

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
- Caso studio: smart cities
- Big Data e le scienze sociali/umane
- Conclusioni: sfide e opportunità'

Vero e Falso

- La scienza ha come primo obiettivo quello di identificare la verità attraverso l'analisi dei fatti
- Verità non sembra avere più importanza in questo momento storico: post-truth
- Verità viene verificata attraverso l'osservazione: esperimenti nelle scienze

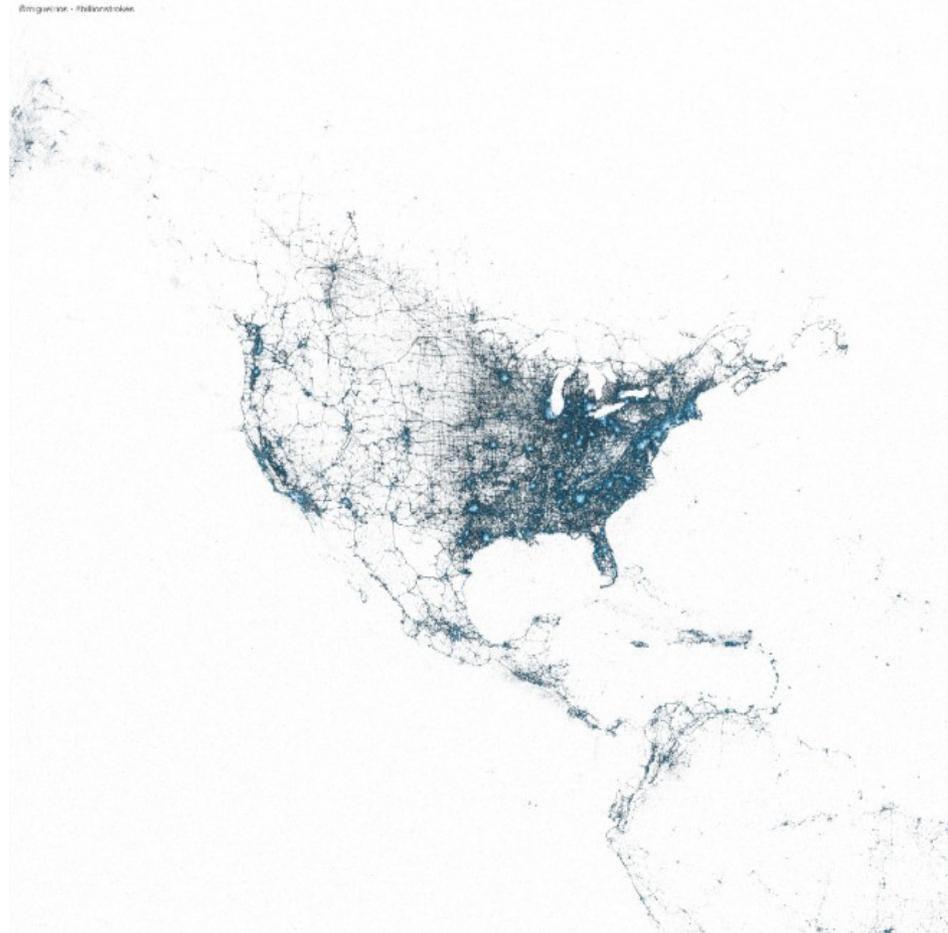
Big Data e verita'

- Big Data insieme alle nuove tecniche e algoritmi permettono di scoprire patterns e trends.
- In genere i dati sono prodotti senza una specifica ipotesi ma sono l'effetto collaterale di qualche altra attività'.
- Le ipotesi o verita' quindi emergono dai dati.

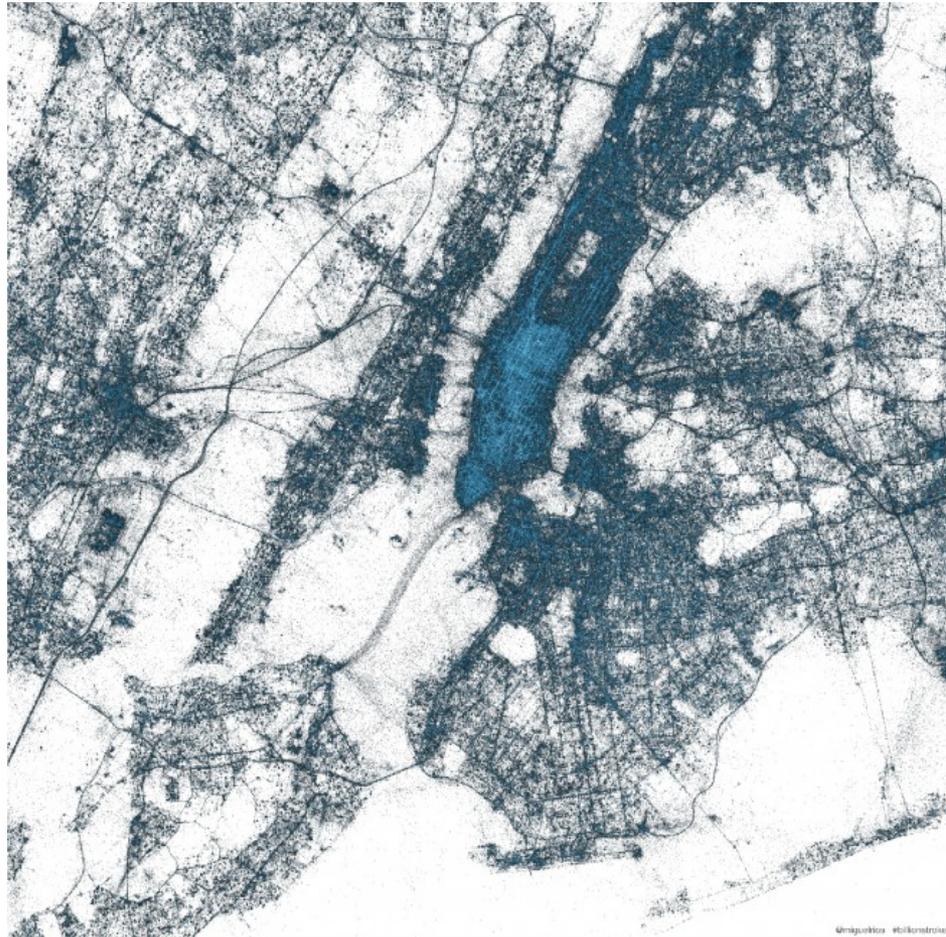
Empirismo e Big Data

- 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 puo' essere identificato da chiunque riesca a decodificare visualizzazioni e risultati statistici
- E' vero???

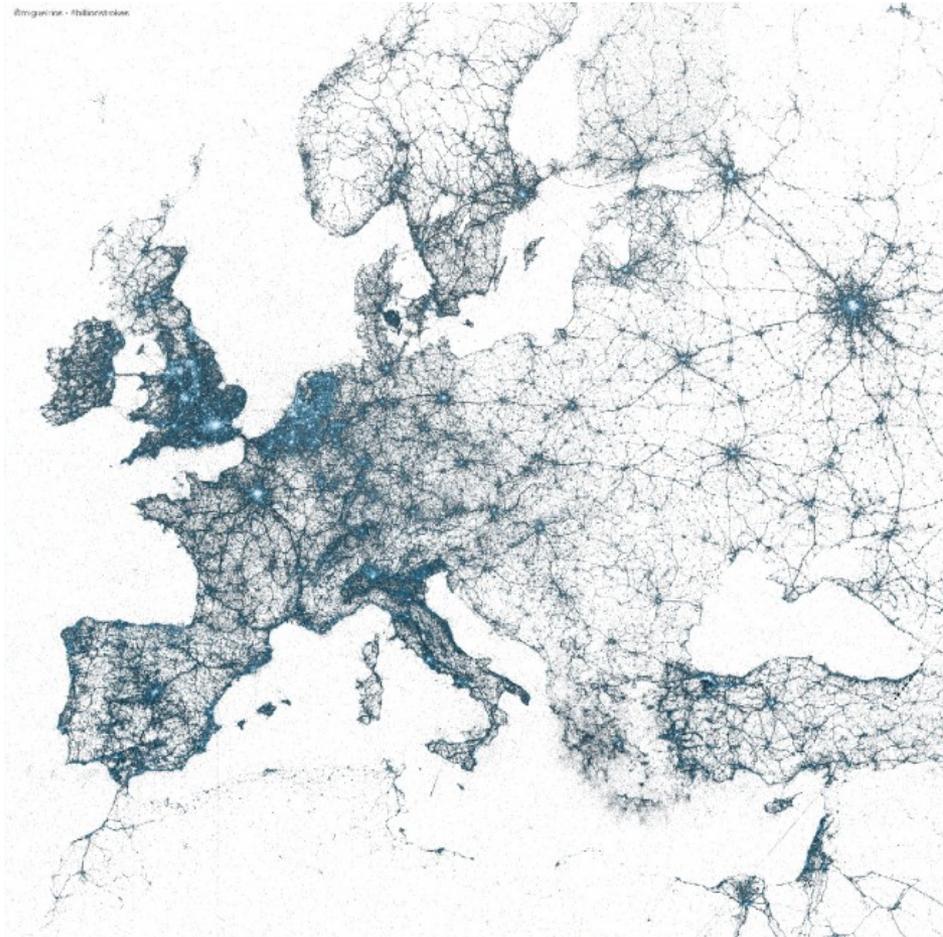
Visualizzazione Tweets: USA



Visualizzazione tweets: NY



Visualizzazione Tweets: EU



Esempio: Piperino

- Piperino: Con le peggiori intenzioni

Le vostre recensioni (319)

