



SIETARITALIA

Society for Intercultural Education  
Training and Research

**Identità multiculturali:**  
il senso e il valore dell'appartenenza nel mondo globale  
Milano, 5-6 maggio 2017

# “DIGITAL CRUMBS” AND SOCIAL NETWORKS TO ANALYZE AND ACT ON THE INTEGRATION PROCESSES OF MIGRANTS

M. Bogetti C. Bosco V. Patti



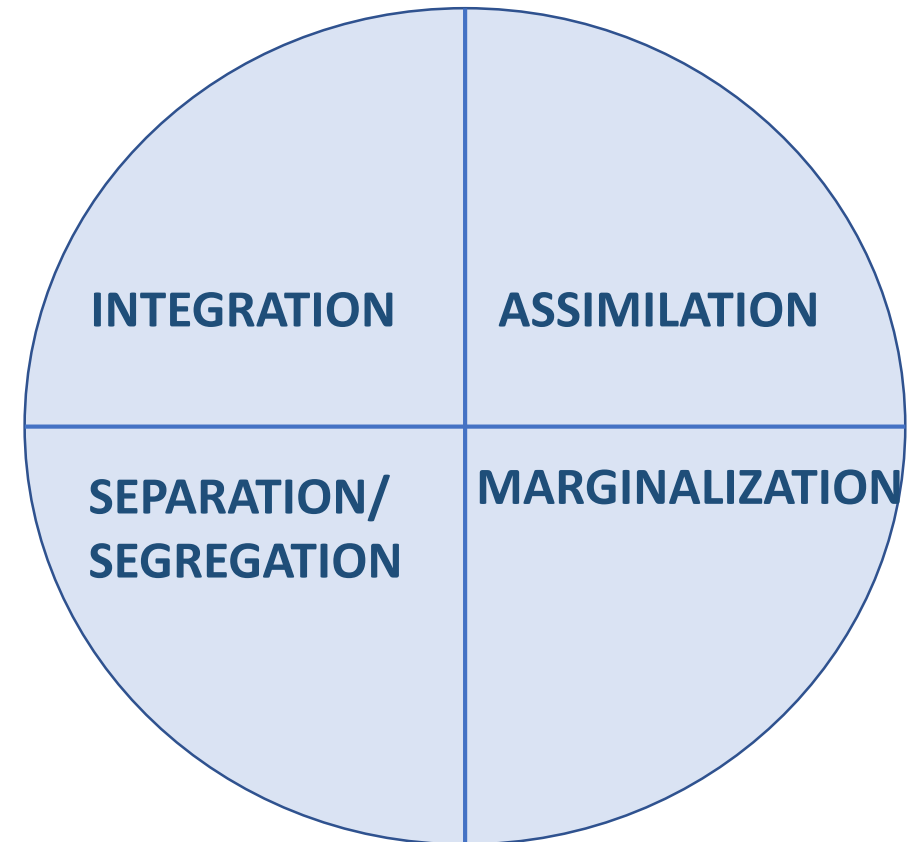
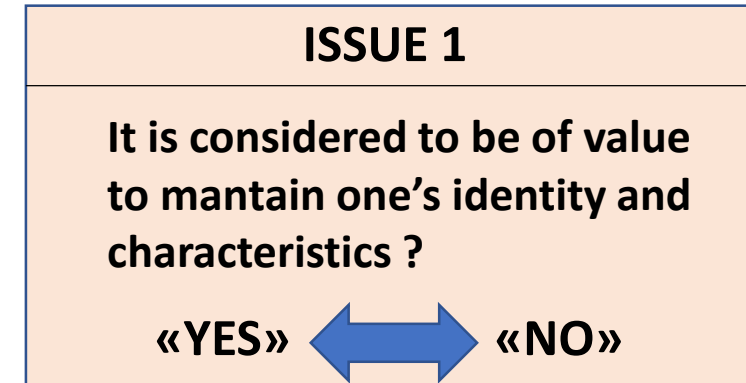
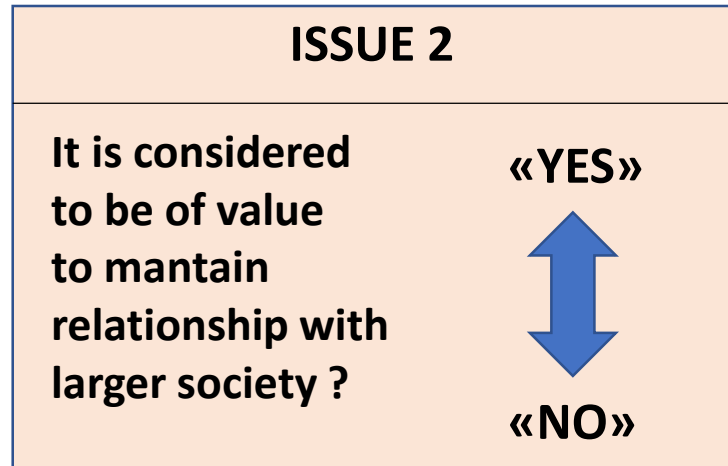
di.unito.it

