

Σημασιολογική Ανάλυση Κειμένου για την εξέταση κοινωνικών φαινομένων και συμπεριφορών

Maria Pontiki

Computational Linguist

Institute for Language & Speech Processing (ATHENA RIC/ILSP), Greece

www.athenarc.gr

Introduction

- **Semantic Analysis**

- extract meaning from text

- build structures that approximate concepts from a large set of documents

- e.g. Topic Modelling, Event Analysis, Sentiment Analysis



Aspect Based Sentiment Analysis in Customer Reviews



Outline

- **Examining Xenophobia in Greece during the economic crisis:
A computational perspective**
- **Annotation exercise**



Examining Xenophobia in Greece during the economic crisis: A computational perspective



www.athenarc.gr

<http://xenophobia.ilsp.gr>

Xenophobia

"**attitudes, prejudices and behavior that reject, exclude and often vilify persons, based on the perception that they are outsiders or foreigners to the community, society or national identity.**"
(UNESCO definition)



Xenophobia in Greece

- Xenophobia as a violent practice against predefined Target Groups (TGs) of Interest
- Violent Practice:
 - physical attacks (event analysis)
 - verbal attacks (sentiment analysis)
- TGs: Albanians, Pakistani, Romanians, Syrians, Islam/Muslims, Roma, Germans, Jews, Immigrants and Refugees.



Examining Xenophobia in Greece during the economic crisis: A computational perspective

Research Questions



Prejudices
&
Stereotypes

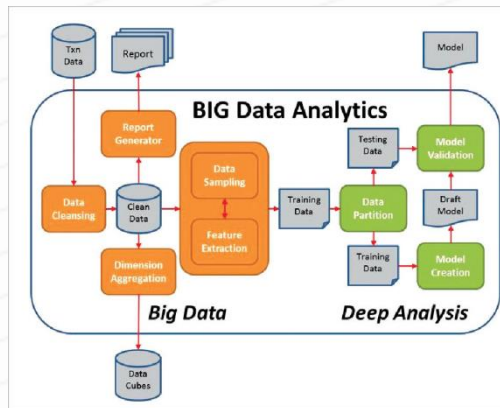


Economic
crisis

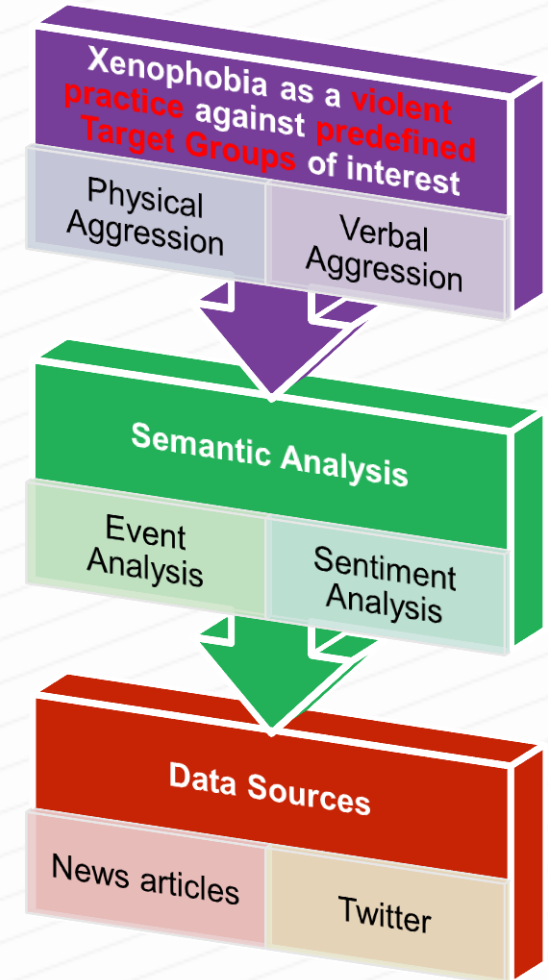


The Basic
Factor of
xenophobic
sentiments
among
Greeks

Methodology



Research Strand



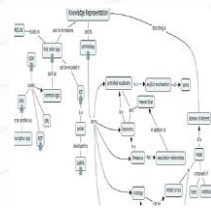
Methodology



Data
Collection



Data
Exploration



Knowledge
Representation



Data Analysis



Visualization



Interpretation



Data Collection: News Articles

➤ 7 Newspapers

➤ ≈ 3.600.000 articles

Avgi	Kathimerini	Eleftherotypia	Rizospastis	TaNea	In.gr	Naftemporiki
792.715	282.621	429.364	725.108	330.190	428.880	649.259
1996-2015	2002-2006, 2009-2012	2002-2006, 2008-2014	1995- 07/2016	1997-2007	1999- 21/09/2016	2000- 21/09/2016

