post-authoritarian Greece (1996-2014)
- protest events /actors-participants/repertoires/agendas
- the evolution and transformation of contentious claims over time

*Document*

- protest cycles in Greece
- “the dynamics of contention” (Kriesi et. al. 1995; Grande & Kriesi 2013) in Greece
RESEARCH FRAMEWORK

• Collective action
• The role of social movements for social transformation
• Research agenda after 2000:
  • migration and citizenship (Koopmans et al. 2005);
  • the emergence of a European public sphere (Koopmans and Statham 2010; Della Porta and Caiani 2009);
  • the politics of unemployment (Giugni 2009);
  • the spread of new forms of participatory politics and democracy (Della Porta 2007; 2009);
  • the policing of protest (Della Porta, Peterson, and Reiter 2006).
RESEARCH FRAMEWORK

• Protest event analysis (*Diani & Kousis 2014*)
• Claims analysis: Political Discourse Approaches in order to extend the context of political claims making, and become more sensitive to discursive messages (*Koopmans & Statham 1999*) (*Koopmans; Statham; Giugni; Passy 2005*) (*de Wilde; Koopmans; Zürn 2014*)
• Research agenda set by large comparative European projects
• Relevant research projects
<table>
<thead>
<tr>
<th>Title</th>
<th>Research Centre</th>
<th>Time reference</th>
<th>Topic</th>
<th>Method</th>
<th>Resources</th>
<th>Sampling Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRODAT</td>
<td>WZB: Berlin Social Science Center (Rucht)</td>
<td>46 years (1950-1996)</td>
<td>Protest event analysis for the whole time span</td>
<td>event data analysis</td>
<td>PRODAT1: Newspapers (1) PRODAT2: Newspapers (2) and Police Records</td>
<td>Monday Newspaper Edition &amp; all editions of the 4th Week per Month</td>
</tr>
<tr>
<td>Europub</td>
<td>WZB: Berlin Social Science Center (Koopmans)</td>
<td>8 years (1980, 1985, 1990, 1995, 1999-2002)</td>
<td>Degree of European Integration of 7 Countries (3) In 6 Institutional Sectors</td>
<td>political claim analysis/Semi-structured Interviews</td>
<td>Newspaper Archive, Internet</td>
<td></td>
</tr>
<tr>
<td>MERCI</td>
<td>WZB: Berlin Social Science Center (Koopmans)</td>
<td>7 years (1992-1998)</td>
<td>Immigration And ethnic relations in 5 European Cities</td>
<td>political claim analysis</td>
<td>1 Newspaper per Country (4)</td>
<td>Every Monday, Wednesday and Friday</td>
</tr>
<tr>
<td>Title</td>
<td>Research Centre</td>
<td>Time reference</td>
<td>Topic</td>
<td>Method</td>
<td>Resources</td>
<td>Sampling Method</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Responsivegov</td>
<td>University of Leicester (Morales)</td>
<td>30 years (1980-2010)</td>
<td>18 Governments’ (1) responsiveness to different forms of public opinion expression.</td>
<td>Content analysis &amp; protest event analysis/ opinion polls</td>
<td>Newspapers &amp; news agencies archive, existing databases of protest events (2)</td>
<td></td>
</tr>
<tr>
<td>GGCRISI</td>
<td>Freie Universität Berlin &amp; University of Crete (Roose &amp; Kousis)</td>
<td>4 years (2009-2013) [research in progress]</td>
<td>Assigning responsibility for Eurozone Crisis, according to the related public debate in Greece &amp; Germany</td>
<td>Political Claim Analysis, Protest Event &amp; Interpretative Framing Analysis (Conversation Analysis of accountability) (3)</td>
<td>5 Newspapers per Country &amp; Reuters, Internet (Parties and Civil Society Positions)</td>
<td>1 Day per week (rotation)</td>
</tr>
<tr>
<td>POLPART</td>
<td>Stichting Vu-Vumc (Klandermans)</td>
<td>5 years (2014-2018) [research in progress]</td>
<td>Different types of Political Participation (national &amp; international level) (4)</td>
<td>Movements &amp; Parties Publications Meta-analysis, International Research Archives comparison, focus group &amp; Questionnaires (1000 respondents)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Research Centre</td>
<td>Time reference</td>
<td>Topic</td>
<td>Method</td>
<td>Resources</td>
<td>Sampling Method</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>----------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>LOCALMULTIDEM</td>
<td>Universidad de Murcia (Morales)</td>
<td>1 year (2006)</td>
<td>Degree of Integration of Immigrants in 6 European Cities</td>
<td>Political Claim Analysis</td>
<td>1 Newspaper per City (1)</td>
<td>Every Monday, Wednesday and Friday</td>
</tr>
<tr>
<td>ESSi</td>
<td>The Netherlands Institute for Social Research &amp; City (November of 2010)</td>
<td>1 Month</td>
<td>Feasibility study (Use of Media Claims as a Research tool supplementary to the ESS / for 8 Countries) (2)</td>
<td>Media claims analysis</td>
<td>2 Newspapers per Country/ Greece: ‘Ta Nea’ &amp; ‘Kathimerini’</td>
<td>Every day Editions except for Weekends</td>
</tr>
<tr>
<td></td>
<td>University London</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIVEWHAT</td>
<td>University of Geneva (Giugni)</td>
<td>3 years (2013-2016)</td>
<td>Alternative Crisis Adaptability forms of European Citizens, individually, collectively and institutionally (20 countries)</td>
<td>Political Claim Analysis and Action Case Analysis</td>
<td>5 Nationwide Newspapers per Country &amp; internet (facebook pages, blogs, etc)</td>
<td>Representative Sample of 1000 public interventions and alternative adaptability forms</td>
</tr>
</tbody>
</table>
ESSi: Developing event reporting