# ACCULTURATION STRATEGIES

## Immigration, Acculturation, and Adaptation

John W. Berry

*Queen's University, Ontario, Canada*



# INTEGRATION OR ASSIMILATION?

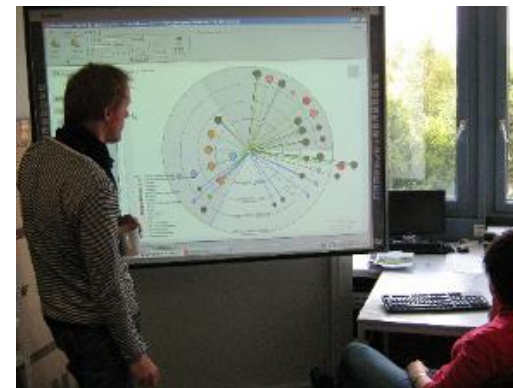
- In the past, the prevailing paradigm for solving problems related to immigration is based on the notion of *assimilation*.
- This paradigm assumes the existence of a valuable *mainstream culture* and proposes a strategy for immigrants integration which is based on the deletion of their culture in favour of the mainstream one.
- It has been applied in the USA for building a multi-ethnic society in the previous century
- But the paradigm of *assimilation* of immigrants to a mainstream culture is not a suitable strategy for Europe at the present time.
- It's assimilation desirable?
- *Immigration* may be seen as a source of *cultural enhancement*. Immigrants must respect rules and values of the hosting country, learn languages to improve the quality of their life, but also contribute to the development of hosting countries by their own culture and language.
- There is a culture to be considered as the mainstream one?
- Determining who is migrant and who is not is not a matter of biology or culture. Classifying different forms of migrants to be related to different positions in society, and also fossilize in prejudices the opinions about new comers, is only a matter of political choices.
- *Integration should be better configured as an interactive and bidirectional process.*

A young Moroccan woman  $t_1$ , last year in high school

**Red:** Spaniards  
**Green:** Moroccans in Morocco  
**Blue:** Moroccans in Spain  
**Gray:** The others  
**SIZE:** Closeness (1-5, the larger the closer)

A young Moroccan woman  $t_2$ , trying to find a job

**Red:** Spaniards  
**Green:** Moroccans in Morocco  
**Blue:** Moroccans in Spain  
**Gray:** The others  
**SIZE:** Closeness (1-5)



# BIG DATA: UN GLOBAL PULSE TAXONOMY\*

**1. Data Exhaust:** digital services create networked sensors of human behavior.

- Passively collected transactional data from people's use of digital services
  - Mobile phones: Call Detail Records (CDR) from mobile phones - i.e. log of calls for billing purpose with basic metadata
  - Purchases (in-store and online credit cards) and financial transfers
  - Web searches, and search engines trends and analytics --" Google flu"-style
  - Geolocation and all kind of individual / personal / local sensors on computers, phone, watch, bracelet, necklace, etc + motion/sound/photo/video capturing / processing, etc
- Operational metrics and other real-time data collected by UN agencies, NGOs and other aid organisations to monitor their projects and programmes: e.g. stock levels, school attendance, IDP & refugee registration, etc.