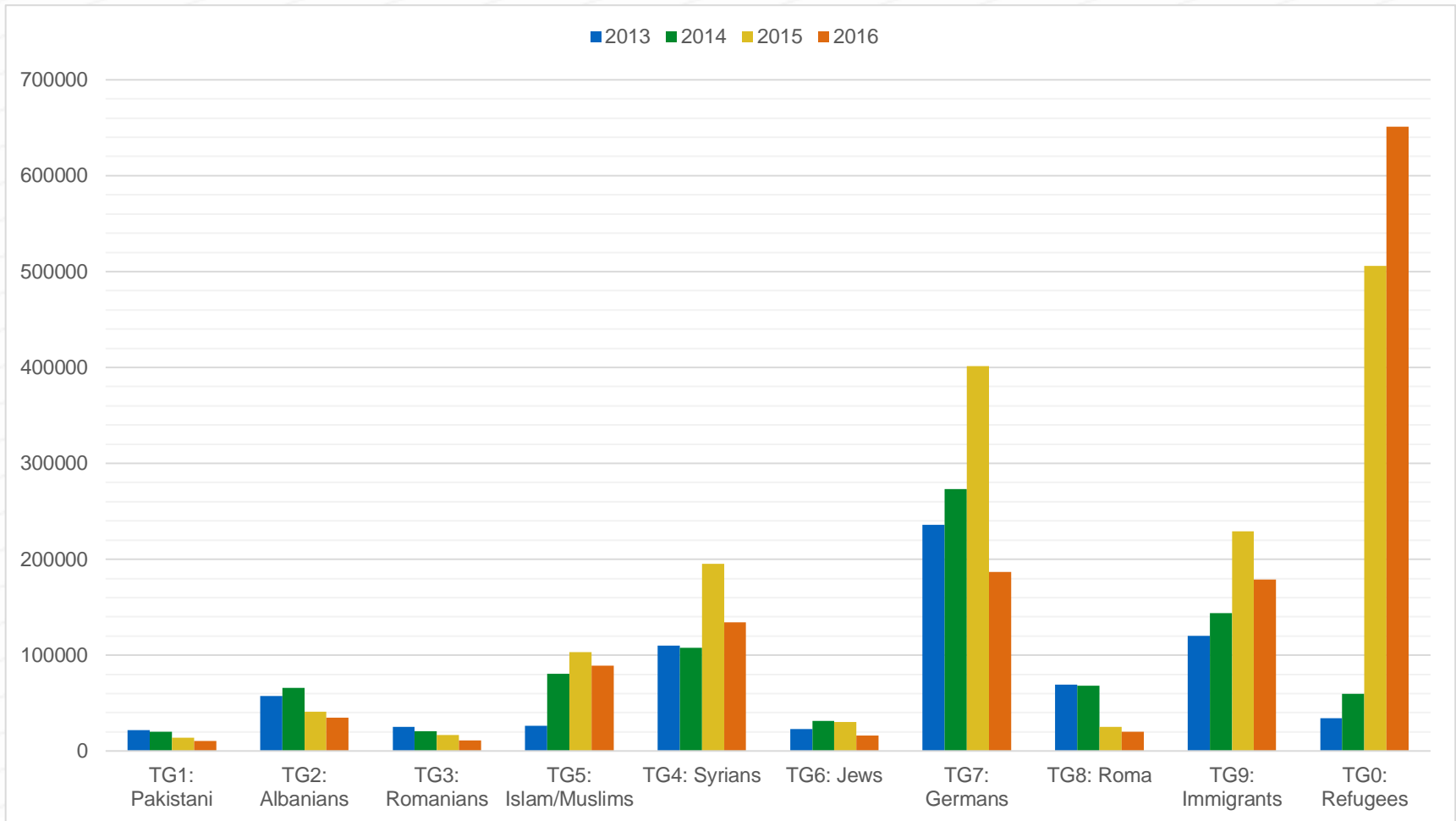
➤ Time period: 1995-2016



Data Collection: Twitter

➤ **≈ 500.000.000 Tweets**

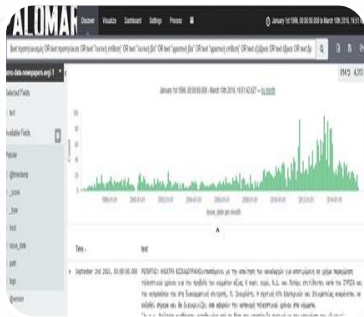
➤ **Time period 2013-2016**



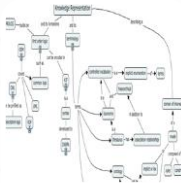
Methodology



Data Collection



Data Exploration



Knowledge Representation



Data Analysis



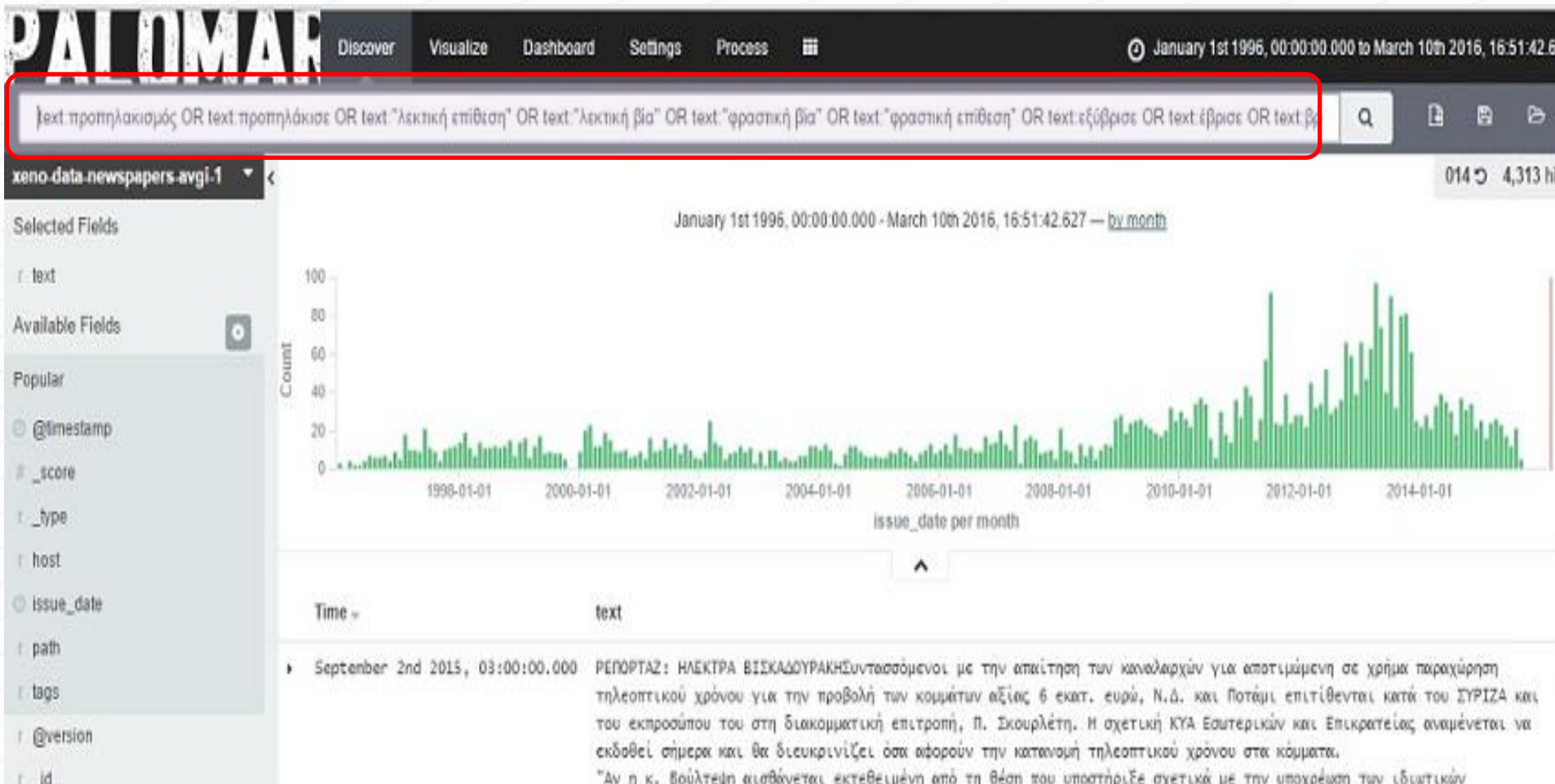
Visualization



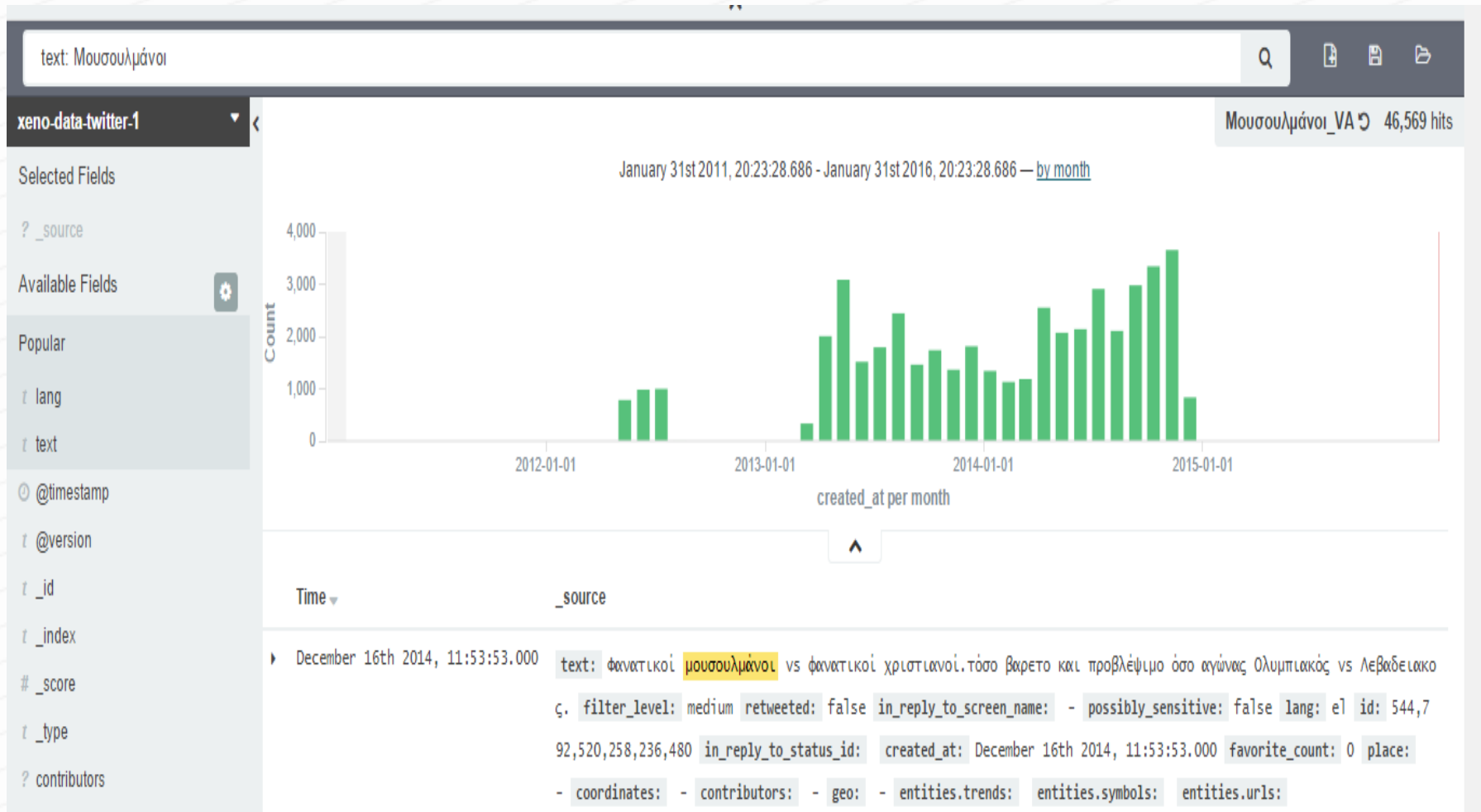
Interpretation



Data Exploration: News



Data Exploration: Tweets



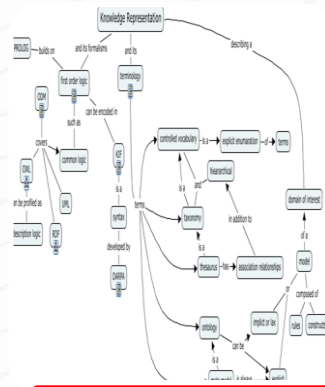
Methodology



Data Collection



Data Exploration



Knowledge Representation



Data Analysis



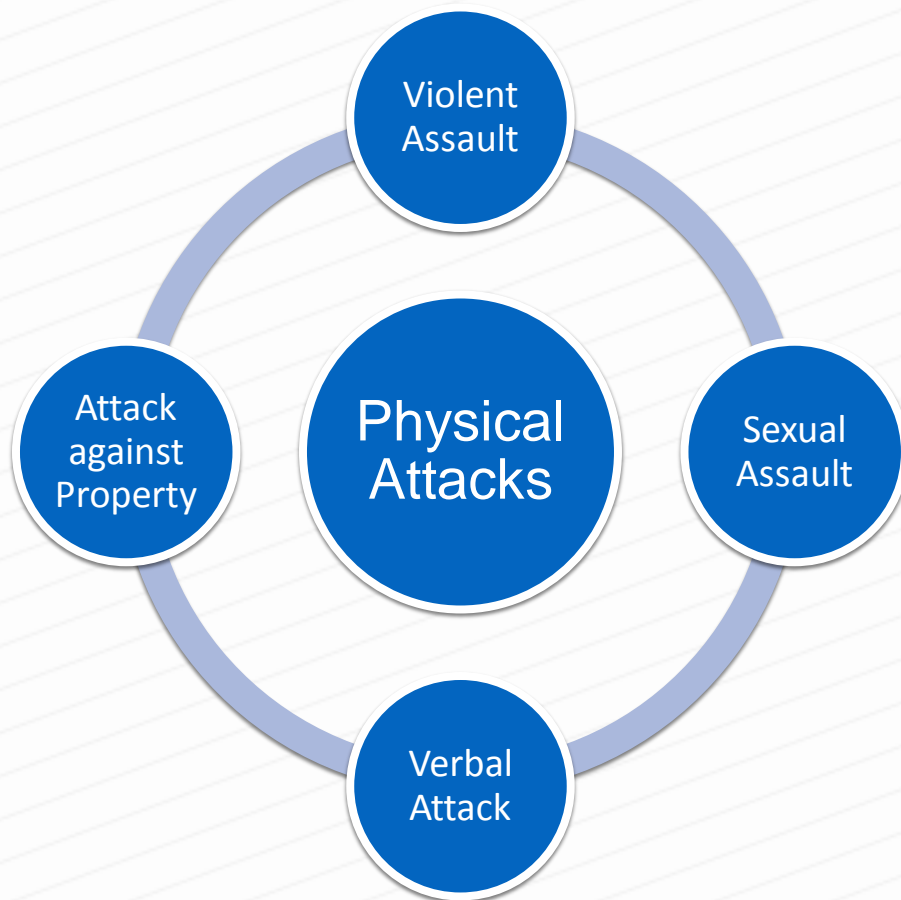
Visualization



Interpretation



Knowledge Representation: Codebook – Event Taxonomy



EVENT

- The word or phrase representing an event type under examination, which is located within the text.
- Features: *Event type*.