• Event Data/Event monitoring in European Social Survey
• Supplementary tool to ESS q/naire
• Measuring context in cross-national survey
• Media claims
PROMAP: a collaborative research project

SoDaMap
"Social Data Mapping"
EKKE

Panorama
Partner: Hellenic Statistical Authority

Socioscope
Partner: Athena Research Centre
ILSP-IMIS

IMIS
Institute for the Management of Information Systems
ILSP
Institute for Language and Speech Processing
The need for automated content analysis with NLP tools

• 1. Social Media as a supplementary resource
• 2. “Big” Data
• 3. Identify trends over larger time spans

• University of Zurich - Wuest et.al 2013 (The politicization of Europe- A comparative study of six European countries 1970-2010)
PROMAP Methodology I

- PCA (Political claims analysis) *Koopmans & Statham (1999)*; Koopmans (2002); de Wilde, Koopmans & Zürn (2014)
- Unit of analysis: claim
- Elaboration of new coding scheme
- Refinement of codebook (14ver.)
- Sampling strategy
  - Selection of daily newspapers: Kathimerini (rightist-center rightist), Avgi (leftist)
  - Sample: 2 days per week for the entire period (Wednesday and Friday)
- Pilot Coding
PROMAP Methodology II

Application of a computational method developed by ILSP (suite v.0.4) Iterations (Machine learning) in collaboration with ILSP team. Geospatial mapping and development of cube collection (RDF) by IMIS (socioscope)

• Constraints
  • Availability of selected newspapers (Eleftherotypia)
  • Availability of newspaper data in html or xml format
  • Copyrighted material
  • Selection Bias- Media content
  • Restrictions of automated NLP tools
Definition of a claim in PROMAP

Rephrasing Koopman’s (2002) and Statham’s (2005) definitions:

• “A claim is defined as a purposive unit of strategic or communicative action in the public sphere aiming at a collective stake and having the form of physical or verbal action”.

• A valid claim is not explicitly political- a non political claim may have political implications.
The structure of a claim (variables)

1. Location of the claim in time and space (WHEN and WHERE is the claim made?)
2. Claimant: the actor making the claim (WHO makes the claim?)
3. Form of the claim (HOW is the claim inserted in the public sphere?)
4. The addressee of the claim (AT WHOM is the claim directed?)
5. The substantive issue of the claim (WHAT is the claim about?)
6. Object actor: who would be affected by the claim if it is realized (FOR/AGAINST WHOM?)
7. The justification for the claim (WHY should this action be undertaken?)