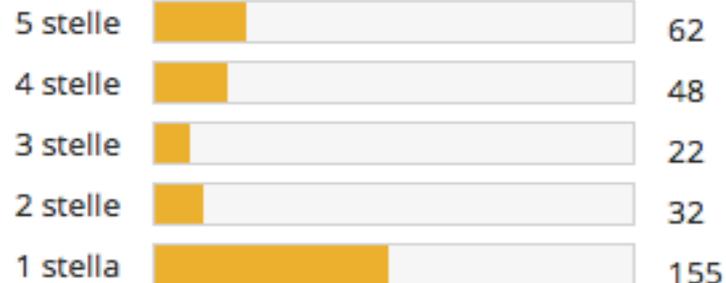
La voce della critica

Note legali

Media clienti: 2,47 di 5



319 recensioni



Bollier 2010

- As a large mass of raw information, Big Data is not self-explanatory. And yet the specific methodologies for interpreting the data are open to all sorts of philosophical debate. Can the data represent an ‘objective truth’ or is any interpretation necessarily biased by some subjective filter or the way that data is ‘cleaned?’. (2010, p. 13)

I dati parlano da soli

“Who knows why people do what they do? The point is that they do it and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves”

Anderson, C. (2008) 'The end of theory, will the data deluge makes the scientific method obsolete?', Edge, Available at: http://www.edge.org/3rd_culture/anderson08/anderson08_index.html

Scienza senza ipotesi

“Correlation is enough. . . .We can analyze the data without hypotheses about what it might show.

....

Correlation supersedes causation, and science can advance even without coherent models”

Anderson, C. (2008) ‘The end of theory, will the data deluge makes the scientific method obsolete?’, Edge, Available at: http://www.edge.org/3rd_culture/anderson08/anderson08_index.html

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)

Esempio analisi sentimenti Twitter

- <https://twitter.com/search-advanced>
- Ricerca: Trump positivi

Esempio Trendsmap



□ <https://www.trendsmap.com/>

Esempio: Infographics

- <http://www.webanalyticsworld.net/2010/10/how-executives-use-social-media.html>

Domande

- E' necessario capire il mondo per scegliere/
vendere/...?
- E' necessario conoscere le cause dietro determinati
patterns?
- E' sufficiente essere in grado di prevedere?
 - ▣ algoritmi che suggeriscono acquisti

Vero e Falso

- Verita' non sempre basata sui fatti, ma anche sull'esperienza
- Verita' legata ai fatti (dati tecnici) vs. Verita' legata ai sentimenti (importanza dell'arte e della letteratura)
- Analisi tecnica dei dati, ma anche importanza delle storie che sono ad essi legate => smart city vs. Utopic urbanism

Social Media e verita'

- 2011 Eli Pariser: filter bubble => essere esposti solo all'informazione legata al nostro profilo (ricerca personalizzata in Google)
- Personalizzazione e' diventata ancora piu' estrema ora che l'informazione e' nelle mani di un gruppo limitato di piattaforme SM
- Personalizzazione implicita nel concetto di Social Networks che rendono visibile quello che l'utente e i suoi amici vogliono vedere

Esempi

Twitter e la politica

Twitter e le relazioni

Twitter e la politica

- Twitter puo' essere usato per scopi politici
- Impatto comunicativo
- Polarizzazione e segregazione
- Rischio:
 - ▣ mancanza di diversita'
 - ▣ segregazione all'interno della propria comunita'
- Opinioni possono diventare ancora piu' estreme

Conover et al. (2011) “ *Political polarization on Twitter*”. AAAI conference on Weblogs and SM.

Gruppi e comunicazione

- Analisi delle reti della comunicazione politica
- Che cosa si analizza?
- Quali strumenti sono necessari?
- Che risultati ci possiamo aspettare?
- Importanti:
 - ▣ Retweets
 - ▣ Mentions
 - ▣ Hashtags (#)

Metodologia

- Rete degli utenti e tweets
- Analisi in raggruppamenti (clusters) basata su retweet e mentions
- Classificazione manuale degli utenti per analizzare la rete
- Interpretazione della struttura della comunità'

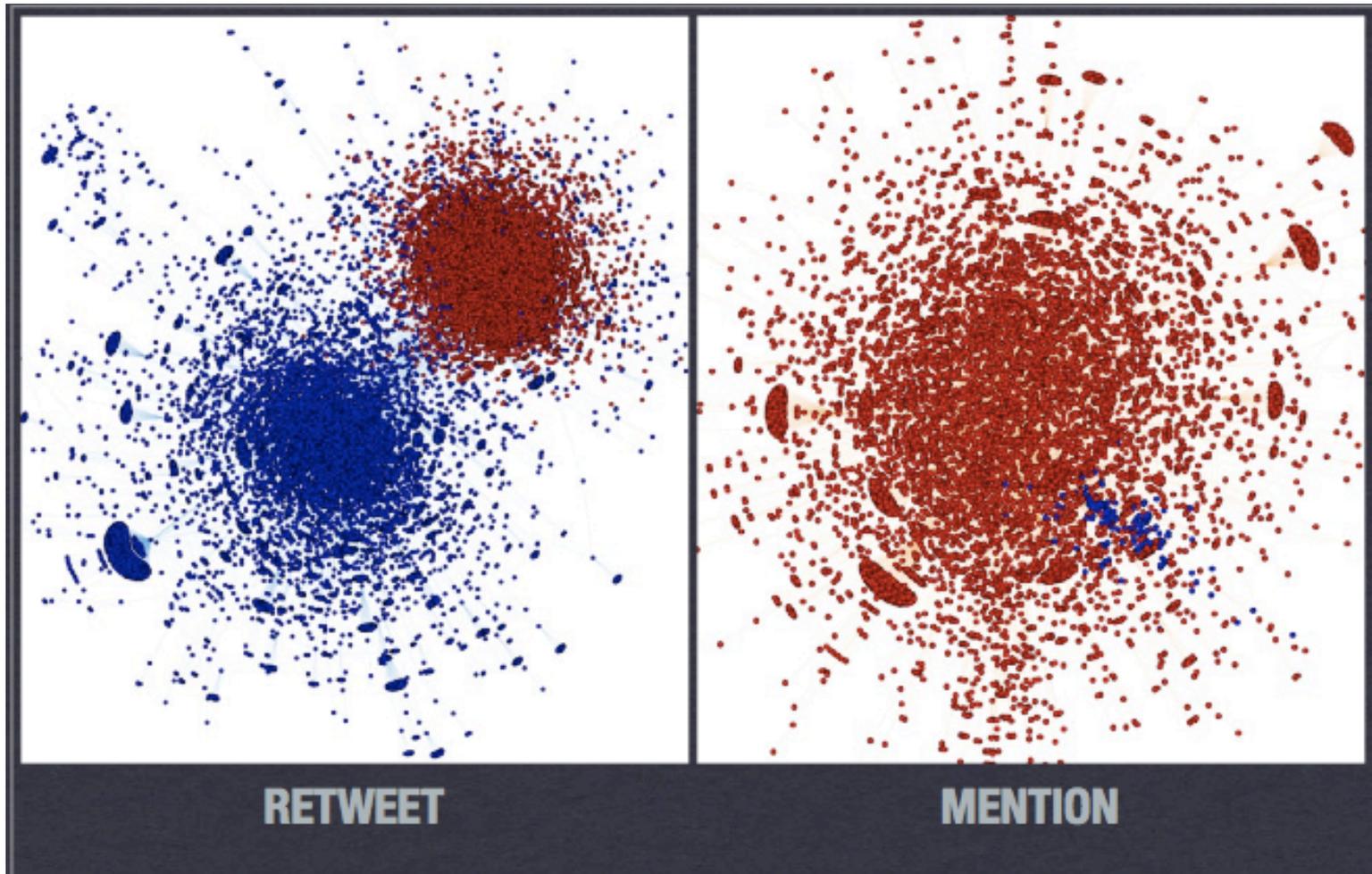
Analisi

- Impatto dei Social Media sulla comunicazione politica
- Retweets: segregazione
- Mention: interazione tra diverse opinioni alimentata da utenti con fini politici utilizzando #
- Uso di #: danno la possibilità' ad utenti di scegliere contenuti che altrimenti non considererebbero

Risultati

- Retweet e mentions sono usati in modo diverso in Twitter.
- Loro influenza sulla comunicazione politica e su come si diffonde l'informazione nella rete.
- Non casuale ma conseguenza dell'uso di # specifici utilizzati da utenti per influenzare la comunicazione.