ACTOR

- The entity that performs each event instance.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

TARGET

- The entity to whom the action is addressed.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

LOCATION

- Where the event took place.
- Features: *Category*.

TIME

- When the event happened.
- Features: *Day, Month, Year*.

CONFIDENCE

- Captures whether in the article there is any indication that an Actor of an assault may not be the actual perpetrator.
- Features: *Degree*.



Codebook – Event Taxonomy

EVENT

- The word or phrase representing an event type under examination, which is located within the text.
- Features: *Event type*.

ACTOR

- The entity that performs each event instance.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

TARGET

- The entity to whom the action is addressed.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

LOCATION

- Where the event took place.
- Features: *Category*.

TIME

- When the event happened.
- Features: *Day, Month, Year*.

CONFIDENCE

- Captures whether in the article there is any indication that an Actor of an assault may not be the actual perpetrator.
- Features: *Degree*.

Στις 28 Μαρτίου 2013, στο κέντρο της Ξάνθης, 30χρονος Έλληνας μουσουλμάνος **τραυματίστηκε με μαχαίρι** από μέλη της Χ.Α.



Codebook – Event Taxonomy

EVENT

- The word or phrase representing an event type under examination, which is located within the text.
- Features: *Event type*.

ACTOR

- The entity that performs each event instance.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

TARGET

- The entity to whom the action is addressed.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

LOCATION

- Where the event took place.
- Features: *Category*.

TIME

- When the event happened.
- Features: *Day, Month, Year*.

CONFIDENCE

- Captures whether in the article there is any indication that an Actor of an assault may not be the actual perpetrator.
- Features: *Degree*.

Στις 28 Μαρτίου 2013, στο κέντρο της Ξάνθης, 30χρονος Έλληνας μουσουλμάνος τραυματίστηκε με μαχαίρι από **μέλη της Χ.Α.**



Codebook – Event Taxonomy

EVENT

- The word or phrase representing an event type under examination, which is located within the text.
- Features: *Event type*.

ACTOR

- The entity that performs each event instance.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

TARGET

- The entity to whom the action is addressed.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

LOCATION

- Where the event took place.
- Features: *Category*.

TIME

- When the event happened.
- Features: *Day, Month, Year*.

CONFIDENCE

- Captures whether in the article there is any indication that an Actor of an assault may not be the actual perpetrator.
- Features: *Degree*.

Στις 28 Μαρτίου 2013, στο κέντρο της Ξάνθης, **30χρονος Έλληνας μουσουλμάνος** τραυματίστηκε με μαχαίρι από μέλη της Χ.Α.



Codebook – Event Taxonomy

EVENT

- The word or phrase representing an event type under examination, which is located within the text.
- Features: *Event type*.

ACTOR

- The entity that performs each event instance.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

TARGET

- The entity to whom the action is addressed.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

LOCATION

- Where the event took place.
- Features: *Category*.

TIME

- When the event happened.
- Features: *Day, Month, Year*.

CONFIDENCE

- Captures whether in the article there is any indication that an Actor of an assault may not be the actual perpetrator.
- Features: *Degree*.

Στις 28 Μαρτίου 2013, στο κέντρο της **Ξάνθης**, 30χρονος Έλληνας μουσουλμάνος τραυματίστηκε με μαχαίρι από μέλη της Χ.Α.



Codebook – Event Taxonomy

EVENT

- The word or phrase representing an event type under examination, which is located within the text.
- Features: *Event type*.

ACTOR

- The entity that performs each event instance.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

TARGET

- The entity to whom the action is addressed.
- Features: *Summary, TG, Nationality, Age, Sex, Status*.

LOCATION

- Where the event took place.
- Features: *Category*.

TIME

- When the event happened.
- Features: *Day, Month, Year*.

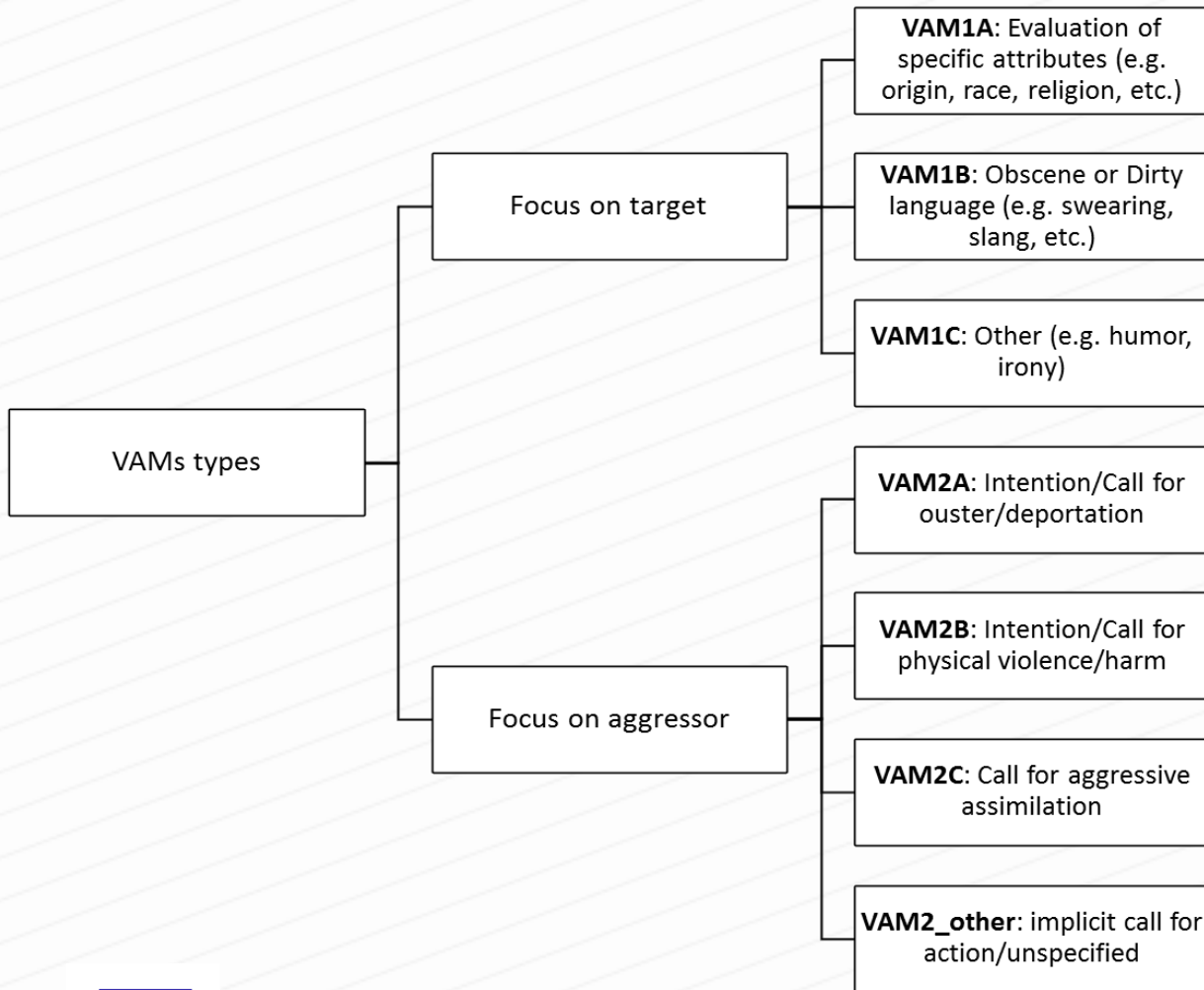
CONFIDENCE

- Captures whether in the article there is any indication that an Actor of an assault may not be the actual perpetrator.
- Features: *Degree*.

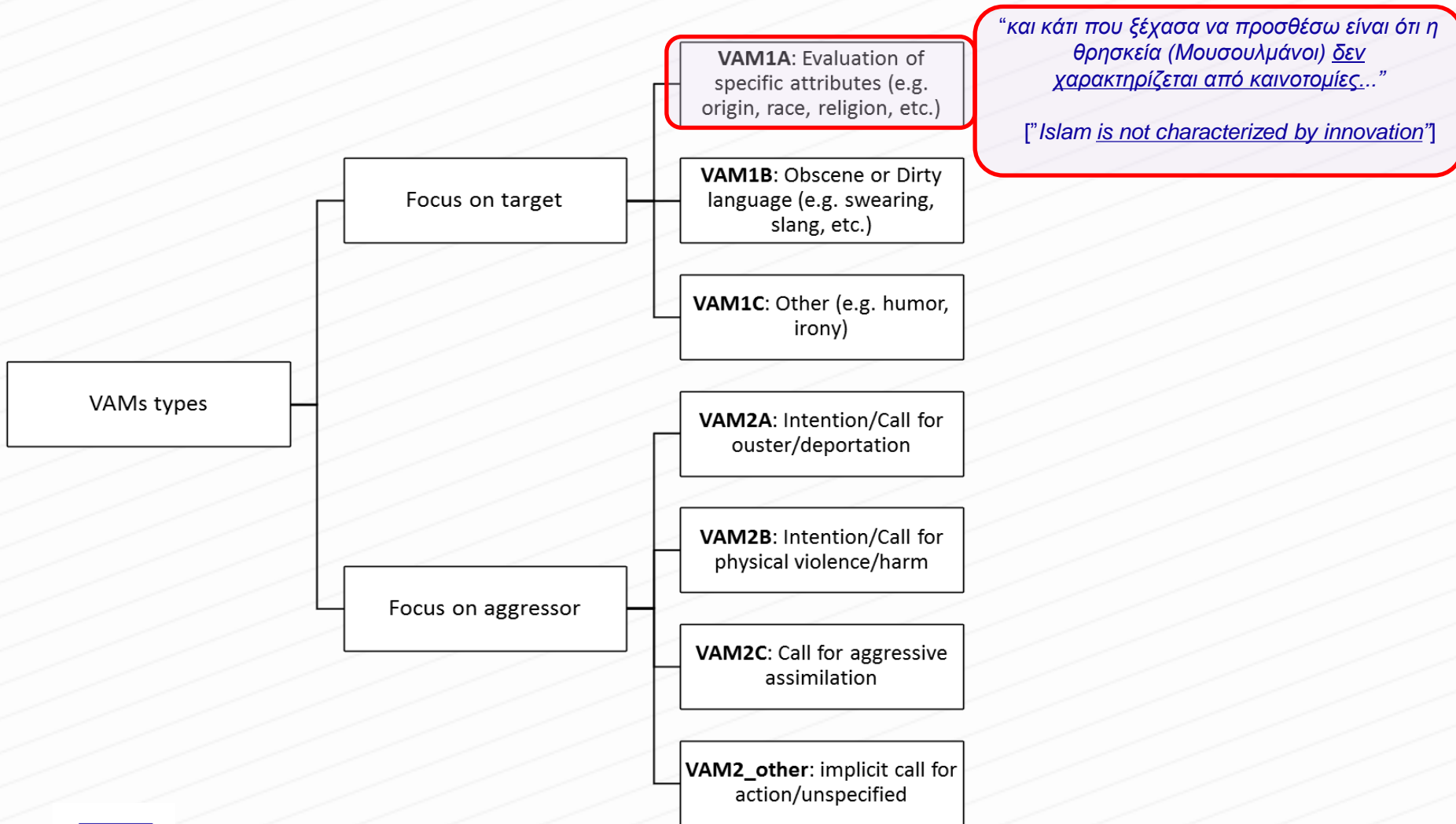
ΣΤΙΣ **28 Μαρτίου 2013**, στο
κέντρο της Ξάνθης, 30χρονος
Ελληνας μουσουλμάνος
τραυματίστηκε με μαχαίρι από
μέλη της Χ.Α.