* Source: Ruud Koopmans, 2002, *Codebook for the analysis of political mobilisation and communication in European public spheres*, Deliverable D2.1 of the project "The Transformation of Political Mobilisation and Communication in European Public Spheres"
### Examples of claims

<table>
<thead>
<tr>
<th>WHO (SUBJECT ACTOR)</th>
<th>HOW (FORM)</th>
<th>AT WHOM (ADDRESSEE)</th>
<th>WHAT (ISSUE)</th>
<th>FOR/AGAINST WHOM (OBJECT ACTOR)</th>
<th>WHY (FRAME)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The employees of the Greek Public Power Corporation</td>
<td>went on a strike</td>
<td>stating that the government</td>
<td>should not further privatize the Power Corporation because</td>
<td>the Greek people</td>
<td>benefit from the public ownership of the Corporation</td>
</tr>
<tr>
<td>The policemen and the fire-fighters</td>
<td>launched a work-to-rule action</td>
<td>asking the Deputy Minister of Finance</td>
<td>to satisfy their institutional and salary related demands</td>
<td>for themselves (policemen and fire-fighters)</td>
<td></td>
</tr>
<tr>
<td>The municipal authorities and citizens of Melissia</td>
<td>organized a rally</td>
<td></td>
<td>in order to save the surrounding forest</td>
<td></td>
<td>because the natural character of the forest should be maintained</td>
</tr>
</tbody>
</table>
Application of computational method (Computational Social Sciences suite v0.4)

Analysis: Requirements Specification document based on the Codebook

Data: Collection, Preprocessing, Filtering

Design: Domain Modelling, System architecture

Implementation & Testing: tools and resources, sophisticated linguistic rules, machine learning, Basis: ILSP-NLP suite

Integration [CSS suite v0.4]: Claim extraction, Actor/Addressee & Issue detection, Time/Location Recognition

Maintenance
Natural Language Processing pipeline

Form
Actor/Addressee
Issue
Time/Location

Named Entity Recognition
Phrase Extraction
Dependency Parsing
Co-reference Resolution
Aggregation

Claims Detection
EKKE ILSP-NLP
Summarize
Preprocessing procedure

- **Input**: collection of articles from newspapers from 1996 to 2014 in **XML** format
- Each file in the dataset is a unique article.
- **Output**: a collection of text files (TXT)

**Original data from Newspapers**
- Transform XML files to TXT files
- \( N^1 = 330.000 \) articles

**Apply a basic filter (Lexicon)**
- Lemmas (nouns): strike, protest, demonstration etc.
- Locate articles that contain these words

**Output: new collection of articles**
- \( N^2 > 50.000 \) articles that contain the nouns identified in the previous step.
- Creation of newspaper archive
Search for terms
Filter by year or by section
Saves the query and the collection of articles
Steps

<table>
<thead>
<tr>
<th>Καθημερινή</th>
<th>Αυγή</th>
<th>Σύνολο</th>
</tr>
</thead>
<tbody>
<tr>
<td>314527</td>
<td>226462</td>
<td>540989</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Καθημερινή</th>
<th>Αυγή</th>
<th>Σύνολο</th>
</tr>
</thead>
<tbody>
<tr>
<td>171616</td>
<td>191268</td>
<td>362884</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Επεξεργασμένα άρθρα</th>
</tr>
</thead>
<tbody>
<tr>
<td>Καθημερινή</td>
</tr>
<tr>
<td>Αυγή</td>
</tr>
<tr>
<td>Σύνολο</td>
</tr>
</tbody>
</table>
Η Αυγή, 11 Σεπτεμβρίου 2013
Μαχητική συγκέντρωση και πορεία διαμαρτυρίας στη Θεσσαλονίκη
Θεσσαλονίκη: Την ώρα που οι εκπαιδευτικοί διαδήλωναν μαζικά στο Σύνταγμα, στη Θεσσαλονίκη εκατοντάδες καθηγητές της δευτεροβάθμιας εκπαίδευσης συγκεντρώθηκαν στην πλατεία Αγίας Σοφίας και στη συνέχεια πραγματοποίησαν πορεία διαμαρτυρίας σε κεντρικούς δρόμους της πόλης, φωνάζοντας συνθήματα κατά της κινητικότητας, αλλά και ενάντια στο ταξικό «Νέο Λύκειο» του Κ. Αρβανιτόπουλου.