Multi-mode communication



Dati

- Dati raccolti durante 6 settimane in occasione delle elezioni del congresso US di mezzo termine nel 2010
- 355 milioni di tweets
- Necessario selezionare
- Come?

Identificazione di contenuti politici

- Identificazione tweets con almeno un # politico
- Identificare la co-occorrenza dei tag
- Uso di seed tags (es. #p2, #tcot)
- Identificazione di # che sono presenti contemporaneamente in almeno un tweet

Risultato finale

- Identificazione di 66 # unici
- In totale 252300 tweets

Esempi

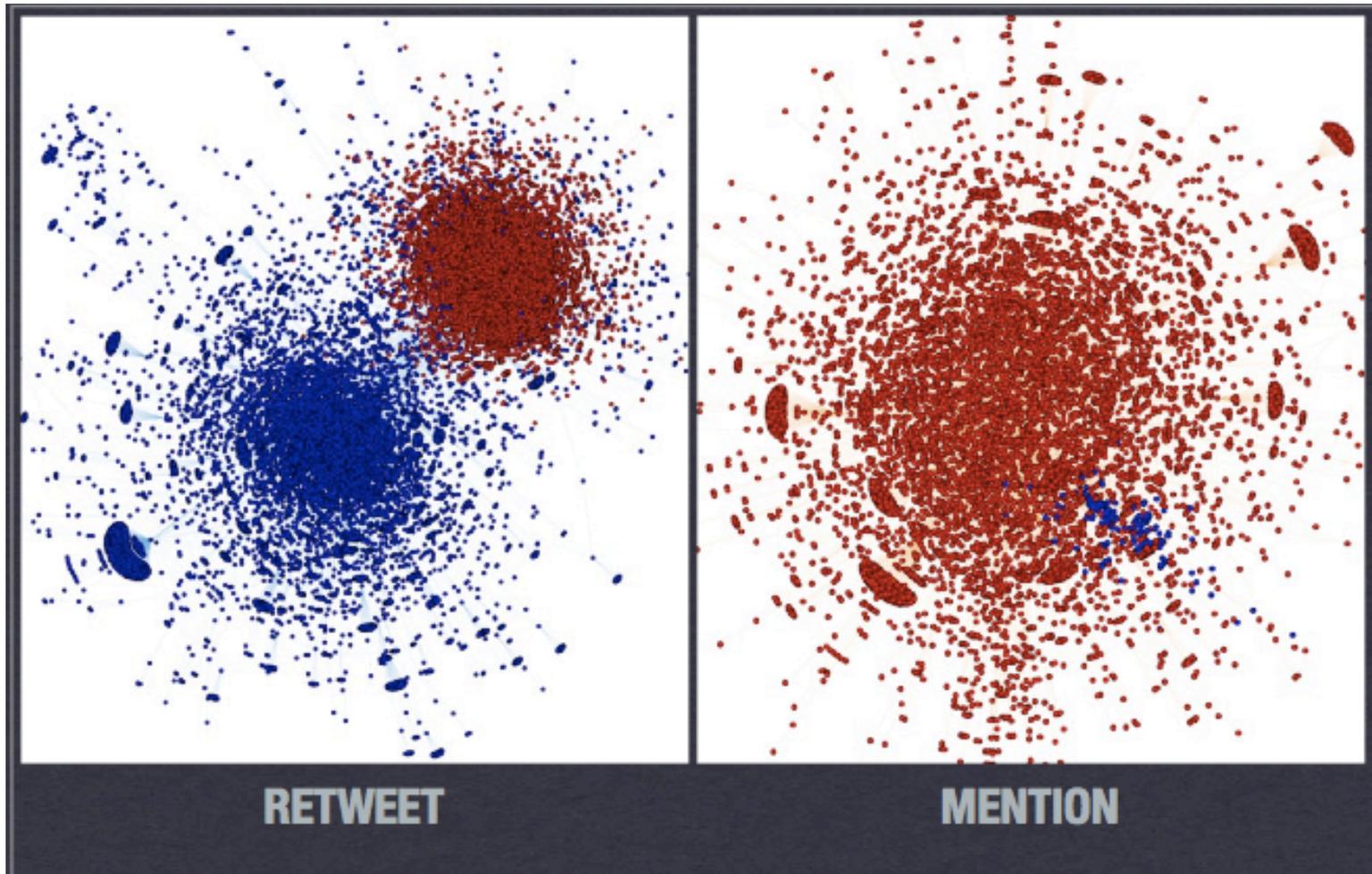
Table 1: Hashtags related to #p2, #tcot, or both. Tweets containing any of these were included in our sample.

Just #p2	#casen #dadt #dc10210 #democrats #dul #fem2 #gotv #kysen #lgf #ofa #onation #p2b #pledge #rebelleft #truthout #vote #vote2010 #whyimvotingdemocrat #youcut
Both	#cspj #dem #dems #desen #gop #hcr #nvsen #obama #ocra #p2 #p21 #phm #politics #sgp #tcot #teaparty #tlot #topprog #tpp #twisters #votedem
Just #tcot	#912 #ampat #ftrs #glennbeck #hhrs #iamthemob #ma04 #mapoli #palin #palin12 #spwbt #tsot #tweetcongress #ucot #wethepeople

Rete e comunita'

- Rete basata su retweets e mention
- Informazione passa da A a B
- RT: 2 raggruppamenti di utenti che diffondono contenuti all'interno della loro comunita'
- MN: non ci sono raggruppamenti distinti

Multi-mode communication



Analisi dei contenuti

- Raggruppamenti sono basati sulle caratteristiche della rete comunicativa (chi parla con chi)
- Sono basati anche sui contenuti della discussione?
 - Similarita' di coseno tra 2 utenti all'interno dello stesso raggruppamento e tra due raggruppamenti diversi
- RT: utenti nel raggruppamento A hanno un profilo piu simile tra di loro che non a quello degli utenti del raggruppamento B
- MN: Somiglianza non attestata

Polarizzazione politica

- C'e' un legame tra i raggruppamenti nella rete dei RT e utenti con opinioni politiche simili?
- Necessario effettuare un'analisi qualitativa dei contenuti
- Conrollare se i tweets di un certo utente sono di sinistra, destra o di orientamento politico non chiaro
- Annotazione dei tweets

Divisioni politiche

Table 4: Partisan composition and size of network clusters as determined by manual inspection of 1,000 random user profiles.

Network	Clust.	Left	Right	Undec.	Nodes
Retweet	A	1.19%	93.4%	5.36%	7,115
	B	80.1%	8.71%	11.1%	11,355
Mention	A	39.5%	52.2%	8.18%	7,021
	B	9.52%	85.7%	4.76%	154

Interazione

- Utenti interagiscono con altri utenti con la stessa opinione (retweet)
- Nella rete MN piu' interazione tra utenti con diverse opinioni

Esempio

User A: Please follow @Username for an outstanding progressive voice! #p2 #dems #prog #democrats #tcot

User B: Couple Aborts Twin Boys For Being Wrong Gender..<http://bit.ly/xyz> #tcot #hhhs #christian #tlot #teaparty #sgp #p2 #prolife

Conclusioni

□ Analisi

- ▣ Retweet, mention, # sono usati in modo differente
- ▣ RT usati per condividere contenuti all'interno della stessa comunita'
- ▣ MN per condividere contenuti attraverso comunita' differenti

Denk e le elezioni politiche olandesi

VERKIEZINGEN

Nep-aanhang online actief voor Denk

De politieke partij Denk heeft op sociale media steun van nepprofielen. Het doel is de publieke opinie te beïnvloeden.