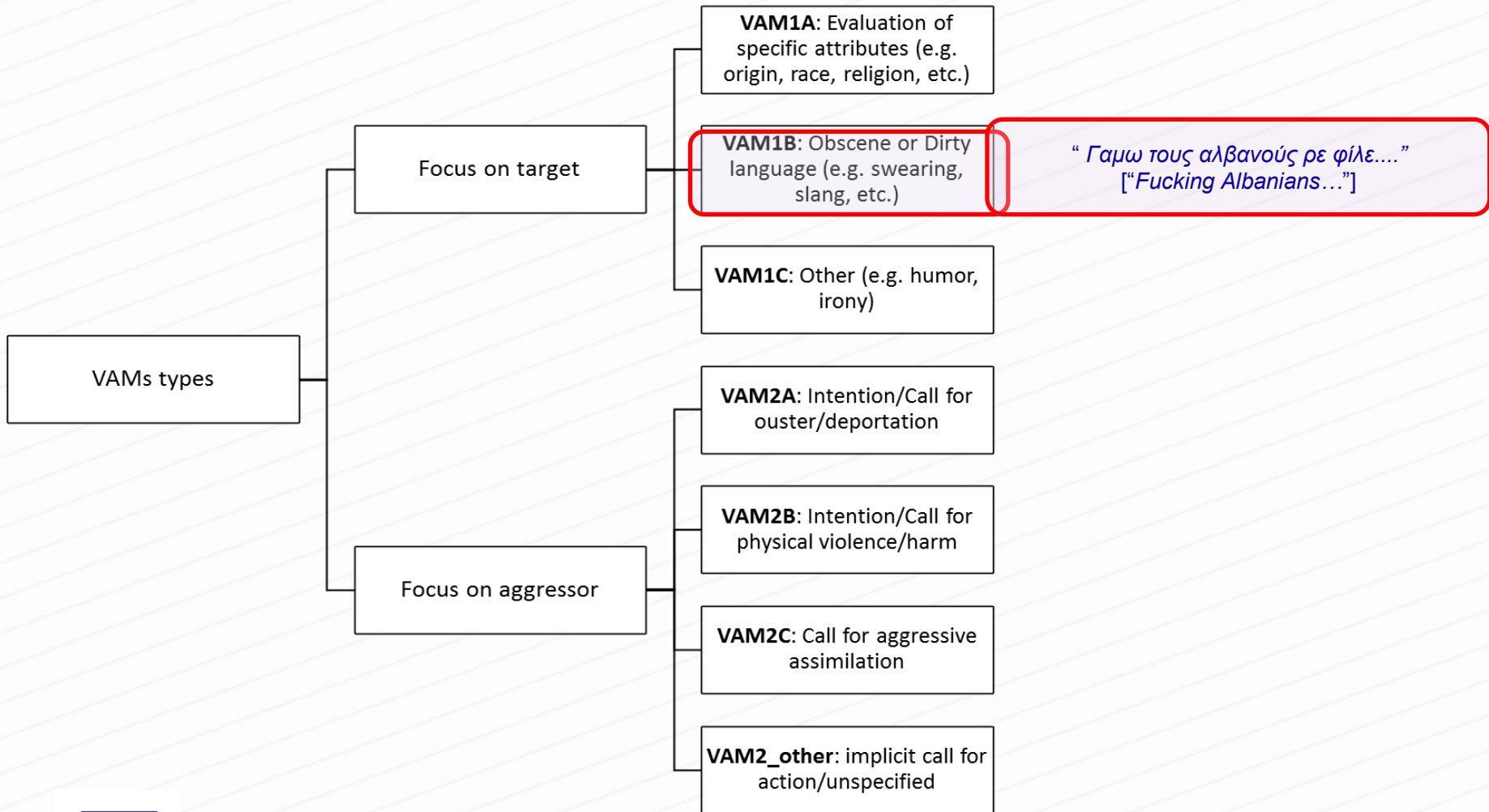
Knowledge Representation: Verbal Aggressiveness Framework



