

Σημασιολογική Ανάλυση Κειμένου για την εξέταση κοινωνικών φαινομένων και συμπεριφορών

Maria Pontiki

Computational Linguist

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



Introduction

- Semantic Analysis
 - →extract meaning from text

 \rightarrow 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**

XENO@GR IF Ceea

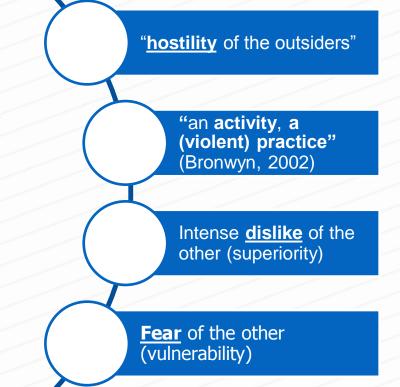


www.athenarc.gr

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)



 \leq

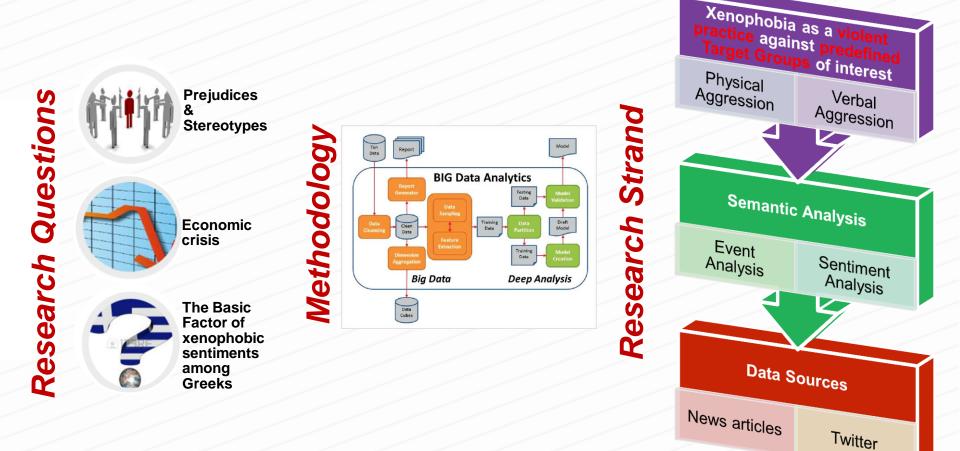
Xenophobia in Greece

 Xenophobia as a violent practice against predefined Target Groups (TGs) of Interest

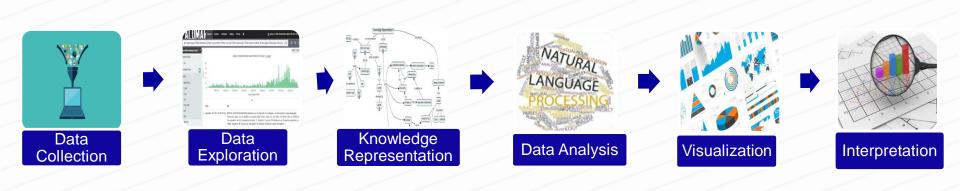
 Violent Practice: -physical attacks (event analysis)
 -verbal attacks (sentiment analysis)

 TGs: Albanians, Pakistani, Romanians, Syrians, Islam/Muslims, Roma, Germans, Jews, Immigrants and Refugees.

