

# Vero e Falso nell'era dei Big Data, Social Media e Apps

Paola Monachesi (Utrecht University)  
P.Monachesi@uu.nl

# Programma

- ~~Da dove vengono i dati?~~
- ~~Quali dati produciamo quotidianamente?~~
- ~~Chi utilizza i Big Data?~~
- ~~Vero — Falso e i Big Data~~
- ~~Vero — Falso e i Social Media~~
- ~~Case studio: smart cities~~
- Big Data e le scienze sociali/umane
- Conclusioni: sfide e opportunità'

# La verita' dei dati

- Nuovo empirismo
- i dati sono obiettivi e neutrali
- nuove tecniche di analisi rivelano la verita' intrinseca dei dati
- dati sono liberati dalla teoria
- Tensione tra mondo del:
  - business (nuove opportunita')
  - accademico (nuove conoscenze)

# Empirismo

- Big Data permettono di:
  - ▣ analizzare interi domini
  - ▣ evitare teorie, modelli, ipotesi
  - ▣ lasciar parlare i dati senza interferenza umana
  - ▣ evidenziare modelli e relazioni tra i dati inerentemente significativi e veri
  - ▣ significato dei dati può essere identificato da chiunque riesca a decodificare visualizzazioni e risultati statistici
- E' vero???

# Nuovo modo di concepire la scienza?

**Table 1.** Four paradigms of science.

Paradigm	Nature	Form	When
First	Experimental science	Empiricism; describing natural phenomena	pre-Renaissance
Second	Theoretical science	Modelling and generalization	pre-computers
Third	Computational science	Simulation of complex phenomena	pre-Big Data
Fourth	Exploratory science	Data-intensive; statistical exploration and data mining	Now

Compiled from Hey et al. (2009).

# Possibili problemi

- Piattaforma e tecnologia determinano come i dati vengono raccolti
- Problema di sampling
- Algoritmi basati su teorie e testati su dati
- Correlazioni possono essere casuali
- I dati sono sempre interpretati e se sono staccati dal contesto possono diventare vuoti di significato
- Apparente semplificazione di realta' complessa

# Scienza dei dati

- Ipotesi originano dai dati invece che dalla teoria
- Decisioni prese su:
  - ▣ quali dati raccogliere
  - ▣ quali tecniche producono interpretazione appropriata
- Attenzione solo a relazioni che sembrano significative
- Approccio deduttivo viene poi applicato
- Possibilita' di combinare vari domini e discipline
- Nuovo modo di produrre teorie, nuova metodologia scientifica

# Scienze sociali vs. Scienze umane

- Scienza dei dati come sviluppo naturale per le scienze esatte
- Scienze sociali?
- Scienze umane?
- Diverse filosofie e diverse metodologie
- SS => fatti, informazione quantificabile e misurabile, statistica, teorie, modelli e simulazioni
- SU => ideologia, opinioni, ...



# Big Data e SS/SU

- Digital humanities
  - ▣ Rigore metodologico
  - ▣ Obiettività
  - ▣ Nuove tecniche complementano le vecchie e permettono nuovi studi e nuove domande
  - ▣ Perdita in profondità di analisi
  - ▣ Pattern descrittivi
- Ma la letteratura è un insieme di dati? E l'arte?

# Big Data e SS/SU

- Computational Social Sciences
  - ▣ Pattern come inizio dell'analisi
  - ▣ Necessita' di spiegare i dati
  - ▣ Societa' non riducibile a fatti
  - ▣ Impossibile spiegare eventi come guerre, razzismo, cultura, politica....

# Domande

- Quali sono i dati di riferimento della mia disciplina?
- Quali sono le metodologie usate per l'analisi?
- Che impatto possono avere per la mia disciplina i Big Data/Social Media/Apps?
- Questi dati sono inerentemente veri nel mio contesto?
- Risposta a queste domande:
  - ▣ All'interno della propria disciplina
  - ▣ All'interno di SS/SU

# I Big Data e i SM nella mia ricerca

- Migranti della conoscenza
- Uso social media da parte degli anziani (del futuro)

# Migranti transnazionali

- Migranti transnazionali creano dei ponti nel collegare spazi urbani locali a quelli globali
- Focus su designers e pianificatori urbani che usano tecnologia (i.e. app developers, programmatori), cultura (i.e. artisti) e sostenibilità (ingegneri).
- Collaborazione con CUHK e Institute of Future Cities
- UU: Institutions seed funding

# Domande di ricerca

- *I migranti della conoscenza come trasformano lo spazio urbano in uno spazio digitale?*
- Focus sulla 'attività' nei social media dei migranti della conoscenza
- **Profilo**
- **Gruppi**
- **Locale vs. globale** (reti, progetti,...)
- **Dinamica** (evoluzione)

# Metodologia

- Estrazione dati twitter
  - ▣ Twitter search API
  - ▣ Dati salvati in un database MySQL database
- Analisi linguistica
- Social Network Analysis
- Visualizzazione dei risultati in una carta

# Analisi linguistica: metodologia

- Analisi della comunicazione nei Social Media
  - ▣ Stile
    - Connettivi
    - Pronomi
    - Parole funzionali
    - N-grams
    - Punteggiatura
    - POS tagging
  - ▣ Contenuti
    - Topic models