Door onze redacteuren
Andreas Kouwenhoven en
Hugo Logtenberg

AMSTERDAM. Voor de politieke partij Denk zijn 'internet-trollen' actief die de partij via Twitter en Facebook steunen en politieke tegenstanders aanvallen. De inzet van trollen wordt besproken in interne communicatie van de partijtop van Denk.

Dat blijkt uit onderzoek van NRC,

dat inzage heeft gehad in meerdere WhatsAppgroepen waarin medewerkers en politici van Denk deelnemen.

De partij van Tunahan Kuzu en Selcuk Öztürk is omstreden omdat zij andere politici van buitenlandse afkomst online verdacht maken.

De appberichten suggereren dat hierbij gebruik wordt gemaakt van zogenoemde trollen, ook wel *sock puppets* genoemd. Het zijn online identiteiten waar een bestaand persoon achter lijkt schuil te gaan, maar die in werkelijkheid worden aangestuurd door individuen of organisaties, met als doel de publieke opinie te beïnvloeden.

In het geval van Denk gaat het om minimaal twintig nepprofielen op Twitter en Facebook. Die zijn samen

verantwoordelijk voor minstens 1.636 berichten en 2.171 'likes' ten behoeve van de Denk-campagne. De bevindingen worden bevestigd door meerdere experts, onder wie Richard Rogers, hoogleraar nieuwe media en digitale cultuur aan de Universiteit van Amsterdam. „Deze valse identiteiten worden ingezet om de beweging groter te laten lijken. Het is een tactiek om invloed uit te oefenen.”

De trollen hebben nauwelijks vrienden of volgers en reageren vrijwel uitsluitend op berichten van Denk en haar jongerenbeweging *Oppositie*. Hun profielfoto's zijn meestal van internet geplukt. De trollen gebruiken onder meer foto's van een Nederlandse student en een Belgische theatermaker. Uit navraag blijkt

dat beiden daarvan niet op de hoogte waren. Het gebruiken van iemands foto zonder toestemming, is in strijd met het portretrecht en de Wet bescherming persoonsgegevens, zegt een woordvoerder van de Autoriteit Persoonsgegevens.

De trollen worden ook ingezet om critici van Denk van repliek te dienen. Zo werd een columnist die op de website *Voorbeeld Allochtoon* een kritisch artikel schreef over Kuzu, belaagd door trollen. Hij besloot hierna de kop van het artikel af te zwakken.

Ook PvdA-minister Lodewijk Asscher en PvdA-Kamerlid Ahmed Marcouch zijn doelwit van trollen. Zo vraagt Denk-politicus Farid Azarkan op 30 juli 2016 bijvoorbeeld in de groepsapp om Marcouch aan te pak-

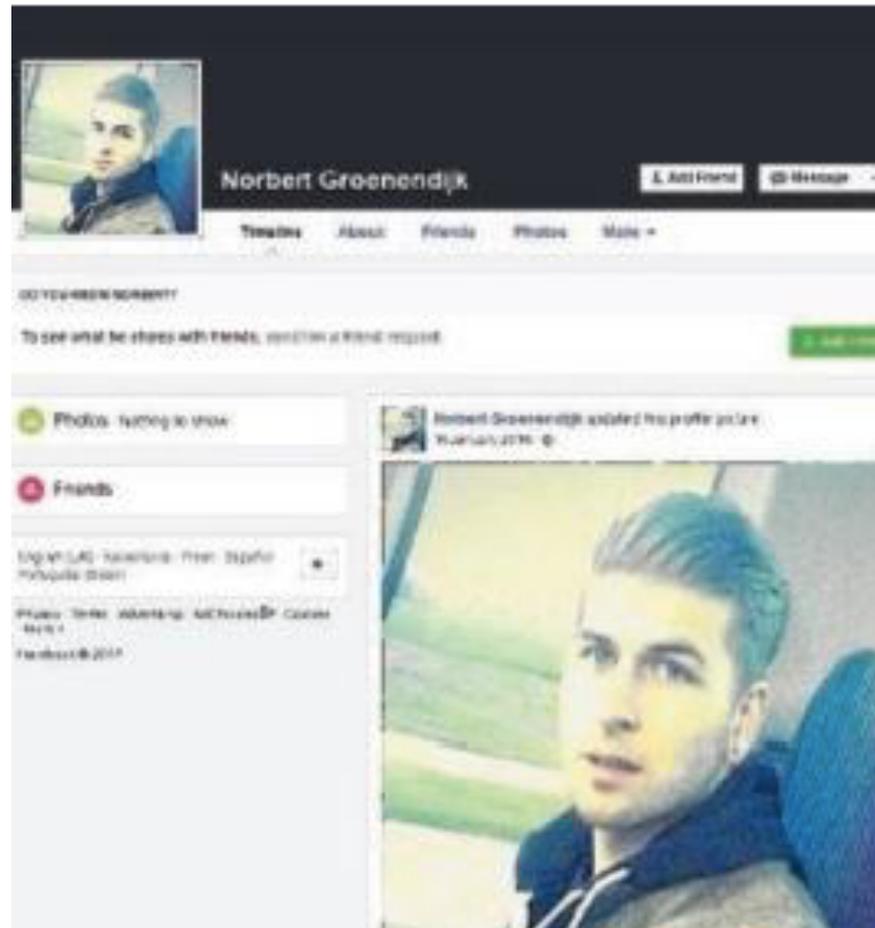
ken nadat deze kritiek uitte op Denk. Azarkan: „Kan er iemand op Marcouch antwoorden: Jij bent helemaal niet geïnteresseerd in een wezenlijke bijdrage van DENK. Jij wilt alleen vaststellen dat DENK niet deugt. Erg doorzichtig en heel hypocriet.” Daarop reageert Enes Yigit, de voorzitter van de jongerenorganisatie van Denk: „Zal ik het doen met een troll account?” Azarkan: „Zeker.” Yigit: „Ga ik doen.”

Of andere politieke partijen in Nederland trollen gebruiken, is onbekend.

Denk heeft, ondanks meerdere verzoeken, niet gereageerd op vragen van NRC.

[+] **De trollen van Denk** pagina 4-5

Denk e le elezioni politiche olandesi



Denk e le elezioni politiche olandesi



HAN

Pedagogiek

Waar op traward

HAN • Bachelor opleidingen • Pedagogiek • Kennismaken en behalen • Studeren in

Jaap, student Pedagogiek: 'Om hbo-pedagoog te worden behandelt iedereen een andere weg'

Jaap is tweedeklasser Pedagogiekstudent en vertelt over zijn 1e studiejaar Pedagogiek aan de HAN in Nijmegen. Je hebt met je medestudenten allemaal hetzelfde doel, namelijk hbo-pedagoog worden, maar iedereen behandelt een andere weg.'

Jaap, weet je nog toen je dacht 'deze studie past echt bij mij, ik heb goed gekozen'?

Jaap: De vaardigheden in het 1e jaar van Pedagogiek vond ik erg interessant -de 3 door gespreksvaardigheid zowel ouders als jongeren in hun krachten zetten vind ik bijvoorbeeld interessant. Je leert er terecht te luisteren. Je hebt gezellige conversaties en kijkt ook veel naar echt goed luisteren is een hele kunst.

Maakt het goed luisteren is het ook belangrijk om actief vragen te stellen om iemand zo goed mogelijk te kunnen begrijpen. De mens is vaak geneigd om wat men eigen voorkeuren ergens over te geven. Wat moet je ook in aanpakken je gevormd bent door je eigen ervaringen. De kunst is om met een houding van niet weten een gesprek in te gaan, zodat je leert iemand echt te begrijpen.

Wat vond je van het 1e jaar HAN Pedagogiek?

Wat 1e jaar van Pedagogiek heb ik vooral ervaren, de een leeractiviteit met het vak Pedagogiek. Naast de week-activiteiten, de stage en de theorie staat je open omvulling met het meeste bij. Het is belangrijk om erachter te komen wie jij bent en welke kwaliteiten jij bent om spouders en leerlingen te ondersteunen.

Je hebt met je medestudenten allemaal hetzelfde doel, namelijk hbo-pedagoog worden, maar iedereen behandelt een andere weg. We hebben allen een andere achtergrond en hebben onze eigen voorkeur voor een bepaalde doelgroep.