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



Methodology

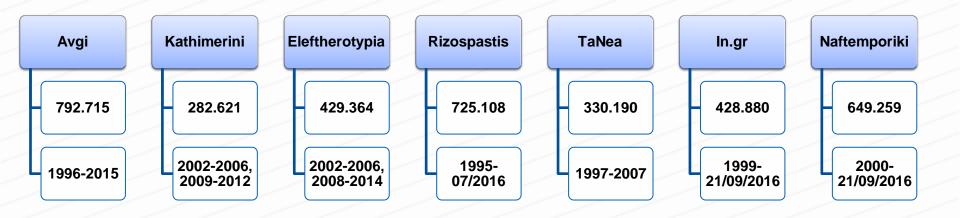




Data Collection: News Articles

7 Newspapers

➤ ≈ 3.600.000 articles



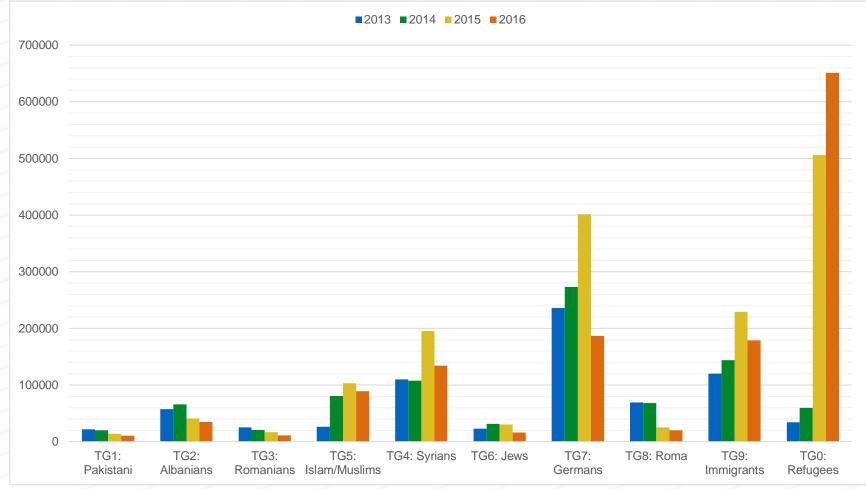
Time period: 1995-2016



Data Collection: Twitter

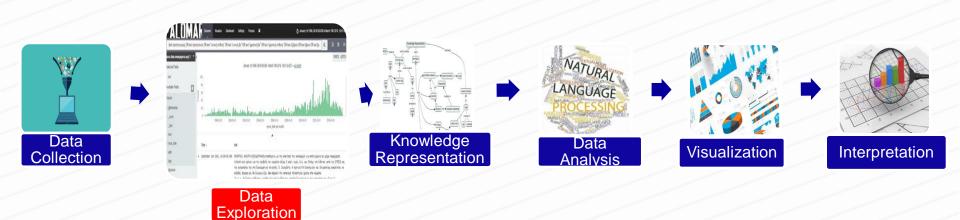
➤ ≈ 500.000.000 Tweets

Time period 2013-2016



 \leq

Methodology

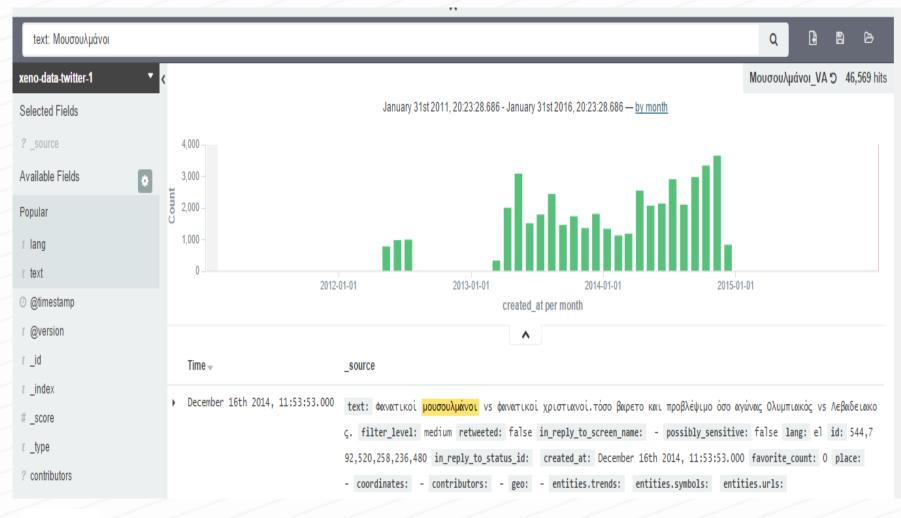


Data Exploration: News



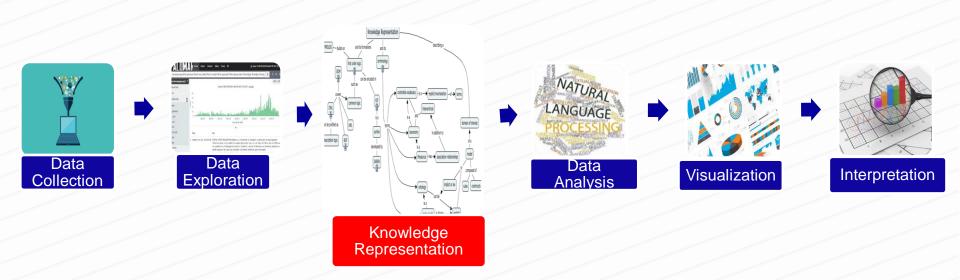
 \ge

Data Exploration: Tweets



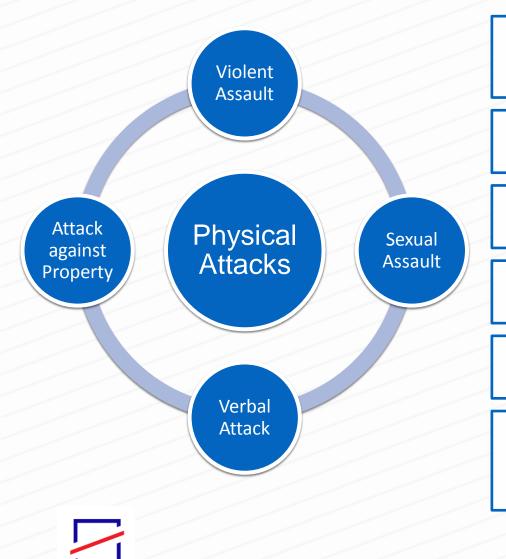
 \ge

Methodology





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.

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.

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.

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.



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.

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.