# GRAGE

*Green and Grey in Europe: elderly living in urban areas*

# GRAGE: Partecipanti



Università degli Studi di Macerata



Ulster University



Waag Society



University of Oulu



University of Utrecht



Institute for Environmental  
and spatial development



Taras University



City of Oulu Civil and  
Environmental services



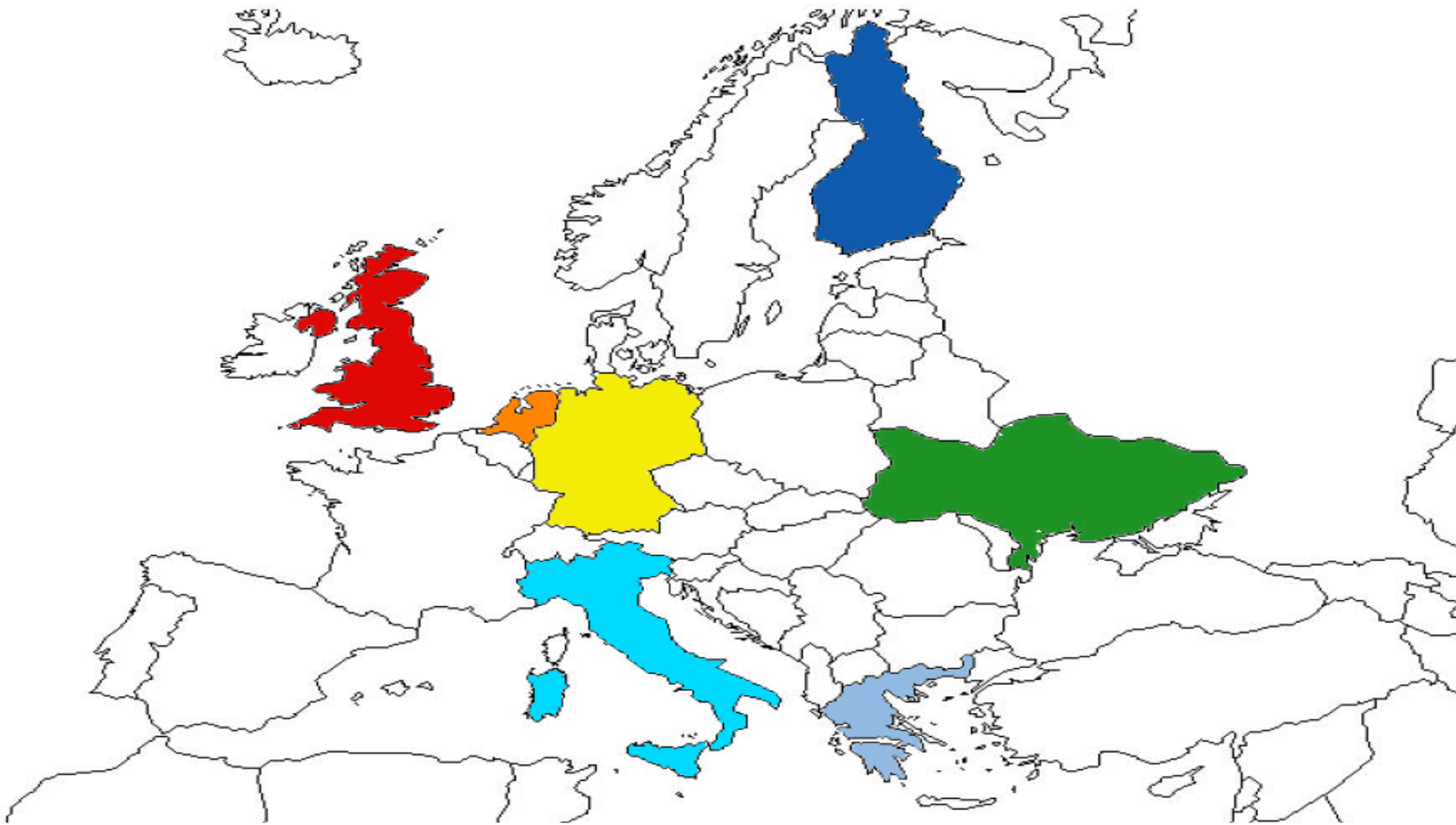
European Quality Institute



Militos



**GRAGE: Paesi**



# ICT e LT per anzianita' attiva nelle citta'

- Analisi del comportamento degli individui con eta' 55+ per stimolare anzianita' attiva
- Collezione dati relativi a comportamento attraverso dispositivi personali
- Individuazione comunita' e comportamento
- Uso di ICT per aumentare mobilita' nelle citta'
- ICT per supportare anzianita' attiva

# Analisi degli anziani dei SM

- Estrazione dati twitter
  - ▣ Twitter search API
  - ▣ Dati salvati in un database MySQL database
- Analisi linguistica
- Social Network Analysis
- Visualizzazione dei risultati in una carta

# Contatti

- Per informazioni e domande:
- Paola Monachesi
- Trans 10
- 3512 JK Utrecht
- Olanda
- Email: [P.Monachesi@uu.nl](mailto:P.Monachesi@uu.nl)
- Web: <http://www.uu.nl/staff/Pmonachesi/>
- LinkedIn:  
<http://www.linkedin.com/in/paolamonachesi/>