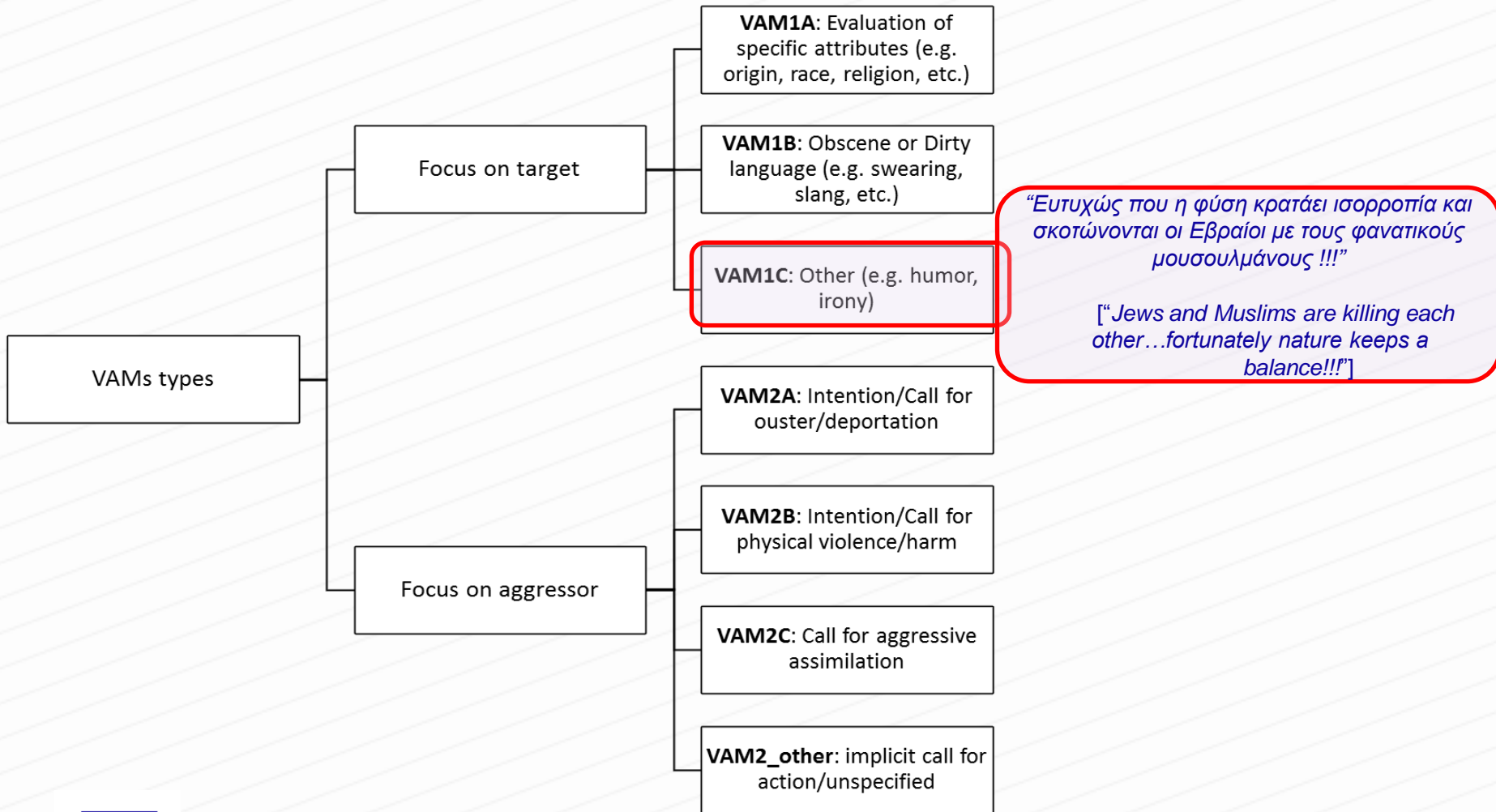
Verbal Aggressiveness Framework



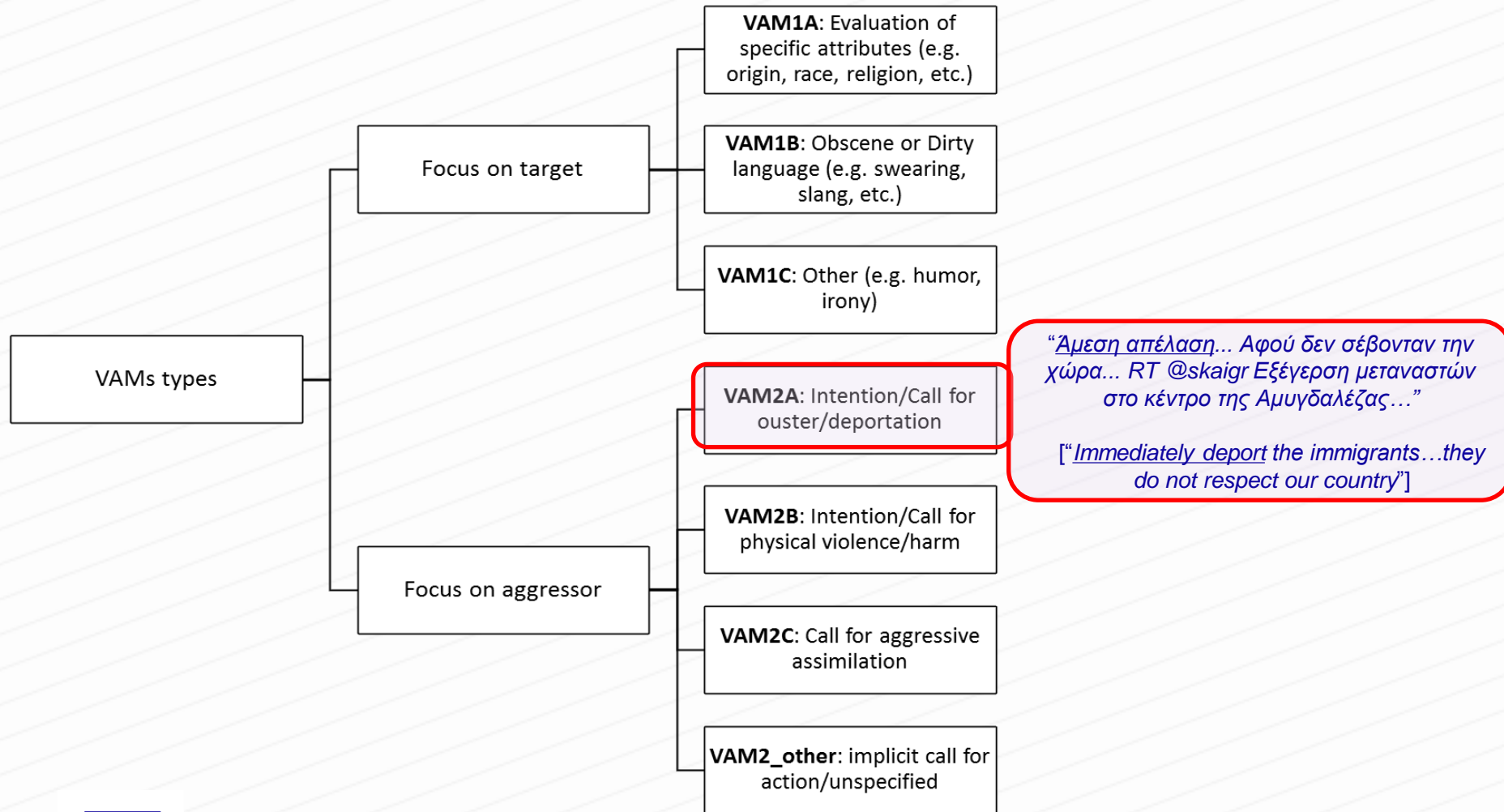
Verbal Aggressiveness Framework



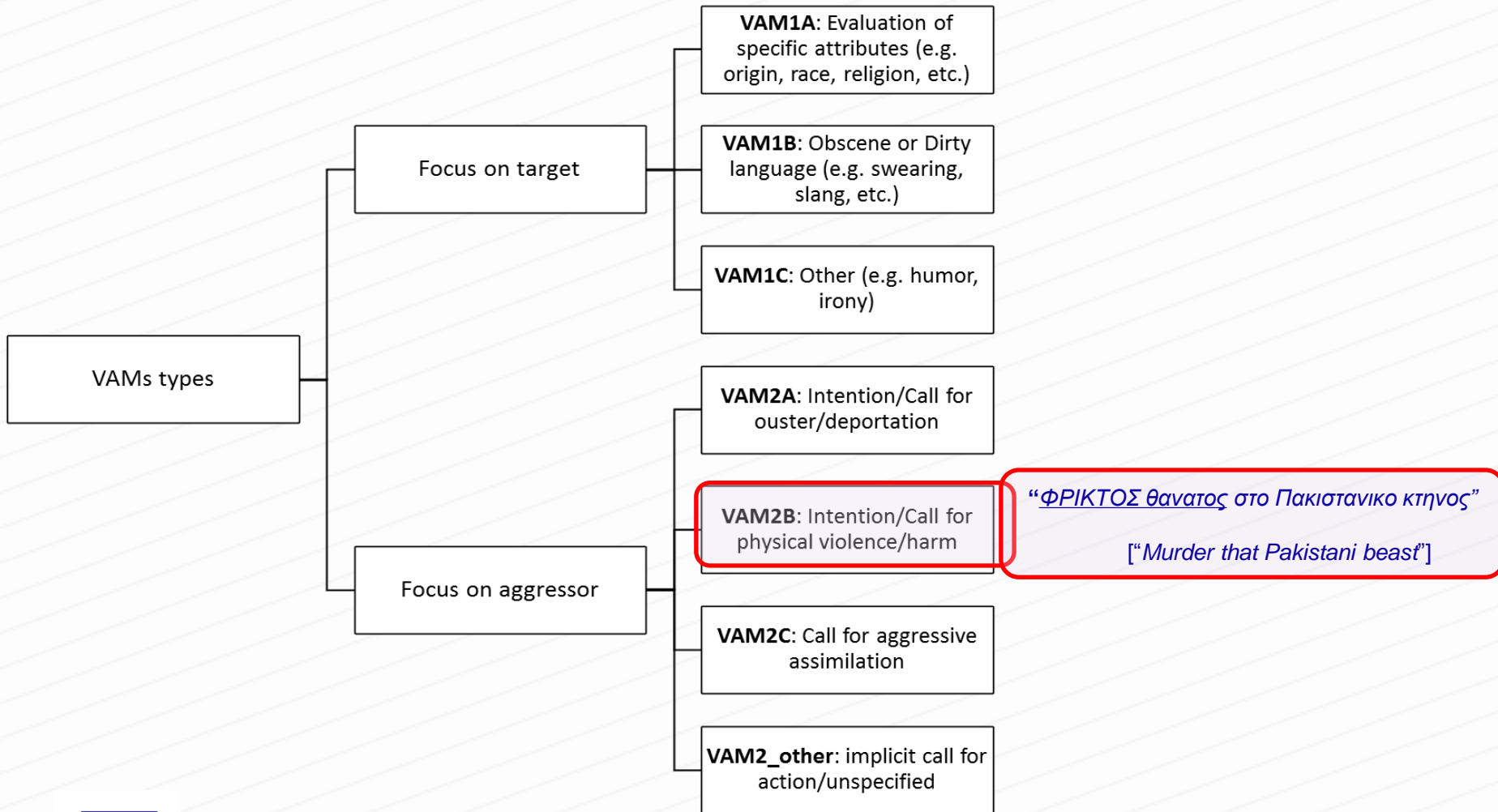
Verbal Aggressiveness Framework



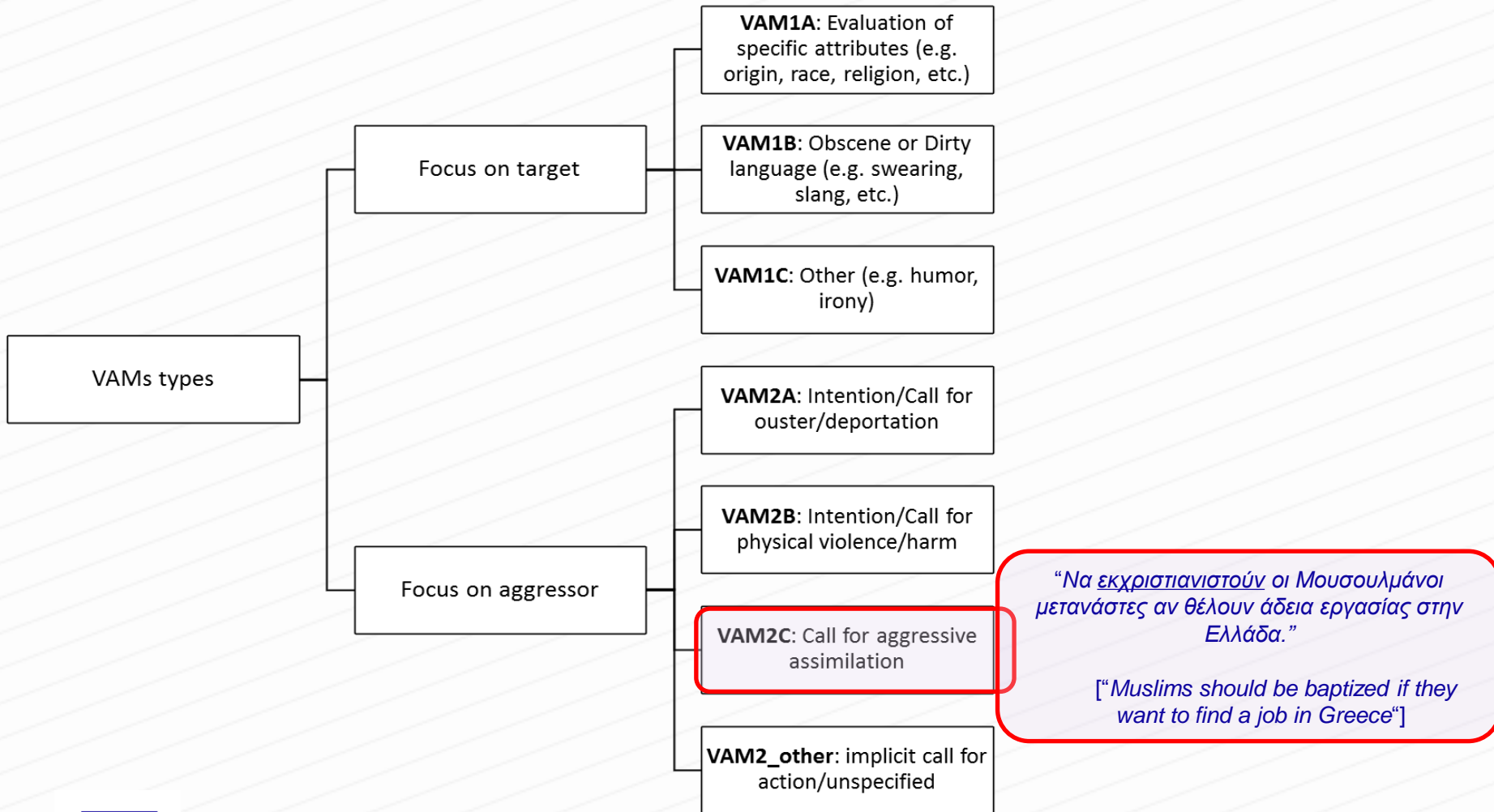
Verbal Aggressiveness Framework



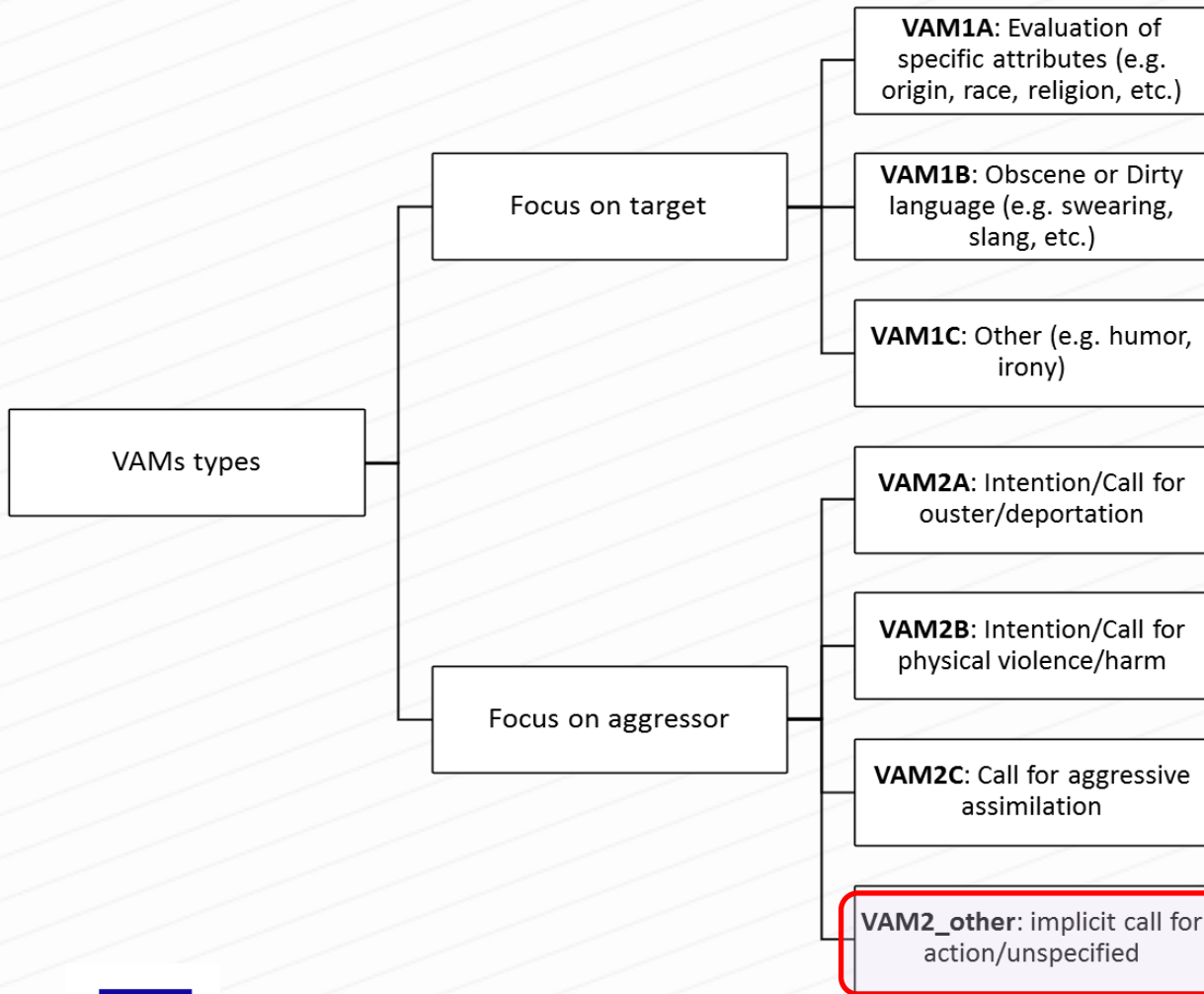
Verbal Aggressiveness Framework



Verbal Aggressiveness Framework



Verbal Aggressiveness Framework



“Θα συνεχίσουμε να κάνουμε τους χαζούς μπροστά στον ισλαμικό κίνδυνο;”

[“We will keep pretending that there is no Islamic danger?”]



Methodology