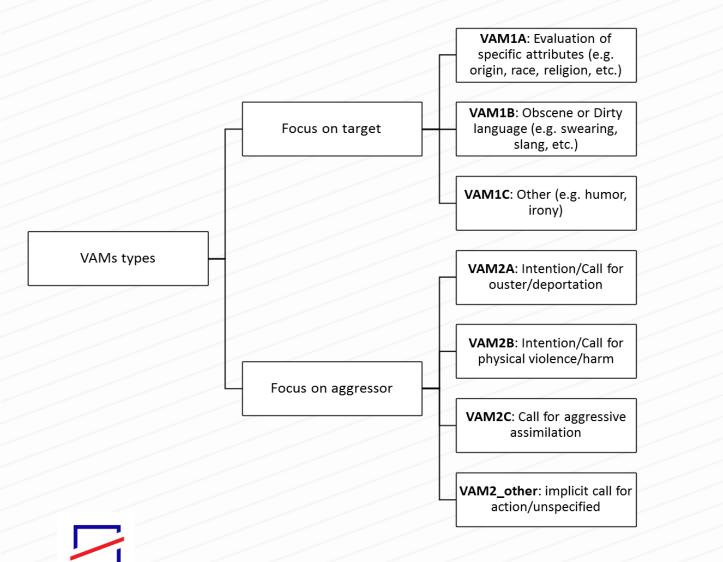
CONFIDENCE

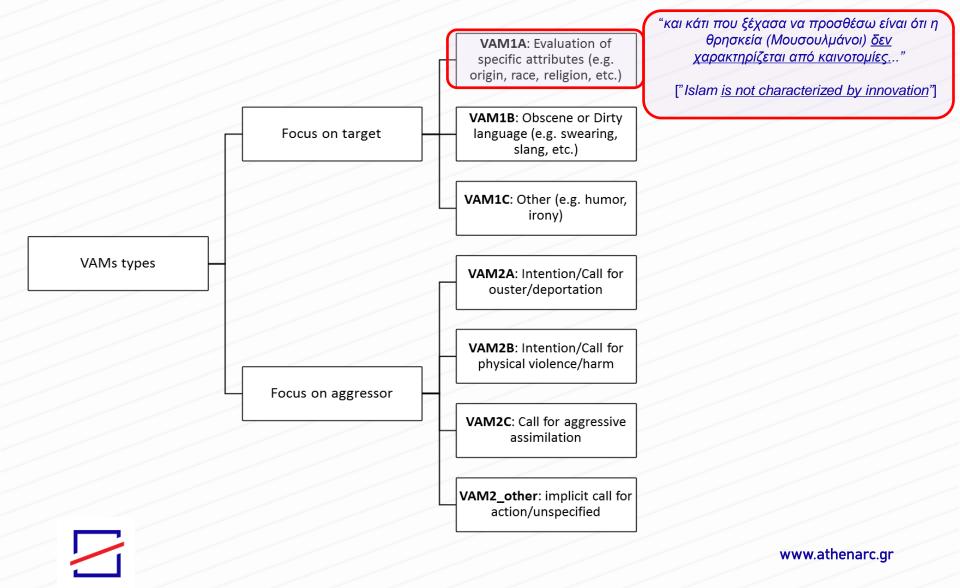
- Captures whether in the article there is any indication that an Actor of an assault may not be the actual perpetrator.
- Features: Degree.

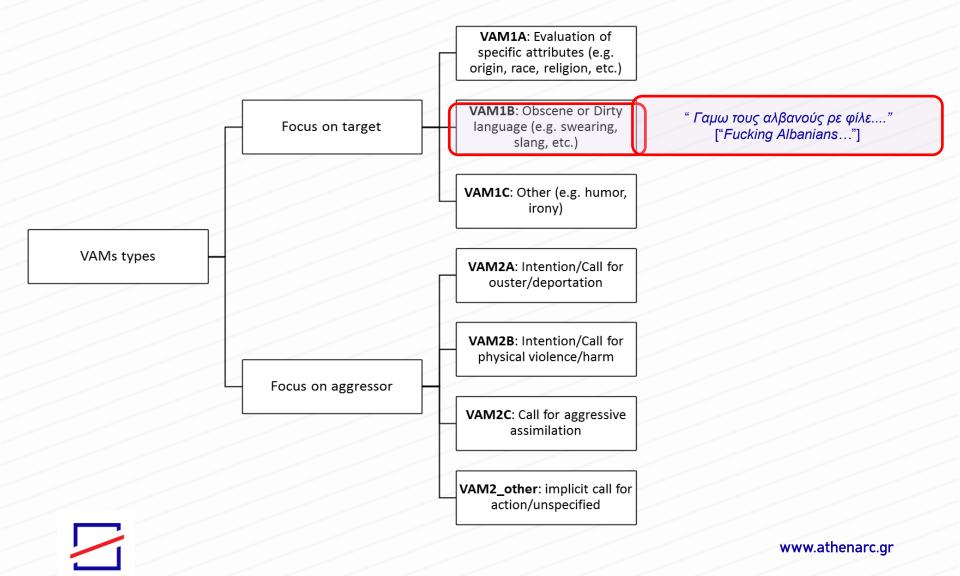
Στις **28 Μαρτίου 2013**, στο κέντρο της Ξάνθης, 30χρονος Ελληνας μουσουλμάνος τραυματίστηκε με μαχαίρι από μέλη της Χ.Α.