Twitter e relazioni

- Analisi della topologia, tipo di relazioni e di come l'informazione si diffonde all'interno dei social media
- Reti online con una struttura interna di cui è possibile identificare l'intensità e il tipo di relazioni
- Legami deboli e forti (Granovetter 1973) anche all'interno dei Social Media
- Legami deboli creano ponti attraverso diverse comunità → Importanti per la diffusione dell'informazione

Twitter e le relazioni

- Legami forti sono spesso all'interno di una stessa comunità
- I contenuti ma anche la frequenza dell'informazione giocano un ruolo nella diffusione dell'informazione.
- Possibile differenziare tra comunicazione personale (MN) e diffusione dell'informazione (RT)
- Grafo con le relazioni tra gli utenti (rete dei followers e followees)
- Grabowicz et al (2012) Social features of online networks: The strength of Intermediary ties in online Social media PLoS ONE 7(1)

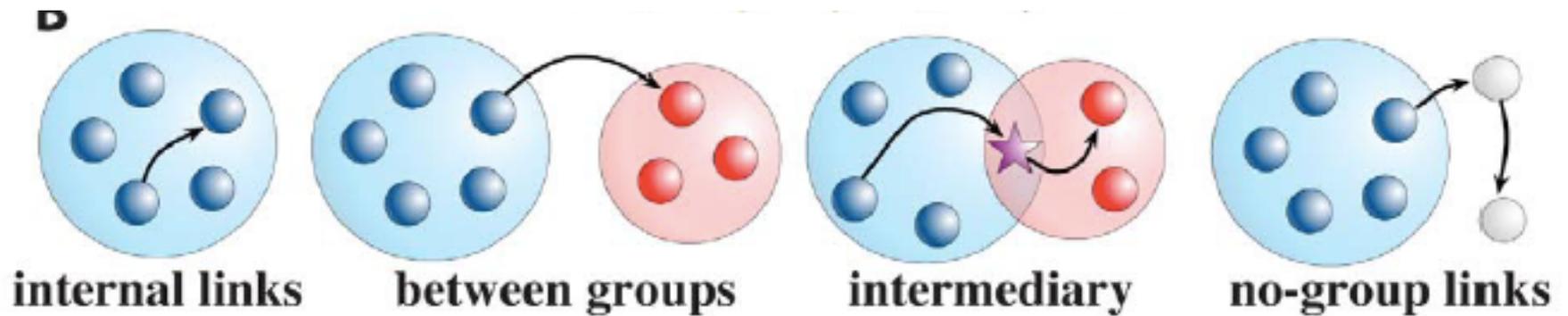
Ipotesi di ricerca

- Analizzare se la struttura della rete e' indicativa riguardo a dove hanno luogo relazioni personali (MT) o la trasmissione di informazioni (RT)

Metodologia

- Identificazione dei followers e followees
- Identificazione delle attività degli utenti: Tweets, MN, RT
- Rete creata sulla base dei followers e followees di ogni utente
- Identificazione dei raggruppamenti: clustering
- Classificazione dei collegamenti (links): interni, tra gruppi, intermediari e collegamenti che non si riferiscono ad un raggruppamento

Links



Risultati

- MT sono piu' frequenti all' interno dei raggruppamenti o tra gruppi simili
- RT sono piu' frequenti tra gruppi (specialmente se differenti) o nel caso di intermediari tra gruppi.
- Relazione con la teoria:
 - Legami deboli sono importanti per diffondere le informazioni tra gruppi diversi
 - Legami forti sono presenti all'interno di un gruppo o tra gruppi simili.

Domande

- Ci sono somiglianze e/o differenze tra i due articoli?
- L'analisi dei dati di Twitter che cosa ci dice riguardo al tipo di comunicazione che avviene all'interno dei raggruppamenti?
- I Social Media polarizzano la comunicazione?

Proprieta' dei dati

- Chi possiede i dati?
- Internet come spazio pubblico anonimo
 - creazione di una propria identita'
 - formazione di relazioni e comunita'
- Reazione a spazio pubblico urbano sempre piu' commercializzato
- Social Media offrono ulteriori possibilita'
 - successo si basa sulla presentazione/promozione di se stessi e non su anonimato (Internet)

Proprieta' dei dati



Web era sinonimo di diversita' ma ora ha reso possibile una centralizzazione dell'informazione (FB, Google) e il risultato e' che siamo diventati sempre piu' impotenti verso governo e societa'

(Hossein Derakhshan – Iranian/Canadian blogger)

Tim Berners Lee e il web aperto

- https://www.youtube.com/watch?v=Jev2Um-4_TQ

Data Selfie

- <https://dataselfie.it/#/about>

Azioni stratagemmatiche

- Privatizzazione dello spazio pubblico: dimensione fisica e digitale.
- Necessita' di creare nuove forme spaziali e processi che possono far emergere nuove comunita' e forme associative
- Citta' come playground in cui nuovi attori danno vita ad azioni stratagemmatiche per riappropriarsi dello spazio pubblico

Banksy

By CARA BUCKLEY and J. DAVID GOODMAN NOV. 5, 2013



Hong Kong Umbrella Movement



Azioni stratagemmatiche

- Nuovi soggetti civici che combattono contro attori istituzionali dominanti
- Comunicazione e media come strumenti di guerriglia
- Citta' come terreno di gioco in cui iniziare azioni stratagemmatiche che permettono di riconquistare e creare spazio pubblico in cui visioni contrastanti si possono confrontare
- Azioni non seguono un percorso definito ma improvvisano traiettorie attraverso spazi ibridi (fisici e digitale, locali e globali) creati da social media e dispositivi mobili

Azioni strategemmatiche

- Sia le azioni che i soggetti che le promuovono sono effimeri e instabili
- Carattere temporaneo permette ai soggetti di sfuggire alle strutture di controllo e di creare nuovi spazi di autonomia
- Norme e strutture vengono invalidate: vuoto come spazio positivo che permette nuove possibilità

Banksy

- Anonimato
- Pubblico coinvolto nella creazione artistica, localmente e globalmente attraverso i media
- Azioni artistiche su spazio urbano per re-territorializzare lo spazio
- Trasformano un luogo anonimo della città' in uno spazio pubblico ludico (Dismaland)
- L'arte di Banksy come arte critica che mina l'ordine egemonico dominante e favorisce una varietà di spazi pubblici.

HKUM

- Occupazione delle aree commerciali per promuovere azioni politiche (75 gg di occupazione)
- Social media e Internet usati per mobilitare e collegare dimostranti
- Azione strategica di ritirata da Internet: uso di Firechat per evitare il controllo su Internet
- Si basa su tecnologia Bluetooth, connettività' migliora quando ci sono molte persone collegate e vicine

HKUM

- Obiettivo e' sfuggire il controllo del governo , ma anche ricreare lo spazio pubblico digitale attraverso l'uso di mesh networking
- Banksy e HKUM hanno un obiettivo comune: sfuggire ai meccanismi di controllo e ricreare uno spazio pubblico democratico

Monachesi and Turco (2017) New urban players: *stratagematic* use of media by Banksy and the Hong Kong Umbrella Movement. *International Journal of Communication*.



Citta' e Big Data

Citta' sono grandi produttori di dati



Smart cities

- Due definizioni:
 - Dispositivi che permettono di catturare e monitorare dati e informazioni che permettono di regolare flussi e processi urbani in modo efficiente => ICT per regolare la città' da un punto di vista tecnologico
 - Economia basata su innovazione e creatività' che sono alla base di nuove imprese => ICT per potenziare capitale umano

Kitchin R (2014) The real-time city? Big data and smart urbanism. *GeoJournal* 79: 1–14

Smart cities

- Soluzioni tecnologiche
- Privatizzazioni, neoliberalismo
- Progresso socio-economico
- Sostenibilita'
- Competitivita'
- Centri urbani alla base dell'innovazione

Smart cities

- 5 caratteristiche:
 - ▣ Uso di ICT
 - ▣ Sviluppo basato su business e neoliberal
 - ▣ Focus su dimensione sociale e umana, basata su creatività'
 - ▣ Agenda basata sul sociale a livello di comunita'
 - ▣ Sostenibilita' dell'ambiente
- Possibili tensioni tra questi aspetti

Hollands, R.G. (2008): Will the real smart city please stand up?, *City* 12:3, 303-320

Smart Cities e Big Data

- Importanza dei Big data per creare questa visione
- Informazione trasparente
- Dati neutrali
- Liberi da ideologia
- Dati ci dicono la verità su situazioni economiche e sociali
- Open data che possono essere usati per sviluppare apps sulla città'