**2. Online Information** – web usage and content as a sensor of human intent, sentiments, perceptions, and want.

- Web content such as news media, news articles obituaries, e-commerce, job postings, bibliographic databases, online full-text libraries
- Social media interactions (e.g. blogs, Twitter) and social media bulk contents
- Web scrapping from open public online contents (web sites, Instagram, ..., text/photo/audio/video processing and pattern recognition, feature extraction, etc.)

**3. Physical Sensors** – focuses on remote sensing of changes in human activity.

- Remote sensing, weather data + astronomical + earth science data: land use, urban development and topographic changes, etc
- Scanned or image/audio/video recording/transmission/processing + new personal sensors (watch, bracelets, phones, etc.) + home sensors, environmental sensors for pollution, etc.

**4. Citizen Reporting or Crowd-sourced Data** – Information actively produced or submitted by citizens through mobile phone-based surveys, hotlines, user-generated maps, etc; While not passively produced, this is a key information source for verification and feedback

**5. [UNPD: Simulated probabilistic data and agent-based simulations]** – including probabilistic estimations and/or projections with thousands of trajectories, parameters, and multidimensional data arrays (e.g., indicator, location, time, age, sex, etc.)

(\*) Big data for development: challenges & opportunities, p.16 <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobaIPulseJune2012.pdf>



# IMMIGRATION AND HATE SPEECH

- It is instead a matter of fact that [immigration stimulates](#) the generation and diffusion of [hate speech](#) and [hate crimes](#), and several initiatives are promoted for preventing hate speech all around the world.
- Can Artificial Intelligence give a contribute in addressing this challenging task? In the perspective of the [AI for Migrant Integration workshop](#) it can.
- AI can help in increasing the awareness and empathy of receiving populations while avoiding polarization against immigrants.

# OUR PROJECTS

- In the last few years, we developed Sentiment Analysis tools and resources in a multi-lingual perspective and recently started addressing hate speech by AI techniques.
- We recently started a project (*Hate Speech and Social Media*) devoted to the improvement of awareness of hate speech used in social media in high school students.
- We are working in the *IHatePrejudice* project (*Immigrants, Hate and Prejudice in Social Media*).

# IHATEPREJUDICE: GOAL

- Language is crucial:
  - to determine boundaries among groups and minorities
  - to confer immigrants specific social roles, rights and opportunities
  - to describe migrants as others, as outsiders of our society, often as dangerous and unwanted
- But language is also exploited by immigrants to react to the challenges of integration
- Starting from the relevance of language in the immigration phenomena and the huge variety of linguistic data available as digital traces,
- the main objective of the project is the development of computational tools for **accessing and analyzing** small and big **data** from different social media to **acquire** a deeper **knowledge about anti-immigration attitudes**, related to **prejudice** and **hate speech**



# IHATEPREJUDICE: OUTCOMES

The final aim is twofold:

- **transforming analyzed data in actionable knowledge** for the territorial entities raising awareness about discrimination and taking action against it
- **democratizing data**, by increasing accessibility for citizens and thus improving the quality of life through evidence-based decision making. This is a prerequisite **to prevent and reduce prejudice** and resistance to the integration of new comers.