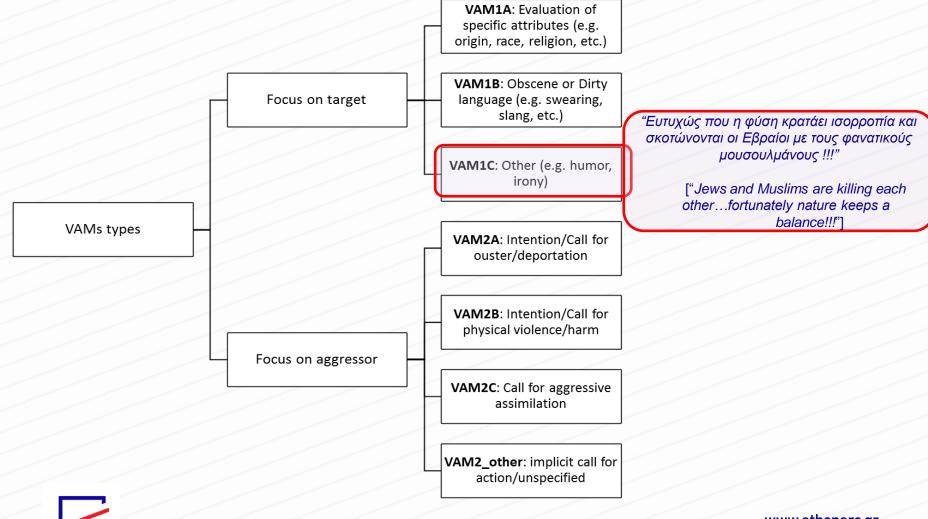
 \geq

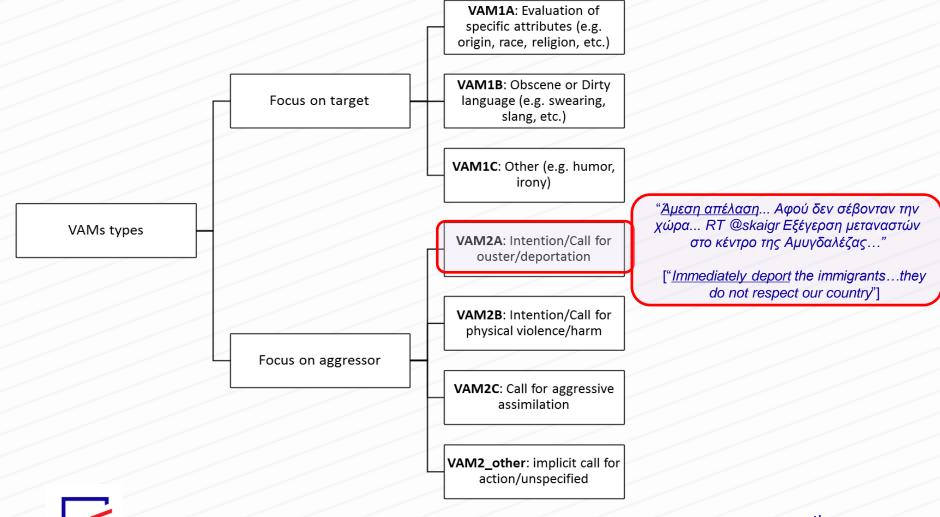
Knowledge Representation: Verbal Aggressiveness Framework