...
<table>
<thead>
<tr>
<th>ΠΟΤΕ (time)</th>
<th>ΠΟΥ (location)</th>
<th>ΠΟΙΟΣ (actor)</th>
<th>ΠΩΣ (form)</th>
<th>ΣΕ ΠΟΙΟΝ (addressee)</th>
<th>ΤΙ (issue)</th>
<th>For/Against whom(object actor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-9-2013</td>
<td>Θεσσαλονίκη</td>
<td>Καθηγητές</td>
<td>Συγκέντρωση/Πορεία</td>
<td>Υπ. Παιδείας Κ. Αρβανιτόπουλος</td>
<td>Κινητικότητα, ταξικό «Νέο Λύκειο»</td>
<td>Καθηγητές</td>
</tr>
<tr>
<td>CDAY=10</td>
<td>CITY=1120701</td>
<td>ACT1S=412</td>
<td>FORMS=7 FORM=72</td>
<td>ACT1S=102 ACTCOUN1=11 ACTPAR1=1</td>
<td>ISFIELD1=90</td>
<td>OBJ1S=412 OBJCOUN1=11 OBJPAR1=0</td>
</tr>
<tr>
<td>CMONTH=9</td>
<td>CYEAR=13</td>
<td>ACTTYP1=1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACTPAR1=0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-9-2013</td>
<td>Πλατεία Συντάγματος</td>
<td>Εκπαιδευτικοί</td>
<td>Διαδήλωση</td>
<td>Υπ. Παιδείας Κ. Αρβανιτόπουλος</td>
<td>Κινητικότητα, ταξικό «Νέο Λύκειο»</td>
<td>Εκπαιδευτικοί</td>
</tr>
<tr>
<td>CDAY=10</td>
<td>CITY=3514501</td>
<td>ACT1S=412</td>
<td>FORMS=7 FORM=72</td>
<td>ACT1S=102 ACTCOUN1=11 ACTPAR1=1</td>
<td>ISFIELD1=90</td>
<td>OBJ1S=412 OBJCOUN1=11 OBJPAR1=0</td>
</tr>
<tr>
<td>CMONTH=9</td>
<td>CYEAR=13</td>
<td>ACTTYP1=1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACTPAR1=0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Findings: Selected forms of action

<table>
<thead>
<tr>
<th>Form</th>
<th>Number of claims identified (1996-2013)</th>
<th>Percentage of total claims identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strike</td>
<td>2100</td>
<td>70.7%</td>
</tr>
<tr>
<td>Public assembly(^1)</td>
<td>718</td>
<td>24.2%</td>
</tr>
<tr>
<td>Occupation/Sit-in</td>
<td>97</td>
<td>3.3%</td>
</tr>
<tr>
<td>Symbolic violence(^2)</td>
<td>55</td>
<td>1.9%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2970</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

1. Public assemblies refer only to peaceful rallies (demonstrations on the move and violent demonstrations/assemblies are considered as different forms of action).
2. Symbolic violence includes the burning of flags, egg throwing, jeering, etc.
Figure 1: Claims per year

- **Strike**
- **Public assembly**
- **Occupation/Sit-in**
- **Symbolic violence**
- **Trend line (strike)**
- **Trend line (public assembly)**

**Key Events:**
- **1996:** National elections (NE)
- **1997:** Teachers' 8-week strike
- **1999:** Olympic Games
- **2004:** New social insurance law
- **2007:** Memorandum
- **2008:** Seafarers' & teachers' strikes
- **2009:** NE/CG
- **2010:** NE/CG
- **2011:** NE/CG
- **2013:** NE/CG

**Abbreviations:**
- NE: National elections
- CG: Change of government
Main findings

• Predominant claims are over labour, employment and educational issues
• Strikes show the most significant variation over time (escalating during the first years of Economic crisis)
• Strikes are largely labour-related and time-dependent
• Trade Unions are the prominent actors of strikes, but shortly after the Olympic Games, and before Memorandum, it seems that they loose their ability to form and synthesize claims
• Symbolic violence occurs mainly in the Capital (Athens) and shows a stable evolution over time
• The expectation of a government change, has a negative effect on protest
A Social Data Analytics platform

www.socioscope.gr
Socioscope Highlights

- **Social Data Analytics**
  - Data Exploration
  - Visual Analytics on social and political data

- **Social Data Visualization**
  - Dynamic Visualizations
  - Combination of data from Different Thematic areas

- **Linked Open Data availability**
  - Reuse of codebook
  - API endpoint for online analysis
Data sources

Dashboards

Data overview

Highlighted Diagrams

Various types of diagrams (bar, line, area, pie)
Choropleth Maps
Hierarchical Exploration (roll-up, drill-down)
Full Text Search

Facets Browsing for filtering

Social conflicts and protests - PROMAP

Facets:
- Newspaper
- Newspaper Issue Date
- Claim Data
- City
- Actor/Claimant Summary
- Addressee Summary
- Addressee Country
- Form of action (Summary)
- Form of action
- Issue field

Mobile Support
References


Thank you for your attention

theosta@ekke.gr