# IHATEPREJUDICE: METHODOLOGY

- The way we are addressing the challenges related to immigration is based on **identifying different patterns of hate and prejudice** in on line expressions.
- The involved phenomena will be studied in a novel and **bidirectional perspective**, e.g. both the movements and communication dynamics of immigrants, and the reactions of local community, media and single citizens will be considered
- We **integrate Natural Language Processing, data and network analytics**, together with **sociological** and **visualization** approaches based on maps showing the diffusion of phenomena as related to geographical locations.
- Data aggregation and visualization will be applied for **sharing** such knowledge **with operators working in the local area** (our case study), but also with citizens for increasing awareness.

# IHATEPREJUDICE: LINGUISTIC ANALYSIS

- Considering the complexity of communication dynamics related to immigration, the development of **novel forms of computational analysis** must sustain the traditional ones, being also more adequate for approaching **big data** like digital traces where causes and consequences of hate and prejudice are diluted
- As we did in the development of multi-lingual resources for Sentiment Analysis about **political debates**, we will
  - collect text samples from different media which include hate speech expressions against immigrants
  - analyze them for detecting relevant expressions
  - annotate a selection according to a scheme that makes explicit **hate expressions, targets and polarization**.

# IHATEPREJUDICE: SENTIMENT ANALYSIS

- The framework integrates techniques for collecting social data and analysing them by data mining and **aspect-based sentiment analysis** and **stance detection**, by applying highly developed natural language processing to deal with the variety of fine-grained aspects involved in hate speech and hostility towards migrants, like e.g. **irony and figurative uses of language**.
- Knowledge is extracted from large linguistic datasets where relevant information is manually annotated and then detected, e.g. **stance** against or pro, presence of **irony** or literal meaning, presence of **hate speech**, presence of **offensive expressions**.

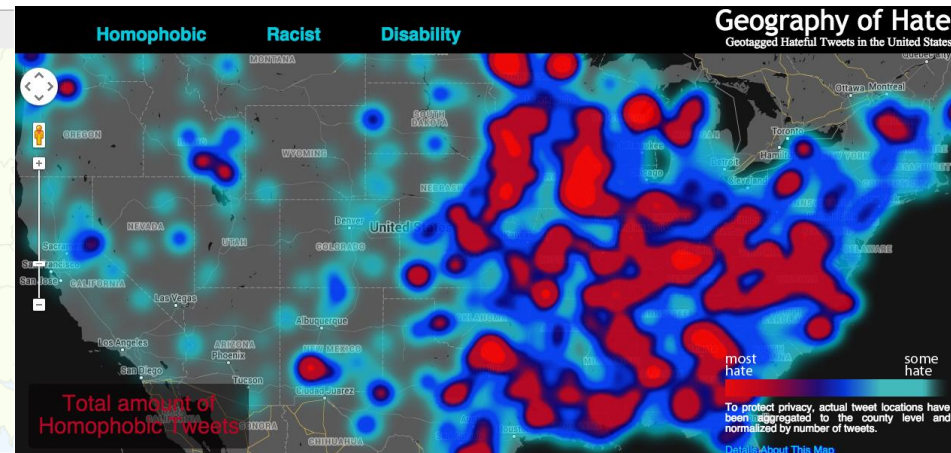
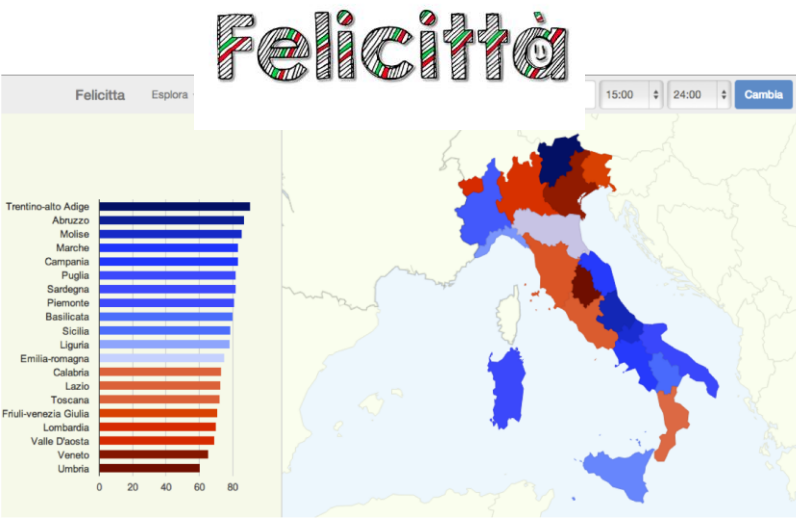
# IHATEPREJUDICE: BUILDING DATASETS for SENTIMENT ANALYSIS