Data
Collection



Data
Exploration



Knowledge
Representation



Data
Analysis



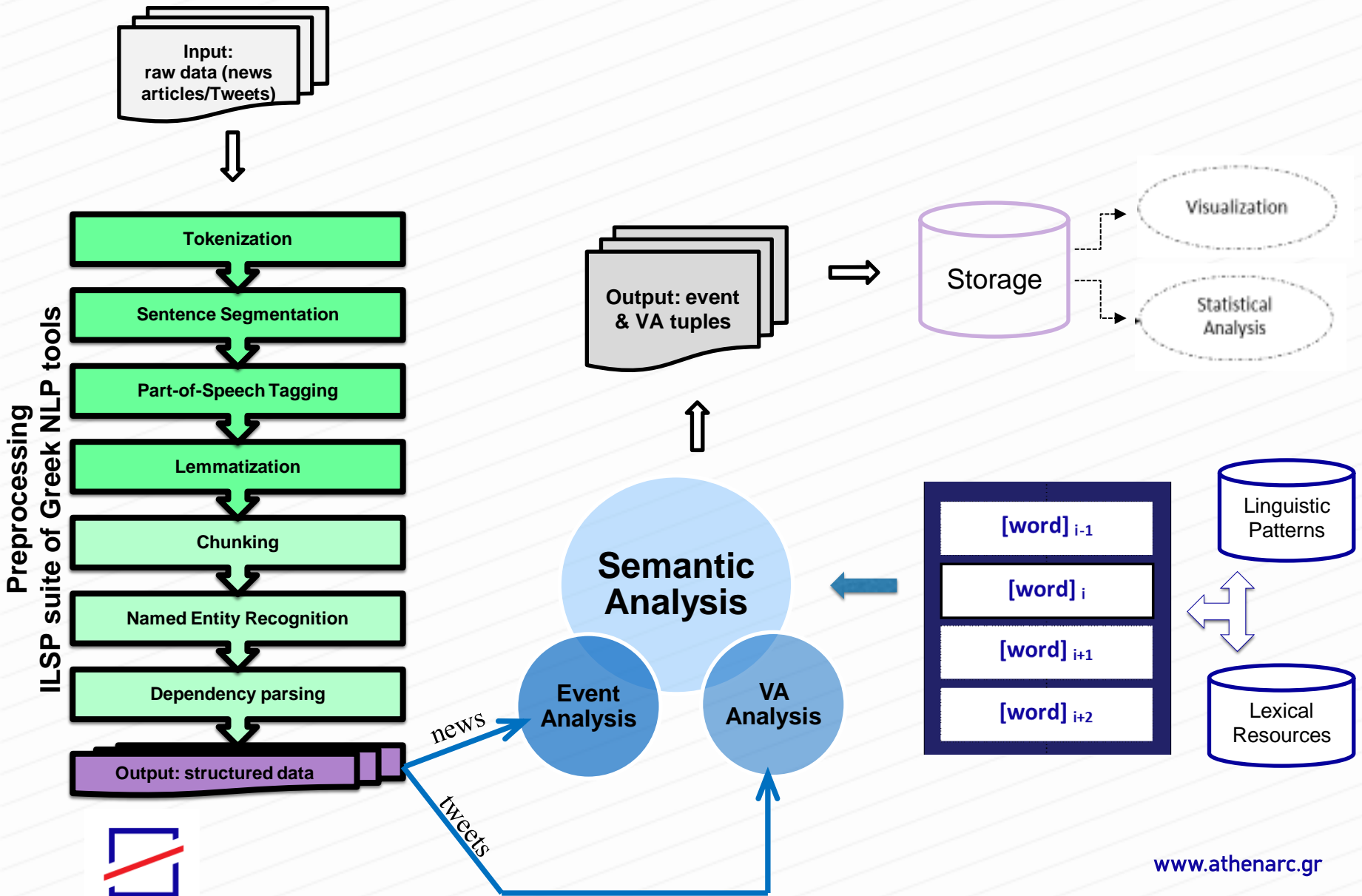
Visualization



Interpretation



Data Analysis Architecture



Verbal Aggressiveness Analysis

Input

- Φρικτός θάνατος στον Πακιστανό της Πάρου!

Output

- <**TG_ID**: 1, **TG_EVIDENCE**: Πακιστανό, **VAM_TYPE**: VAM2B, **VA_EVIDENCE**: θάνατος>

Other metadata:

- ✓ Tweet_Id
- ✓ User_Id
- ✓ Tweet Text
- ✓ Timestamp



Event Analysis

Input

- Στις 28 Μαρτίου 2013, στο κέντρο της Ξάνθης, 30χρονος Έλληνας μουσουλμάνος τραυματίστηκε με μαχαίρι από μέλη της Χ.Α.