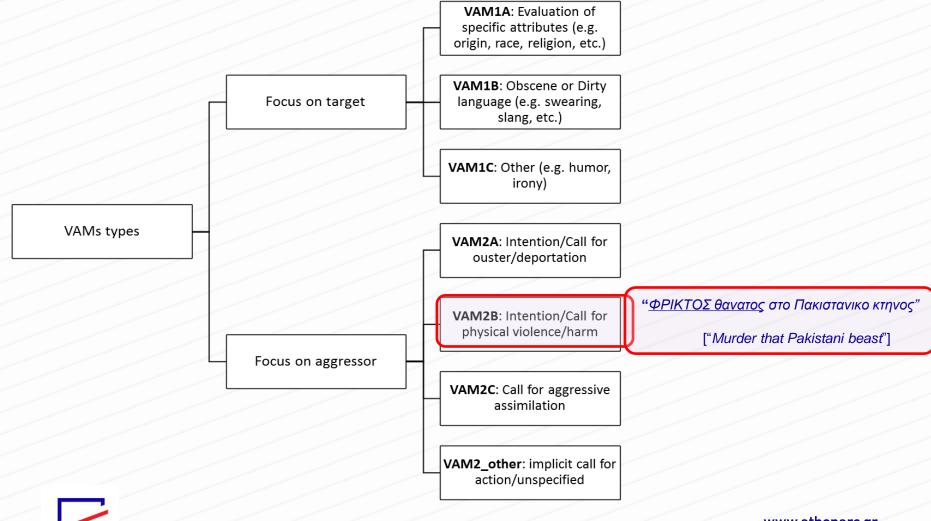


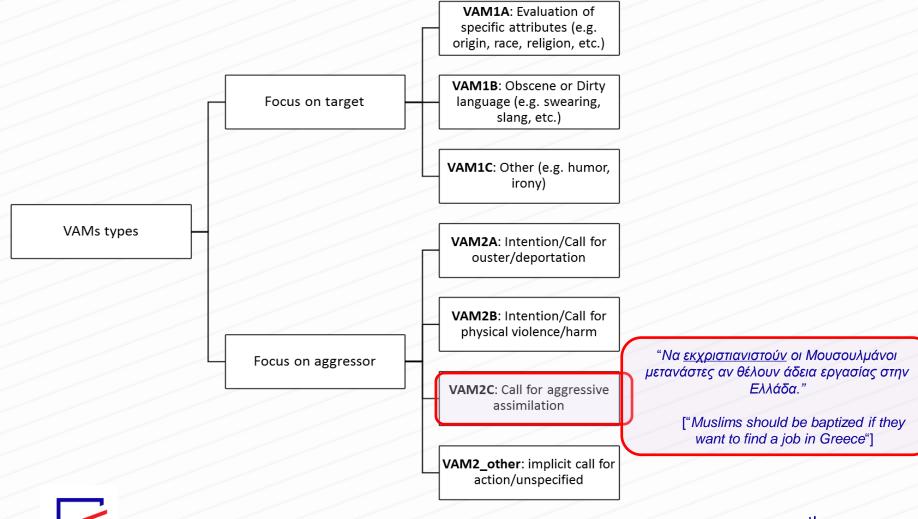


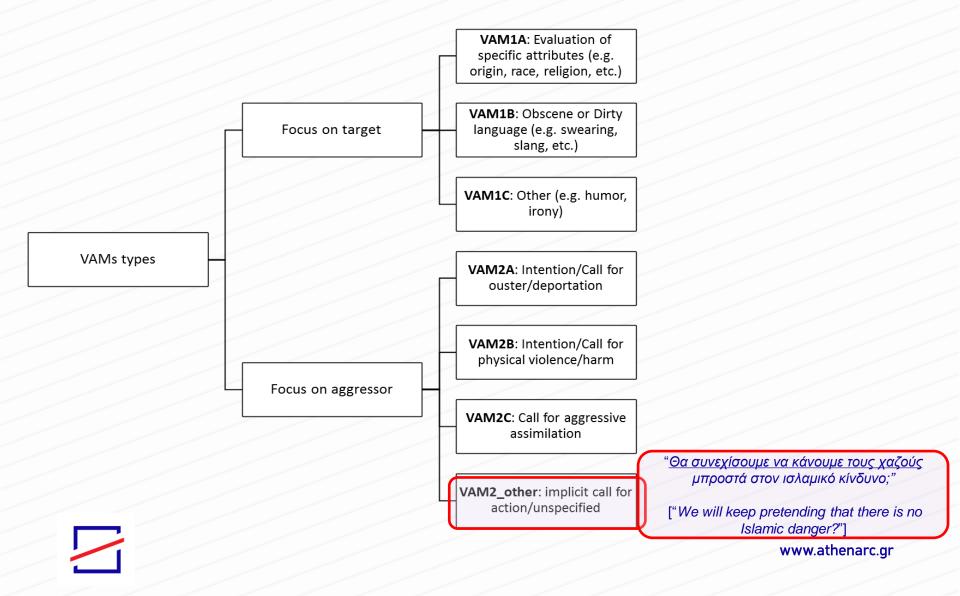












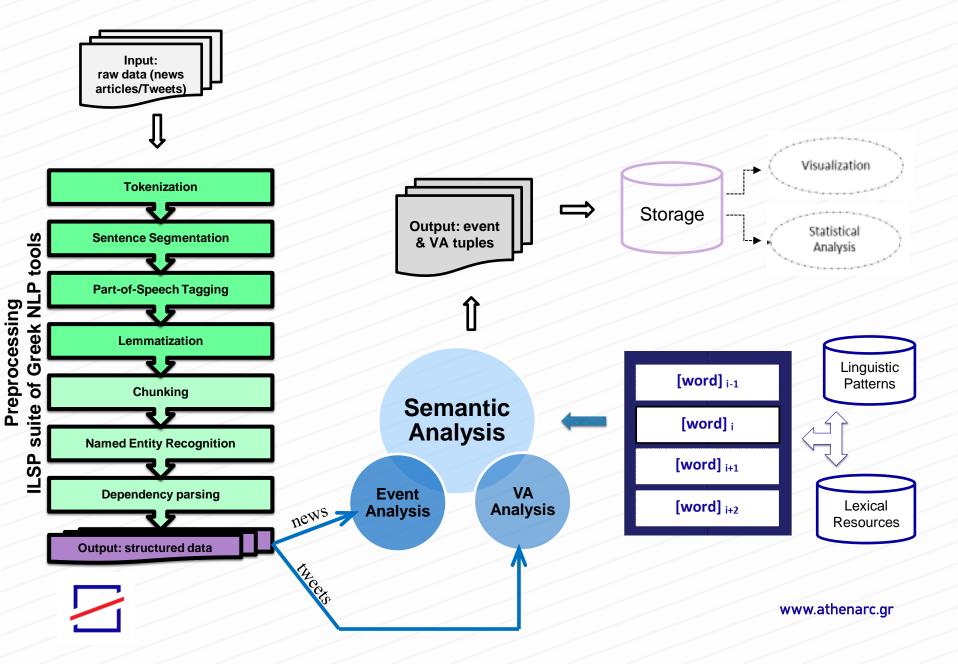
Methodology







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 Ελληνας μουσουλμάνος τραυματίστηκε με μαχαίρι από μέλη της Χ.Α.

Output

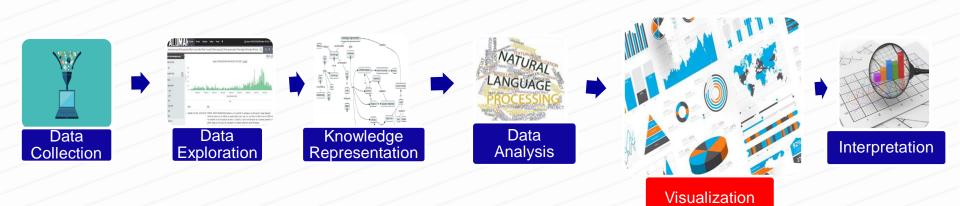
- <ΤΙΜΕ: Στις 28 Μαρτίου 2013, Location: στο κέντρο της Ξάνθης, Target: 30χρονος Ελληνας μουσουλμάνος Event: τραυματίστηκε με μαχαίρι Actor: από μέλη της Χ.Α.>
- ✓ Target_Age: 30
- ✓ Target_TG: 5
- ✓ Target_Nationality: Greek
- Event_Type: Assault against Life
- ✓ Actor_TG: Control Group
- Actor_Summary: Racist, extreme right organizations/groups