testo	stance	ironia	hate speech	offensivo
Nuova gaffe della Boldrini: «Mi sento più italiana quando sto con gli immigrati» <a href="https://t.co/sl77XBVHhP">https://t.co/sl77XBVHhP</a>	NEUTRALE	IRONICO	HATE SPEECH: NO	PER NULLA OFFENSIVO
ROMA – ORRORE ALLA STAZIONE: DUE CARABINIERI ACCERCHIATI DA 15 IMMIGRATI INFEROCITI - Notiziario Online	CONTRO	NON IRONICO	HATE SPEECH: SI	POCO OFFENSIVO
Milano, soldati aggrediti da immigrati: un ferito e un arresto <a href="https://t.co/JbrJk0IbEN">https://t.co/JbrJk0IbEN</a> #thexeon #milanocronaca	NEUTRALE	NON IRONICO	HATE SPEECH: SI	POCO OFFENSIVO
Battipaglia, immigrato piscia in strada nei pressi di una scuola:... <a href="https://t.co/jVKXOVupfK">https://t.co/jVKXOVupfK</a>	CONTRO	NON IRONICO	HATE SPEECH: SI	MOLTO OFFENSIVO
Un giudice blocca in tutti gli Stati Uniti lo stop di Trump agli immigrati islamici <a href="https://t.co/y5Mys5Djrl">https://t.co/y5Mys5Djrl</a>	NEUTRALE	NON IRONICO	HATE SPEECH: NO	PER NULLA OFFENSIVO
Altro che pagarci le pensioni Immigrati sono solo un costo: Un luogo comune molto diffuso in Italia racconta ...	CONTRO	NON IRONICO	HATE SPEECH: SI	MOLTO OFFENSIVO
Italia: immigrati 92.300, emigrati 107.529. Chi è che invade?	CONTRO	IRONICO	HATE SPEECH: SI	ABBASTANZA OFFENSIVO
@dawac72 immigrati a volontà e questi che ha fatto Renzi città invase e disagi ovunque	CONTRO	NON IRONICO	HATE SPEECH: SI	ABBASTANZA OFFENSIVO
MILANO: RIVOLTA IMMIGRATI, ZONA DEVASTATA E CASSONETTI ROVESCIATI   VoxNews <a href="https://t.co/ndzFs1VHOC">https://t.co/ndzFs1VHOC</a>	CONTRO	NON IRONICO	HATE SPEECH: SI	PER NULLA OFFENSIVO
@VincenzoDeLuca È giunto il momento di fare un po di pulizia seria, tra espellere immigrati irregolari e arrestare i nostri malviventi!!!	NEUTRALE	NON IRONICO	HATE SPEECH: NO	PER NULLA OFFENSIVO
Che futuro possiamo garantire a tutti questi immigrati? Che futuro possiamo garantire a noi e ai nostri figli? SIETE DEI CRIMINALI...	NEUTRALE	NON IRONICO	HATE SPEECH: NO	PER NULLA OFFENSIVO
vicino a #Spagna ci sono Marocco/Tunisia, #Italia vicino a #Libia dove parte flusso migratorio. #Salvini la #geografia salla!	PRO	IRONICO	HATE SPEECH: NO	PER NULLA OFFENSIVO
@Roccia9 Leggiti questo articolo per vedere quanto sono amati gli immigrati italiani. <a href="https://t.co/NYaJEVJTj9">https://t.co/NYaJEVJTj9</a> @Italy_Mauri	NEUTRALE	NON IRONICO	HATE SPEECH: NO	PER NULLA OFFENSIVO
Immigrato stupra la figlia di 7 anni   Tutti i Crimini degli Immigrati <a href="https://t.co/b6glbN3FNc">https://t.co/b6glbN3FNc</a>	CONTRO	NON IRONICO	HATE SPEECH: SI	MOLTO OFFENSIVO
Isis a Milano: è la prova delle frottole raccontate da Renzi sugli immigrati ED E SOLO L INIZIO VEDRAI CHE FUOCHI EHH	CONTRO	IRONICO	HATE SPEECH: SI	ABBASTANZA OFFENSIVO
Le "social card" di #Renzi? Sette su dieci agli immigrati <a href="https://t.co/Lb9b08LHvY">https://t.co/Lb9b08LHvY</a> il 67% dei contributi per l'affitto va a non italiani	CONTRO	NON IRONICO	HATE SPEECH: SI	PER NULLA OFFENSIVO
@ACeschia Assolutamente NO..Renzi e il PD stanno distruggendo l'italia..con gli immigrati clandestini.. Berlusconi ha fatto il suo tempo!	CONTRO	NON IRONICO	HATE SPEECH: SI	POCO OFFENSIVO
#Gentiloni ce la farà a bloccare il flusso immigrati in Libia?in Spagna Zapatero c è riuscito a bloccare tutto muri e guardie!#speriamo	CONTRO	IRONICO	HATE SPEECH: SI	ABBASTANZA OFFENSIVO
Sotto la neve alla fermata del bus: studentessa salva un immigrato <a href="https://t.co/2Ud9ICD5TN">https://t.co/2Ud9ICD5TN</a>	PRO	NON IRONICO	HATE SPEECH: NO	PER NULLA OFFENSIVO
Capotreno a giudizio per aver fatto scendere immigrato senza biglietto <a href="https://t.co/Uar6aLRApX">https://t.co/Uar6aLRApX</a> <a href="https://t.co/qae8uDxs61">https://t.co/qae8uDxs61</a>	CONTRO	NON IRONICO	HATE SPEECH: SI	ABBASTANZA OFFENSIVO