Output

- <**TIME**: Στις 28 Μαρτίου 2013, **LOCATION**: στο κέντρο της Ξάνθης, **TARGET**: 30χρονος Έλληνας μουσουλμάνος **EVENT**: τραυματίστηκε με μαχαίρι **ACTOR**: από μέλη της Χ.Α.>
- ✓ **Target_Age**: 30
- ✓ **Target_TG**: 5
- ✓ **Target_Nationality**: Greek
- ✓ **Event_Type**: Assault against Life
- ✓ **Actor_TG**: Control Group
- ✓ **Actor_Summary**: Racist, extreme right organizations/groups

Other metadata:
Event_Id, Medium, Article
Title



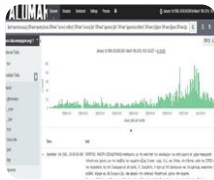
Semantic Analysis Output: Knowledge Databases

Excel interface showing a spreadsheet with columns: B, C, D, E, F, G, H, I, J. The data includes columns for evidence type (TG_evidence), target ID (Target_id), VAM type (VAM_type), VA evidence (VA_evidence), Year, Month, Day, user ID (user_id), and text. The text column contains various entries, including mentions of "B«ΜίλησανB»", "ΑΛΒΑΝΟΣ ΝΤΑΗΣ ΕΠΙΤΕΘΗΚΕ ΚΑΙ ΤΡΑΥΜΑΤΙΣΕ 54ΧΡΟΝΗ ΕΛΛΗΝΙΔΑ!", and "RT @platitudinous: Έξω οι κωλοέλληνες απ'τη Γερμανία!".

Methodology



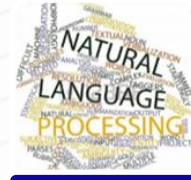
Data
Collection



Data
Exploration



Knowledge
Representation



Data
Analysis



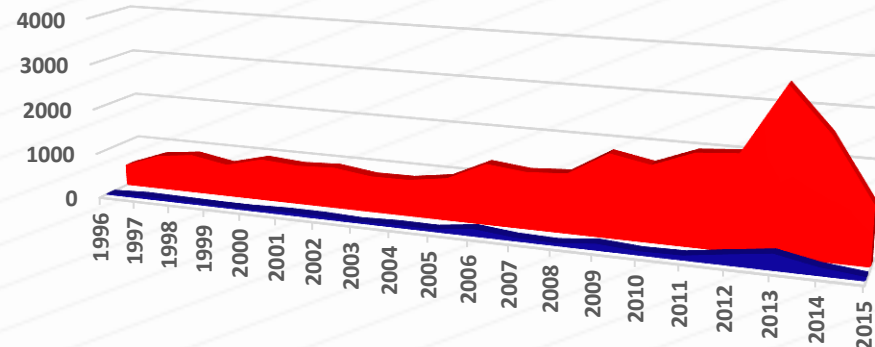
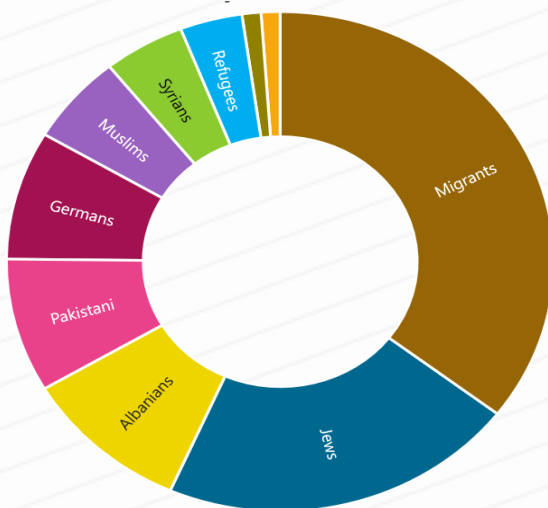
Visualization



Interpretation

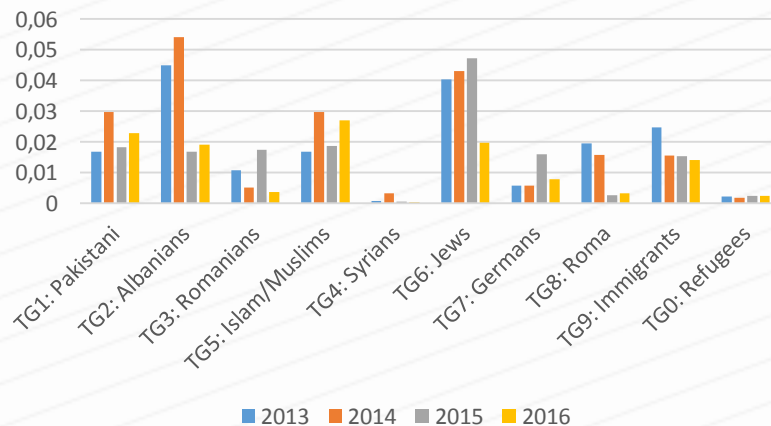


Visualization: Examples



PA Timeline TGs vs CG

Targets of Physical Attacks: Naftemporiki 2000-2016



Per-year VA rate (VAMs/Tweets) per TG



Word Cloud of unique aggressive terms for “Muslims/Islam”

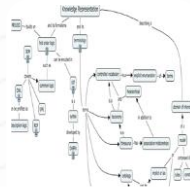
Methodology



Data Collection



Data Exploration



Knowledge Representation



Data Analysis



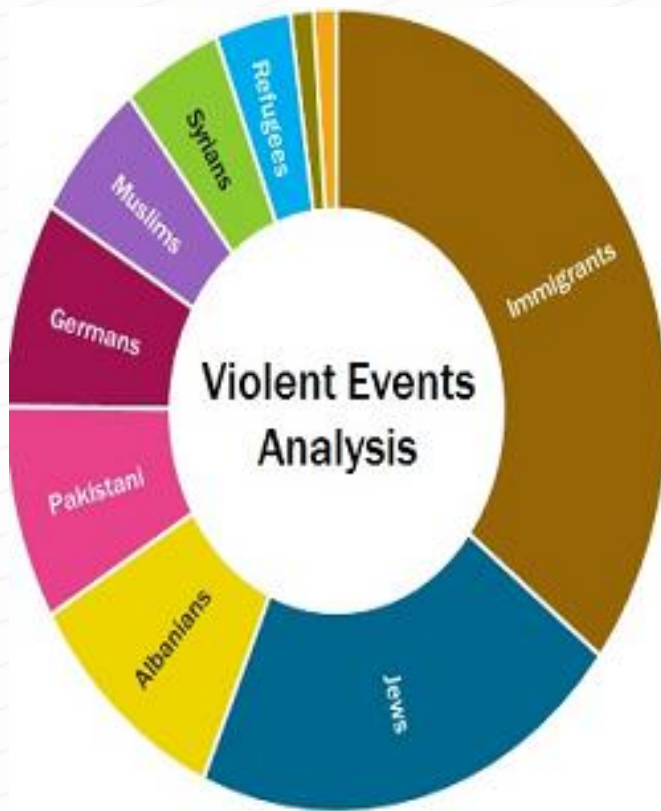
Visualization



Interpretation



Who are the main targets of physical and verbal aggression?



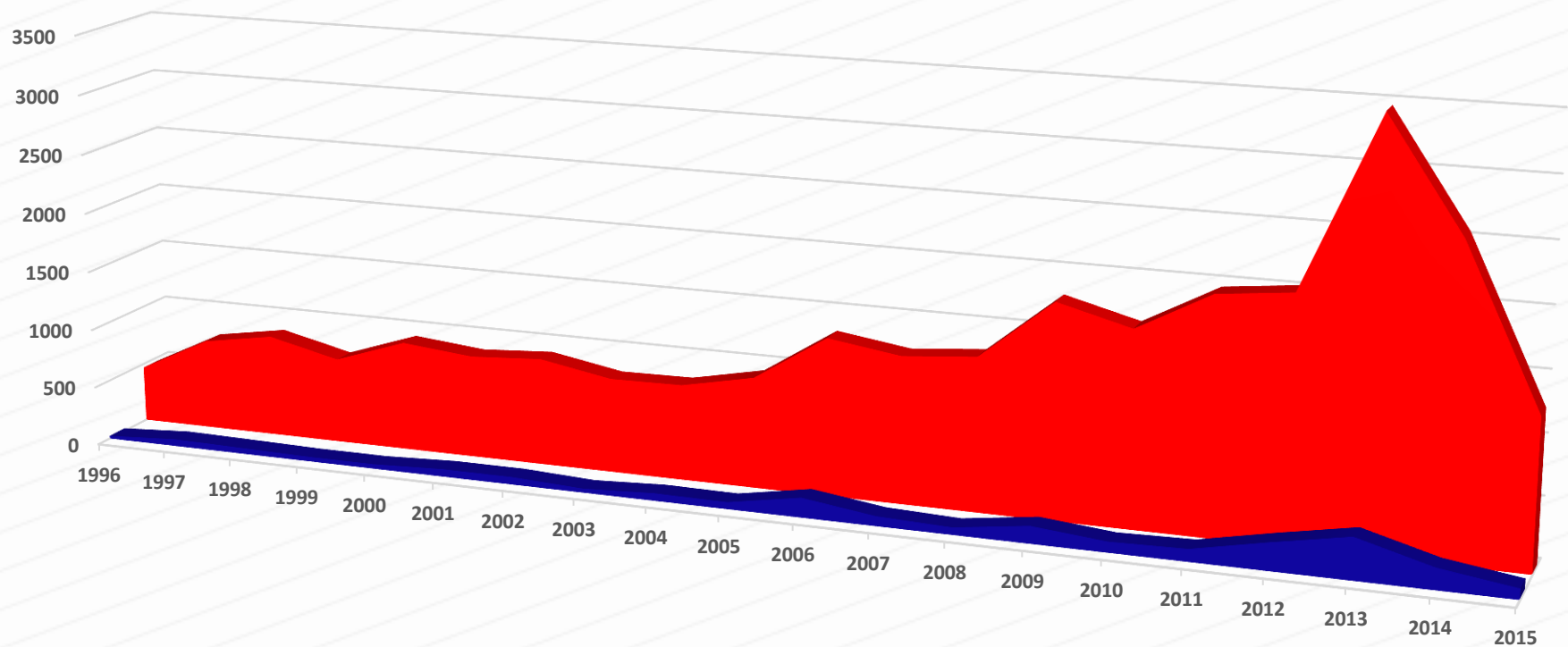
Who are the main targets of physical and verbal aggression?