Other metadata: Event_Id, Medium, Article Title

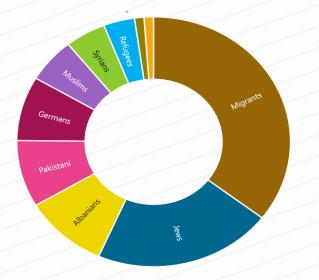
Semantic Analysis Output: Knowledge Databases

x	🕅 🗋 5 - Č 1															困 – ■ ×	
	FILE HOME	E HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW															Maria Pontiki 👻 🔍
1	Cut	Calibri	- 11	• A A =	= = *	>- ₽ ₩	rap Text	General	Ŧ	ŧ			e		∑ AutoSum ↓ Fill •	A A	
P	aste 💉 Format Painte	er BI	<u>u</u> + 🖽 + <u>4</u>	<mark>⊘ - A</mark> - E	≡≡≡ €	E 🔁 🖽 N	erge & Center	- 🚰 - %	€.0 .00 .00 →.0	Conditional Formatting		Cell Styles ∗	Insert *	Delete Forma	t 🅑 Clear 🔻	Sort & Find & Filter ▼ Select ▼	
_	Clipboard	Fai	Font	Gr.		Alignment		G Numb	er 🖓		Styles			Cells	Ed	iting	~
J	1154 💌 :	\times \checkmark	f _x Γαμωτ	τους αλβανα	ούς ρε φίλε												¥
	В	С	D	E	F	G	н	I								J	
1	TG_evidence	Target_id 💌	VAM_type	▼ VA_evide	nce 🔻 Year	 Month 	- Day -	user_id 🔹	text								
42	Αλβανός	TG2	VAM1B	νταής	2013	Oct	08	181544322	Β«Μίλησα	ινΒ» τα ευρι	ήματα στο	σπίτι τω	υν Αλβα	νών νταήδω	v http://t.co/WN	1WNvyh4rh	
47	Αλβανός	TG2	VAM1B	νταής	2013	Aug	25	32207236	ΑΛΒΑΝΟΣ	ΝΤΑΗΣ ΕΠΙΤ	ГЕӨНКЕ КА	I TPAYN	ΛΑΤΙΣΕ 5	4XPONH EA	ΛHNIΔA! http://t	.co/g8rXKR9dv)	(
56	Αλβανα	TG2	VAM1B	Αλβανα	2013	Aug	28	954419090	RT @plati	tudinous: Έξ	ω οι κωλοι	έλληνες	απ'τη Γ	ερμανία! @1	limosnik:"Όσοι έ	χουν αλλοδαπά	ά άτομα σπίτια τ
57	Αλβανα	TG2	VAM1B	Αλβανα	2013	Aug	28	476022997	RT @plati	tudinous: Έξ	ω οι κωλοι	έλληνες	απ'τη Γ	ερμανία! @1	limosnik:"Όσοι έ	χουν αλλοδαπά	ά άτομα σπίτια τ
59	Αλβανα	TG2	VAM1B	Αλβανα	2013	Aug	28	331590396	RT @plati	tudinous: Έξ	ω οι κωλοι	έλληνες	απ'τη Γ	ερμανία! @1	limosnik:"Όσοι έ	χουν αλλοδαπά	ά άτομα σπίτια τ
60	Αλβανα	TG2	VAM1B	Αλβανα	2013	Aug	28	1022706734	RT @Crt_I	vlts: ποσο μ	αλακας ει	σαι? @Τ	imosnil	k Δυστυχώς ό	σοι έχουν αλλοδ	απά άτομα σπί	τια τους αυτή τη
62	Αλβανα	TG2	VAM1B	Αλβανα	2013	Oct	29	773734368	Θες να αν	οιξουν τα σι	υνορα για	τα Αλβα	ανα "ελλ	ληνοπαιδα"	θεοδωρακη η γε	νικως για ολουσ	; εισαι εσυ μια
64	Αλβανα	TG2	VAM1B	Αλβανα	2013	Aug	28	902577704	RT @Crt_	vlts: ποσο μ	αλακας ει	σαι? @Τ	imosnil	k Δυστυχώς ό	σοι έχουν αλλοδ	απά άτομα σπί	τια τους αυτή τη
71	Αλβανός	TG2	VAM1B	ζώο	2014	May	07	17895667	Χτύπησε Ι	και βίασε ηλ	ικιωμένη	. Αλβαν	ό ζώο φ	υσικά (APXE	ΙΟ ΖΩΩΔΩΝ ΕΓΚ/	\HMATΩN) http	://t.co/e4TYtwI9
83	Αλβανός	TG2	VAM1B	χασάπης	2014	May	23	2433442367	RT @mars	ilnik: Απίστε	ευτα πράγ	ματαΈ	στησαν	μνημείο στα	Εξάρχεια για το	ν Αλβανό χασά	πη Ιλίρ Καρέλι! τ
87	Αλβανα	TG2	VAM1B	Αλβανα	2014	Oct	15	2791813033	@usay_gr	Τα Αλβανα	θα καψου	ν τα αρ	χς				
89	Αλβανός	TG2	VAM1B	χασάπης	2014	May	24	237719164	-				-	-			σαν μνημείο στα
	-	TG2	VAM1B	Αλβανά	2013	Aug	28	412082316									://t.co/U9E6wD7
92	Αλβανα	TG2	VAM1B	Αλβανα	2013	Aug	21	626561729	RT @POR	ΓAPORTA: Π	ου σαι ρε Ε	Βενιζέλο	? Σε και	νουν πλάκα	τα Αλβανα? Που	σαι ρε πολυτσα	αμπουκα καθεστ
	-	TG2	VAM1B	Αλβανα	2013	Aug	28	15071813	RT @plati	tudinous: Έξ	ω οι κωλοι	έλληνες	απ'τη Γ	ερμανία! @1	Γimosnik:"Όσοι έ	χουν αλλοδαπά	ά άτομα σπίτια τ
99	Αλβανία	TG2	VAM1B	γαμώ	2014	Oct	15	476022997	ΚΛΑΣΕ ΜΑ	Σ ΜΙΑ ΜΑΝ	TPA @Anc	ient_Kir	ng_FAM	ΙΩ ΤΗΝ ΑΛΒΑ	NIA @PETROSTA	ΑΜΑΤΑΚΟS Αλβ	ανός ποδοσφαις
11	Β Αλβανα	TG2	VAM1B	Αλβανα	2014	Oct	15	637222171	RT @agel	adam: Φαντι	ασου να σι	ηκωναν	οι Σερβ	οι καμια παρ	οομοια σημαια ε	θνικιστικου τυπ	του. 10 χρονια ατ
	-	TG2	VAM1B	γαμώ	2014	Oct	15	1327403546	_				-				του ΠΑΣ ο υπερι
14	3 Αλβανός	TG2	VAM1B	εθνίκια	2014	Oct	15	1581397489	Ρε τα εθν	κια τους Αλ	βανούς /	Αλλα το	αριστερ	ρό τουίτα τζι	αρ μούγκα Ο @	NikoAgo σχολί	ασε άραγε?
		TG2	VAM1B	γαμώ	2014	May	25	1169066552									λων, οταν σε γα
149	9 Αλβανός	TG2	VAM1B	γαμώ	2014	May	25	1499573102								ινοί τέτοια πολ	υλογία σε πιάνει
22	Β Αλβανός	TG2	VAM1B	φλώρος	2014	Nov	02	2488469162						πότε θα γίνι			
24	Β Αλβανός	TG2	VAM1B	κάθαρμα	_	Apr	08	17895667	Πιάστηκε	στην Κύπρο	το αλβανά	ό κάθαρ	μα Αθά	νι Αστρίτ ή Θ	θαθάνι Τίτι ή Θαί	θάν Τζόρτζιο ή /	Ασάνι… http://
25	Β αλβανικός	TG2	VAM1B	κάθαρμα	_	Nov	26	175540480	προς έλλι	νες δεσμοφ	ύλακες κα	ι επιτηρ	ητ'ες το	ου αλβανικοί	ύ καθάρματος, το	ου μακελάρη τι	ις πειραϊκής: β€?
25	9 αλβανικός	TG2	VAM1B	κάθαρμα	2014	Nov	26	106740388	προς έλλη	νες δεσμοφ	ύλακες κα	ι επιτηρ	ητ'ες το	ου αλβανικοί	ύ καθάρματος, το	ου μακελάρη τr	ις πειραϊκής: β€? 🛒
	• → <u>TG2</u>	Albanians	TG3_Romani	ans TG4	_Syrians T	G5_Islam_Mu	slims TG	6_Jews TG7_	Germans	🕂 :	4						Image: A start of the start