Esempio: Londra

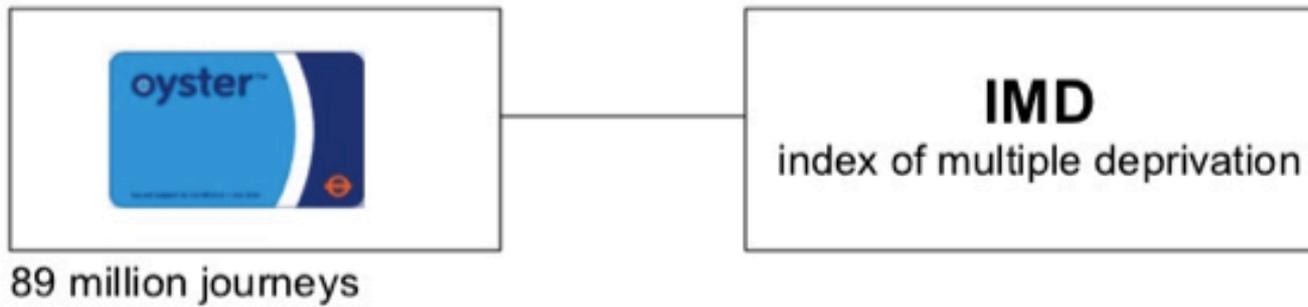
- <http://citydashboard.org/london/>

Smart Cards e comunita' urbane

- E' possibile usare Oyster card per analizzare le varie comunita' cittadine?
- Analisi della relazione tra il flusso urbano dei trasporti pubblici e gli indici dei censimenti delle varie comunita'
- Attraverso l'analisi dei viaggi effettuati si cerca di capire da quali comunita' provengono i viaggiatori
- Goal: monitoraggio spazio urbano

The Hidden Image of the City: Sensing Community Well-Being from Urban Mobility
N. Lathia, D. Quercia, J. Crowcroft, The Computer laboratory, University of Cambridge,
In [Pervasive 2012](#). Newcastle, UK. June 18-22, 2012.

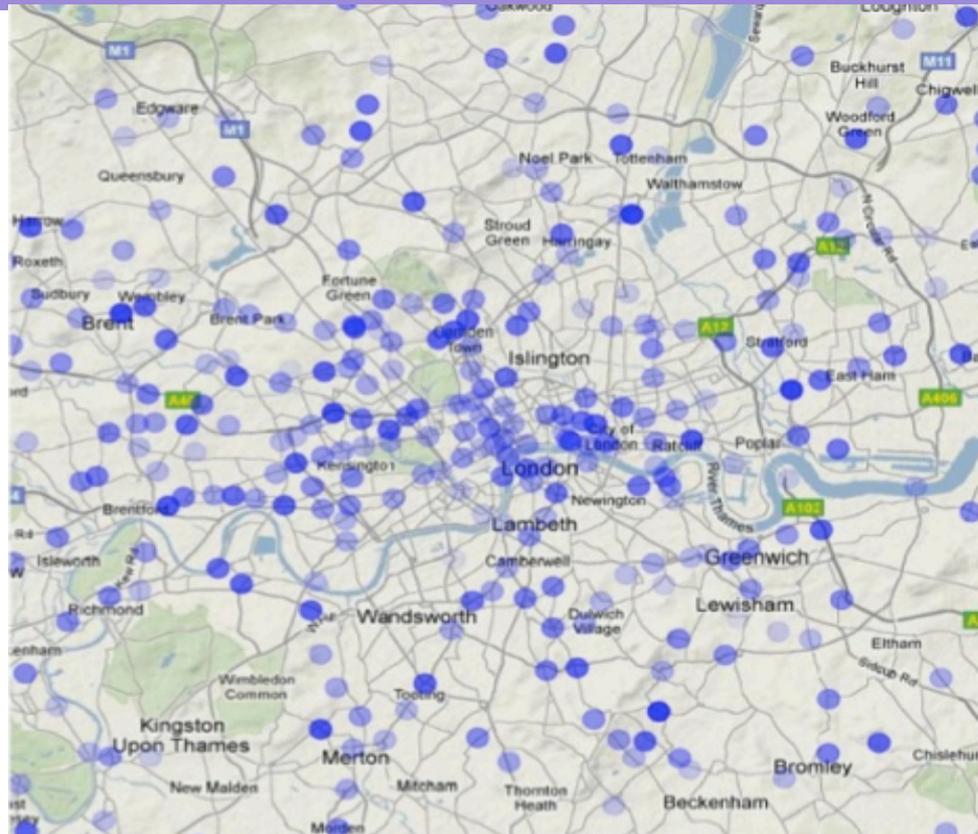
Approccio



Dati

- Dati legati al benessere: IMD
 - ▣ Index of multiple deprivation
 - ▣ Comunita' povere hanno un IMD alto
 - ▣ Comunita' ricche hanno un IMD basso
- Dati oyster card
 - ▣ Tutti i viaggi fatti nel *Marzo 2010*
 - ▣ Pulizia dati: No viaggi bus, inconsistenze
 - ▣ ~76 milioni viaggi, 5.1 milioni di utenti
- Mapping stazioni e valori IMD

Distribuzione geografica valori IMD



Ogni pallino e' una stazione, valori piu' scuri hanno un valore piu' alto IMD

Risultati

- Più le zone sono povere più vengono visitate
- Gli abitanti di Londra tendono a non visitare comunità che hanno un valore IMD simile al loro
- Aree ricche tendono ad attrarre individui che provengono da zone più povere
- Ricchi tendono a non visitare comunità povere
- Segregazione solo in zone povere

Limitazioni

- Indirizzo esatto dei viaggiatori non e' conosciuto
- Non si sa quanto l' Oyster card sia usata nelle varie comunita'
- Non erano a disposizione dati sulla densita' urbana
- Sono state analizzate solo parti della citta' coperte da trasporti pubblici

Chi puo' usare questi dati?

- Pianificatori urbani
- I responsabili delle politiche
 - ▣ Nel prendere decisioni
 - ▣ Infrastrutture di trasporto

Progetto studenti UU: SentiSpot



Analisi dei sentimenti basata su Twitter per identificare l'umore in varie zone della vita notturna di Amsterdam

Metodologia

- Estrazione dati twitter
 - ▣ Twitter search API
 - ▣ Dati salvati in un database MySQL
- Uso di un algoritmo esistente per analisi dei sentimenti
- Visualizzazione dei risultati in una carta
- JS heatmap
 - ▣ <http://www.patrick-wied.at/static/heatmapjs/>

Smart Cities e Big Data

- Cittadini
 - ▣ Informazioni sulla città'
 - ▣ Supporto nelle decisioni
 - ▣ Nuove idee su sviluppo e opportunità'
- Governo
 - ▣ Modo piu' efficiente di gestire la città'
- Aziende
 - ▣ Nuove opportunità' di business a lungo termine

Possibili problemi

- La politica dei dati urbani
 - ▣ Dati non sono ideologici
 - ▣ Sensori non hanno un'agenda politica, riflettono la verità sul mondo
 - ▣ Algoritmi sono neutri e non ideologici basati su obiettività scientifica
- Ma i dati sono complicati
 - ▣ They are never raw
 - ▣ I dati sono risultati di scelte e restrizioni

Possibili problemi

- I dati usati sono determinati da:
 - ▣ Sampling
 - ▣ Tecnologia o piattaforma usate
 - ▣ Il contesto in cui i dati sono generati
 - ▣ Il modo in cui sono classificati
 - ▣ Regole rispetto a privacy, protezione dei dati e sicurezza
- Possono essere importanti, ma i loro limiti non vanno dimenticati

Possibili problemi

- Governo tecnocratico
 - ▣ Vari aspetti della città' possono essere misurati
 - ▣ Problemi possono essere risolti con soluzioni tecniche
 - ▣ Approccio algoritmico che permette soluzioni razionali e imparziali
- Ma anche importanza di altri aspetti come cultura, politica, capitale che giocano un ruolo importante

Possibili problemi

- Hackable city
 - ▣ Servizi e spazi che dipendono da software per funzionare
 - ▣ Problemi di software rendono le città vulnerabili
 - ▣ Possibilità di cyber attack

Possibili problemi

- Interessi delle varie aziende coinvolte
 - ▣ Stretto legame con determinati software
 - ▣ Posizione di monopolio
 - ▣ Marketizzazione dei servizi
 - ▣ Una soluzione per tutte le città dimenticando così l'unicità dei luoghi, culture, popolazioni

Possibili problemi

- Sorveglianza e perdita della privacy attraverso l'accumulo di dati
 - ▣ Cultura del controllo
 - ▣ Dati in possesso di governo e aziende

Esempio

- Possibilita' di analizzare il comportamento collettivo e interattivo dei cittadini online e offline

Come

- **Spazio:** Metodi etnografici per analizzare come gli individui si muovono nello spazio e gli effetti nel creare co-presenza
- **Movimento:** Dispositivi personali per monitorare l'attività degli utenti e i contesti locali
- **Comunicazione:** Language Technology per analizzare uso del linguaggio: social media.