- Physical and verbal aggression as indicators of xenophobic attitudes in Greece seem to be addressed to the same targets.
→ possible correlations between verbal aggressiveness and physical violence?
- According to the 2016 ADL Global 100 survey Greece was the most anti-Semitic country in Europe (69%).
- The image of foreigner as it was constructed in Greece during and after the first wave of migration flow was mainly associated with Albanian nationality (Voulgaris et al., 1995).
- PAKISTANIS came to Greece during the 2000s increased migration flow from Asian countries.
- According to the 2016 PEW survey findings a big majority of Greeks holds an unfavorable view of Muslims (65%).
- It is more likely to verbally attack groups of people framed as IMMIGRANTS rather than REFUGEES due to the different connotations and implications of these two lexicalizations.



Have the xenophobic attitudes raised during the economic crisis?

Escalation of violent incidents in general.

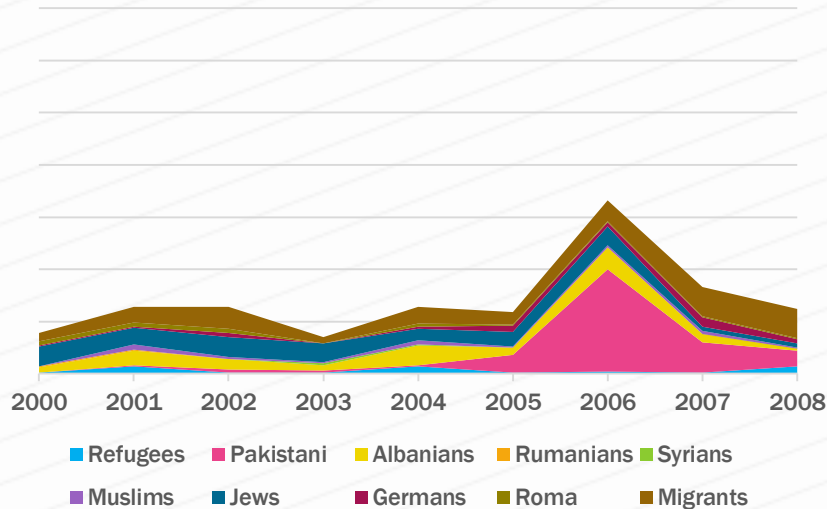


■ TGs ■ CG

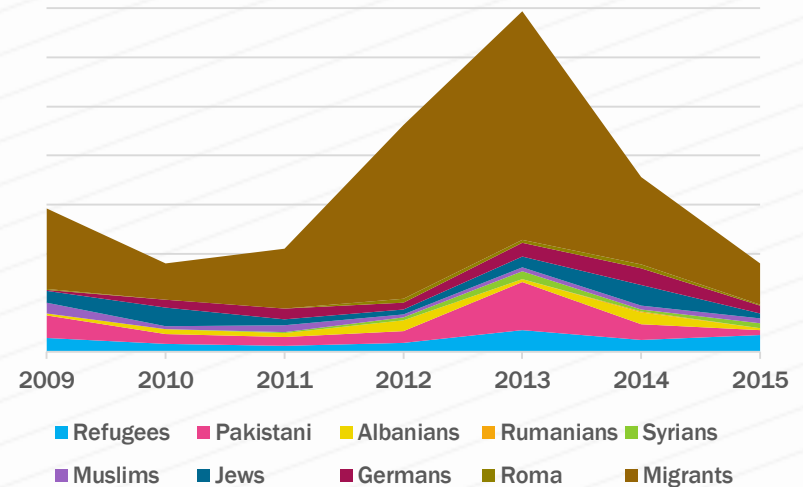
Physical Attacks Timeline

www.athenarc.gr

Have the xenophobic attitudes raised during the economic crisis?



Mostly Attacked TGs before the economic crisis (Avg)

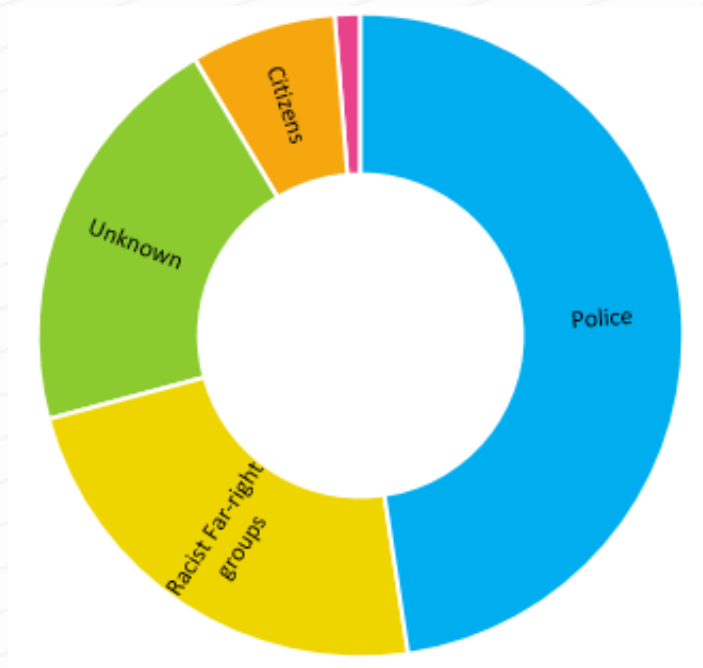


Mostly Attacked TGs during the economic crisis (Avg)

- Continuity is observed against Pakistani, Albanians and Jews.
- New targets seem to emerge depending on the contemporary conditions affecting the country; Germans appear as targeted group, with the attacks against them rising as the economic crisis deepens.

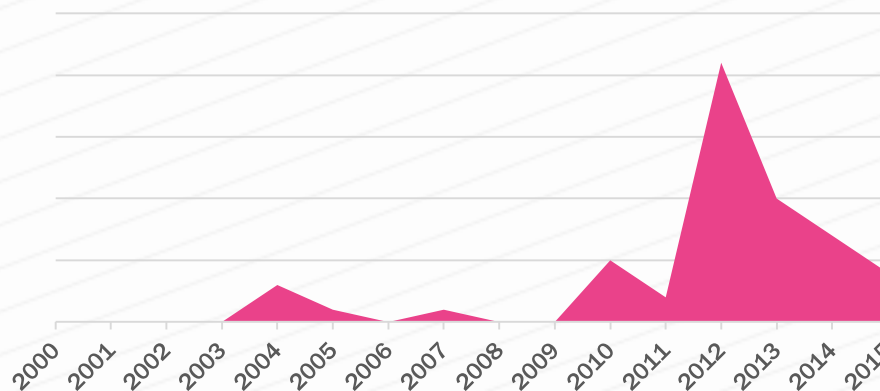


Who are the main actors of physical aggression?



- Unknown: indications that point to features of extreme right groups (e.g. attributes like “wearing black t-shirts”) or “unidentified person” or “a group of people”, without giving any further details.
- Citizens as actors don’t seem to represent a major proportion.

Racist Far-right groups



Are there stereotypes and prejudices against foreigners rooted deeply in the Greek society?



Word Cloud of unique aggressive terms for “Albanians”



Word Cloud of unique aggressive terms for “Pakistani”

- The dominant stereotypes in the construction of the image of **ALBANIANS** are associated with “*crime*” and “*cultural inferiority*”.
- Such crime stereotypes are used also in the construction of the image of **ROMANIANS**, in addition to the sexual decadence (lax rules of sexual morality) stereotype for the Romanian women.
- A continuity of the so-called stereotype of the Balkanian criminal.
- **PAKISTANI:** inferiority, personal hygiene, physical appearance, color skin stereotypes.



Are there stereotypes and prejudices against foreigners rooted deeply in the Greek society?



Word Cloud of unique aggressive terms for "Muslims/Islam"

- Evaluative and dysphemistic terms of insult or abuse to debase core Islamic values, principles, practices.
- Brutal violence, cultural inferiority, sexual abuse, sexist behavior, fanaticism, terrorism and irrationalism stereotypes.
- Notions of Islamophobia.
- Attacks targeted to **MUSLIMS/ISLAM** are triggered by geopolitical events such as the rise of ISIS.
- Attacks against **SYRIANS** target the violent practices of the Syrian rebels and ISIS.
- Attacks against **REFUGEES** and **IMMIGRANTS** are mostly attempts to challenge their identity.



Are there stereotypes and prejudices against foreigners rooted deeply in the Greek society?



Word Cloud of unique aggressive terms for “Jews”

- A perception of a particular enmity towards the Greek nation.
- Identification with the negative aspects of the banking system and global capitalism, frequent appeal to conspiracy theory elements .
- Avarice stereotype.
- Notions of hate-speech.

- A perception of a particular enmity towards the Greek nation, blame attribution patterns of the Greek crisis.
- Memories and symbols of WWII and of Nazi occupation of Greece.
- Lack of delight, cultural inferiority, poor aesthetics (bad taste) stereotype.



Word Cloud of unique aggressive terms for “Germans”



Further Insights

■ *Xenophobia in Greece: Vulnerability-driven or Superiority-based?*

→ Dominance is directed against Albanians and Pakistani.

→ Vulnerability in the case of Jews and Germans.

■ *Xenophobia in Greece: Culturally-rooted or crisis-driven?*

→ when examined in terms of Aggression towards specific predefined TGs of interest, seems to be culturally rooted and not crisis-driven.

→ continuity of deeply rooted stereotypes about specific TGs (e.g. Albanians, Romanians, Jews, Roma).

→ emergence of attacks that are associated with blame attribution patterns of the Greek crisis (e.g. Germans, Jews).



Limitations

- **Validity:** Precision-oriented method, a significant amount of attacks has not been detected and explored.
- **Representativeness:** A single platform study, cannot capture the wider social ecology and diffusion of Social Media. No Social Media platform is representative of the general population.
- **Completeness:** The results presented do not reflect all (aspects/types of) xenophobic attacks against all foreigners.





Maria Pontiki

Computational Linguist

Institute for Language & Speech Processing

(ATHENA RIC/ILSP), Greece



www.athenarc.gr