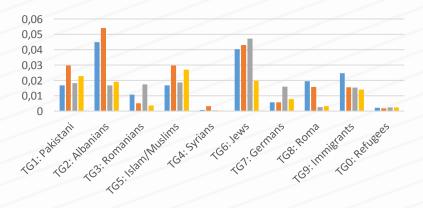
Methodology



Visualization: Examples



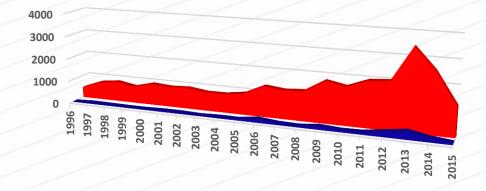
Targets of Physical Attacks: Naftemporiki 2000-2016



■ 2013 **■** 2014 **■** 2015 **■** 2016



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

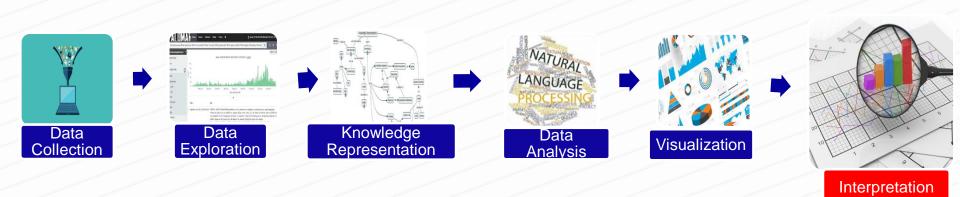


PA Timeline TGs vs CG



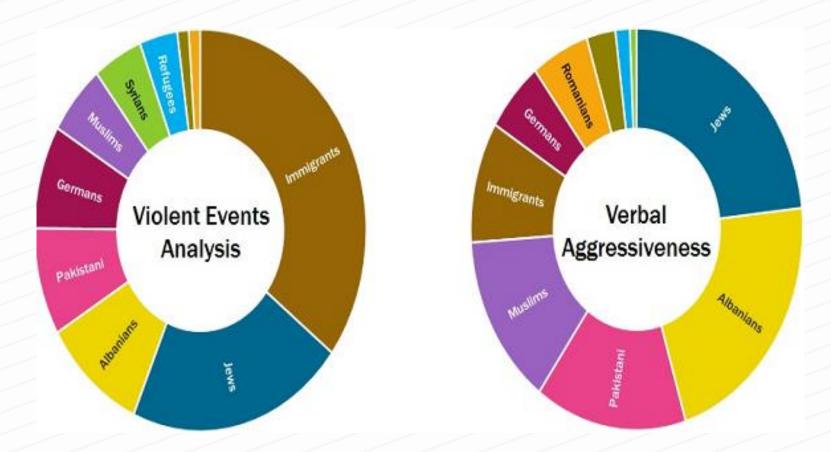
Word Cloud of unique aggressive terms for "Muslims/Islam"

Methodology



 \Box

Who are the main targets of physical and verbal aggression?