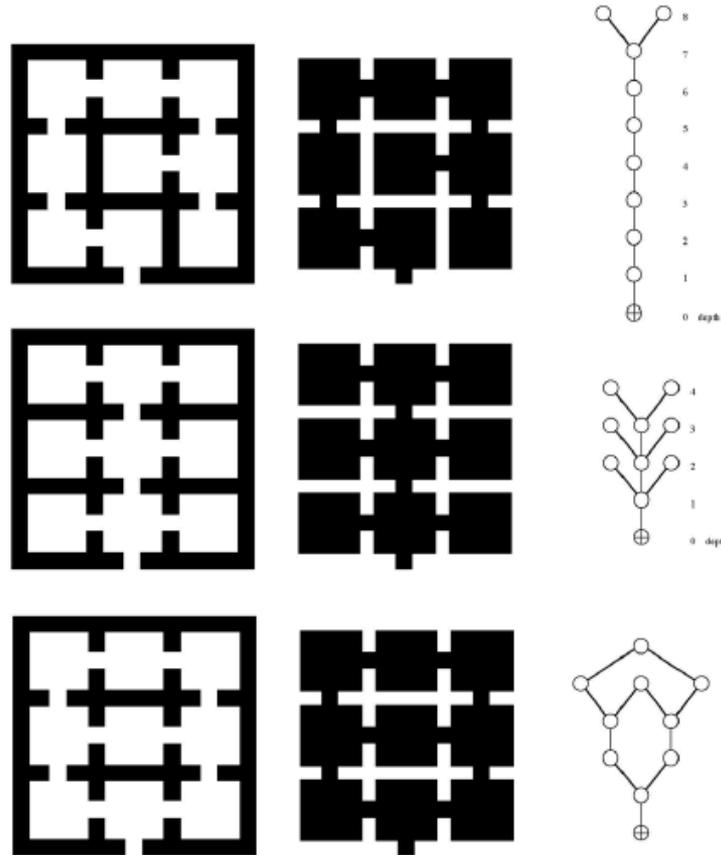
Fusione mondo fisico e mondo virtuale

- Diventa possibile:
 - ▣ Monitorare come gli individui si muovono nello spazio fisico
 - ▣ Analizzare gli effetti del movimento attraverso la traccia digitale che lasciano
 - ▣ Analizzare i dati che sono spesso in linguaggio naturale attraverso tecniche computazionali
 - ▣ Formalizzare l'informazione estratta
 - ▣ Analizzare:
 - Diffusione dell'informazione
 - Scambio di conoscenza
 - Origine dell'innovazione

Analisi dello Spazio

Spazio Fisico

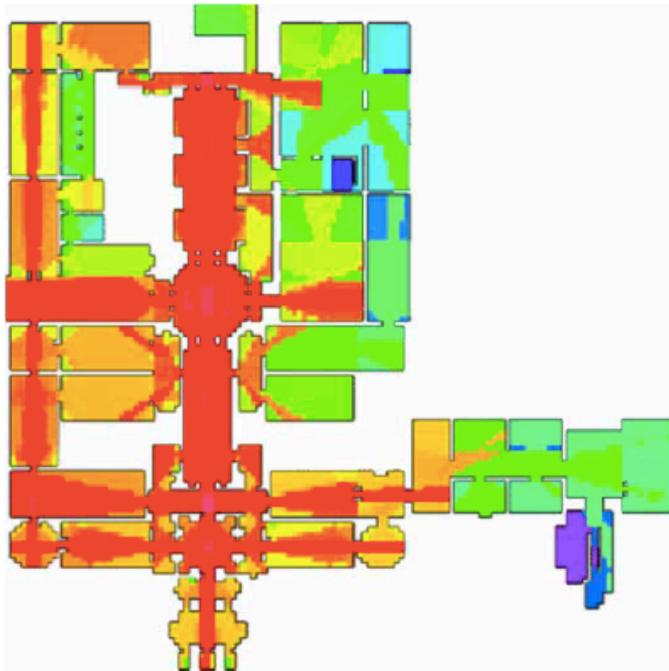
Architectural Space as a Network – The Method of Space Syntax



Spazio fisico e movimento

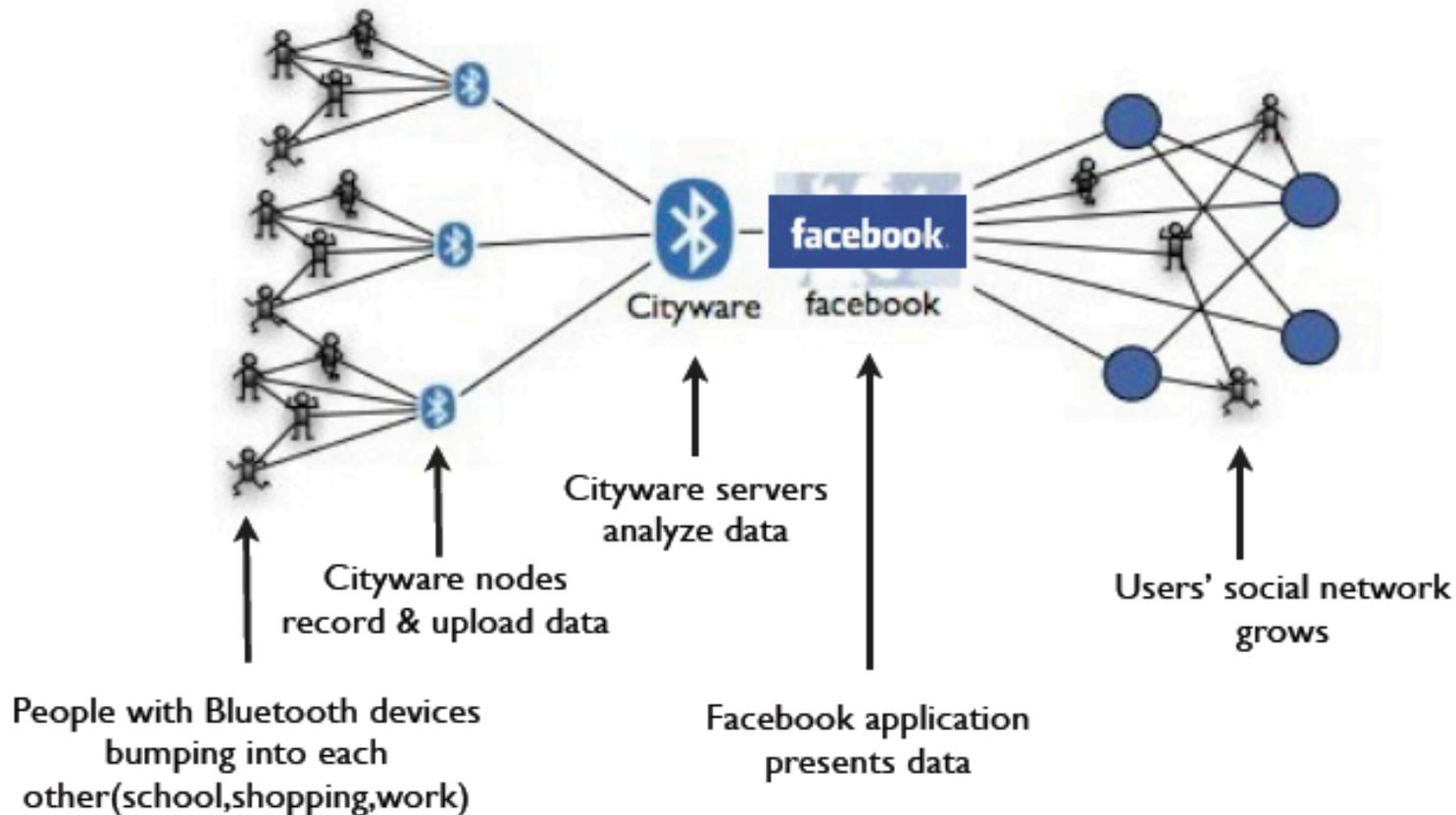
Space Syntax Research: Museums

Spatial configuration and movement flows at Tate Britain



Analisi del movimento

Spatial and transpatial networks



Kostakos and Venkatanathan (2010) Making friends in life and online: Equivalence, micro-correlation and value in spatial and transpatial social networks. Proceedings of IEEE SocialCom, Minneapolis, USA, pp. 587-594