# IHATEPREJUDICE: VISUALIZATION AND MAPS

- **Analyzed data** will be aggregated under several semantic perspectives including geographical and temporal components.
- Geo-social data will be exploited to derive a **geography of hate and prejudices about immigrants** and to study immigrants mobility.
- A **web-based platform** will be designed for supporting interactive access to data and visualize them in hate maps.





# CONCLUSION AND FUTURE WORK

- We are developing an AI multi-disciplinary methodology for addressing the challenges of immigration by improving citizens' awareness.
- Our aim is at transforming analyzed data in actionable knowledge and democratizing data about immigration to prevent and reduce hate and prejudice.
- After the trial on the local context, the methodology will be applied in the next future to a larger area (namely the European one).

# THE MULTI-DISCIPLINARY TEAM



di.unito.it



Cristina BOSCO



Viviana PATTI



Delia Irazu  
HERNANDEZ  
FARIAS



Manuela  
SANGUINETTI



Rossano  
SCHIFANELLA



Giancarlo RUFFO



Marcello  
BOGETTI



Massimo FOLLIS



Marco STRANISCI



Michelangelo  
CONOSCENTI

# Hate Speech Monitoring@UniTo

All the information are (or will be) available at

<http://hatespeech.di.unito.it>

# Our sponsors



Thank you!!!

[marcello.bogetti@unito.it](mailto:marcello.bogetti@unito.it)

[bosco@di.unito.it](mailto:bosco@di.unito.it)

[patti@di.unito.it](mailto:patti@di.unito.it)