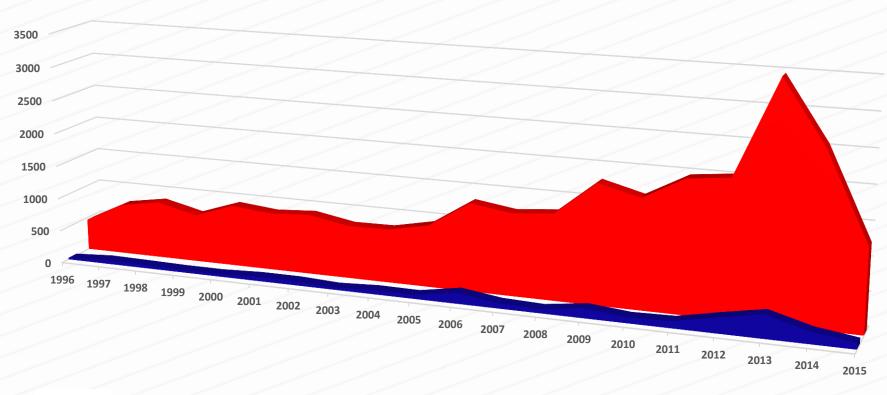
 \leq

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.

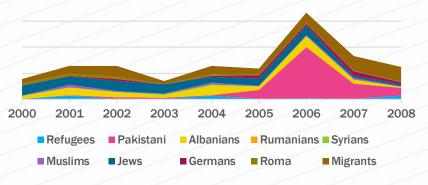


 \leq

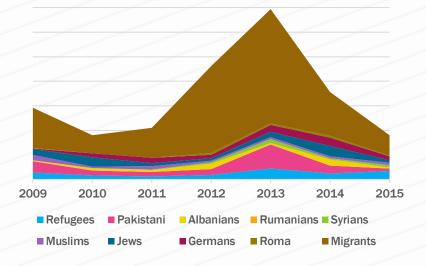
TGs CG

Physical Attacks Timeline

Have the xenophobic attitudes raised during the economic crisis?



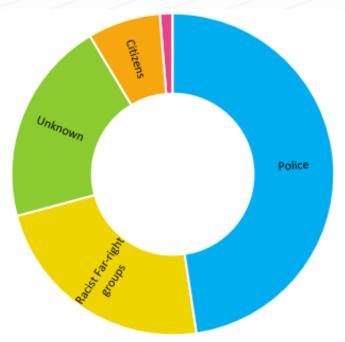
Mostly Attacked TGs before the economic crisis (Avgi)



Mostly Attacked TGs during the economic crisis (Avgi)

- 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

200 2001 2002 2003 2004 2005 2006 2001 2008 2009 2010 2012 2012 2013 2014 2015

 \leq



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

κατσαπλιάς υπερεθνικισμός χαζοχαρούμενος σκάρτος Αλβανα κομματόσκυλο Αλβανα υπερεθνικιστής εκδίωξη ΑΛΒΑΝΑΡΙΑ ιταμός απρέπεια φάρα ανούσιος αλβανιά εριστικός αρχιμαφιόζος επερασμένος αλβανιάρης στυγερός καραγκιοζάκος Φρασύδειλος αλβανιάρης στυγερός ΑλΒΑΝΑΡΙΟ αγέλη καραμπίνα ακατονόμαστος βλάσφημος κουλός Τραμπουκισμός κατΣΑΠΛΙΑΔΩΝ ΚΑΤΣΑΠΛΙΑΔΕΣ

Word Cloud of unique aggressive terms for "Albanians"



- 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.

Word Cloud of unique aggressive terms for "Pakistani"

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, 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.

- 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.

σπεκουλαδόρος βατραχοπέδιλο σπάταλος γερμανά ηθικολογία γερμαναράς εξωφρενικός κομπιναδόρος μεγαλοκαρχαρίαςξεροκεφαλιά μικροψυχία χαραμοφάης σανδάλικωλόχαρτο ξερόλας μπουρδέλο εκδικητικός χασούρα λαϊκιστήςχαρτογιακάς αντιπαθής κυνισμός κακοήθεια ανεύθυνος εγωισμός ανήθικος αδιαλλαξία ανέντιμος νεύρο συνωμοσιολόγος γλείψιμο ανικανότητα διασυρμός αλαζονεία σαδιστής χέστης άλτσα χαχόλος άγαρμπος ράιχ ναζής βόδι ραστος φερέφωνο τσίπα κλεφταράς αγένεια κοιλαράς νας απομίμηση ιερατείο ασέβεια αλαζόνας εθελόδουλος νταβάς ανοησία μειονέκτημα φύρα ημιμαθής βδελυρός έπαρση απατεωνιά άβουλος άκαρδος εμπαθής απροθυμία άφρών μιζαδόρος δύσκολος λαϊκισμός αμετροεπής ξεροκέφαλος υπεροψία αναιδής τεμπελχανάς κοροϊδία παράλογος ξενέρωτος σωβινισμός ανεδαφικός ανικανοποίητος σανδάλιααα παρούφα γερμανοτσολιάςγκάου ΓΕΡΜΑΝΕΛΛΗΝΕΣ βασανιστής παρούφα αργόστροφος Φρικαλέος

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.
- \rightarrow Vulnerability in the case of Jews and Germans.

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

 \rightarrow 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.

GRACIES JEW 26 GAI БЛАІ Ш ПАСИБІ/ Ι/ΕΥΧΑΡΙΣΤ \mathbf{O} RAZZI/ DANK JH G 0/**ХВАЛА**/ THANKYC U

Maria Pontiki

Computational Linguist Institute for Language & Speech Processing

(ATHENA RIC/ILSP), Greece