Tipi di reti

- Reti di incontri (Spatial network)
 - ▣ Individui collegati se erano nello stesso luogo durante l'analisi
 - ▣ Rete Facebook (Transpatial network)
 - ▣ Individui collegati se erano amici su FB
- Rete Fusa
 - ▣ Rete incontri e FB unite
 - ▣ 3 tipi di relazioni: Incontri, FB e “fusi”

Risultati

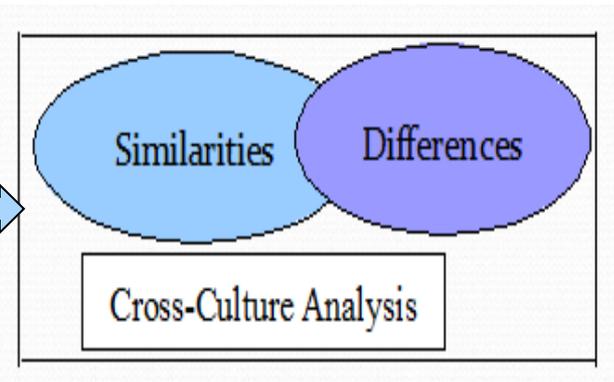
- Reti Bluetooth e Facebook con caratteristiche simili
- Possono essere usate come proxy della rete dell'utente (aspetti simili)
- Rete fusa e' la meno importante
 - Piu' probabile con parenti e colleghi (cf. Granovetter 1973)
- Reti spaziali piu' importanti delle reti transpatial
 - Bluetooth ha il potenziale di registrare "familiar strangers"

Analisi linguistica e interazioni sociali



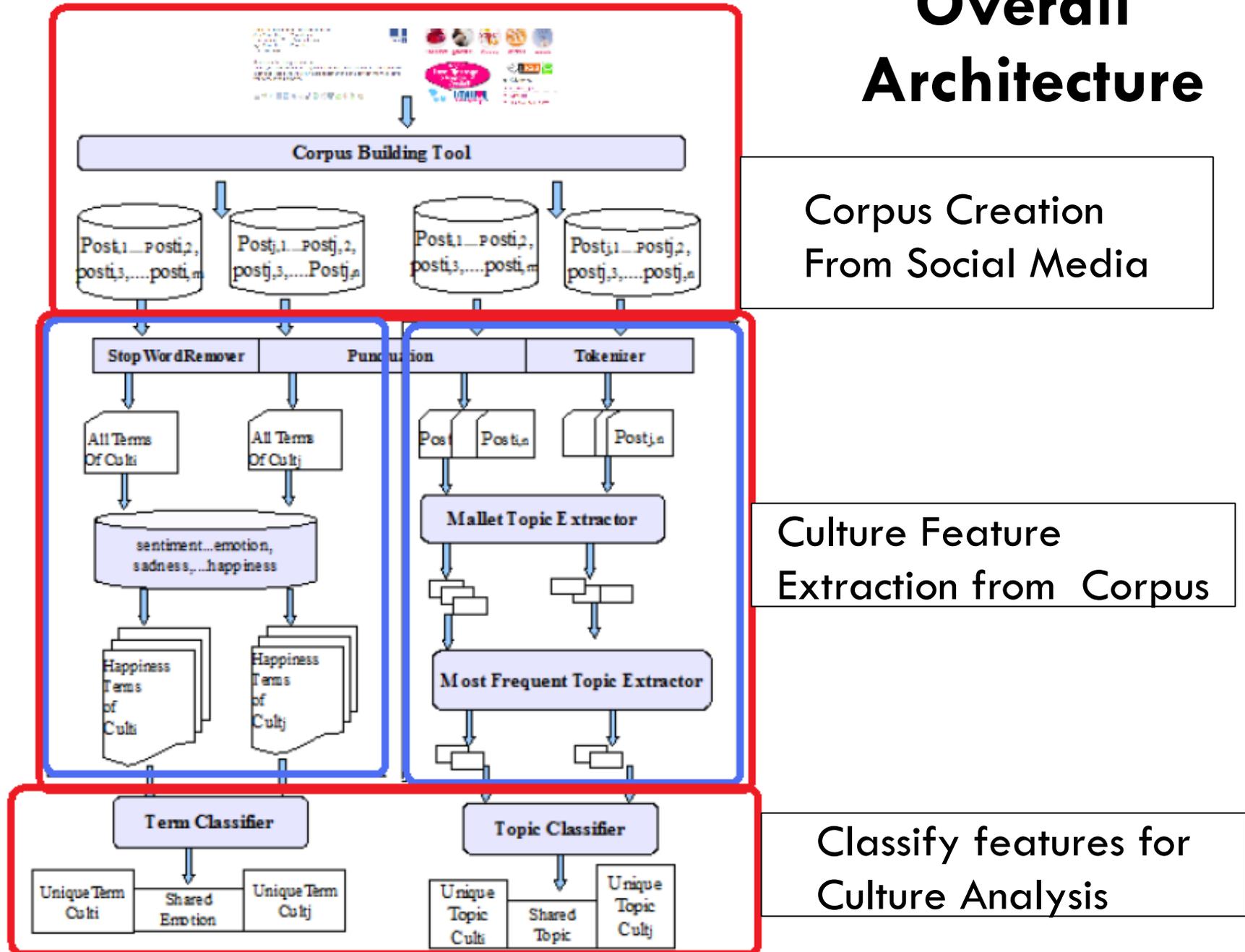
Automatic
Corpus
Creation

Automatic
Culture
Feature
Extraction



Elahi and Monachesi (2012) An Examination of Cross-Cultural Similarities and Differences from Social Media Data with respect to Language Use. In: Proceedings of the Language Resources and Evaluation Conference (LREC 2012)

Overall Architecture



Risultati

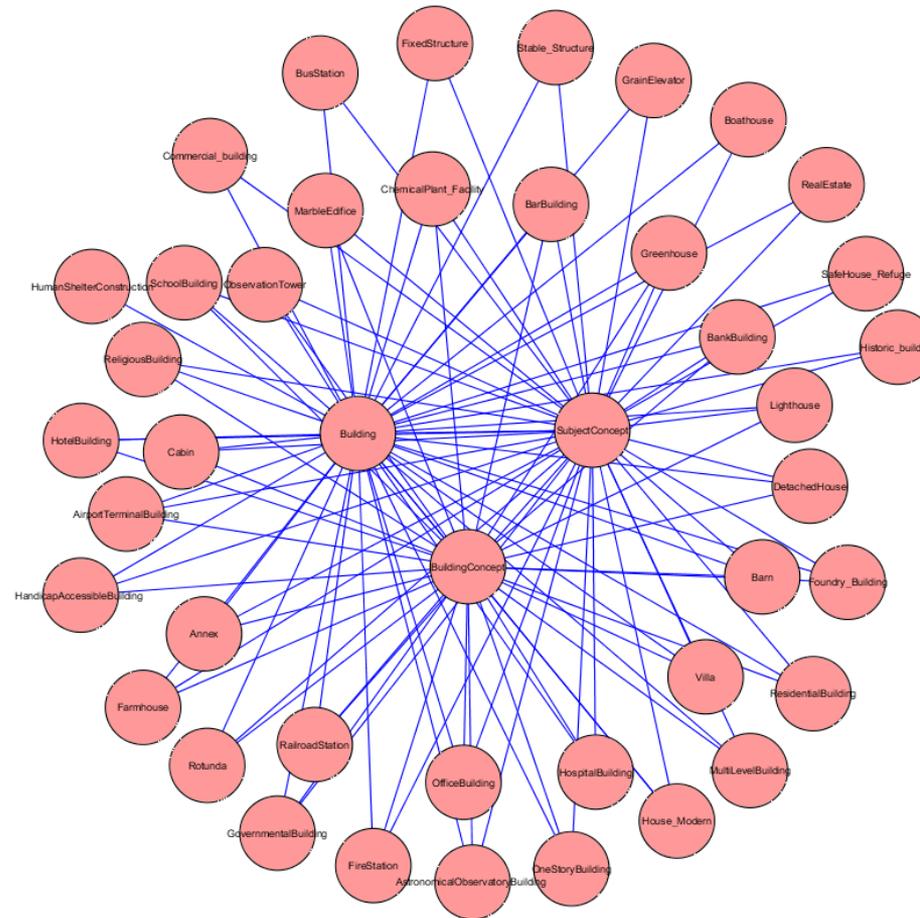
- Discussioni su amore si focalizza piu' sulle emozioni che su religione o sesso (somiglianza)
- Piu' termini relativi a sentimenti positivi/felicità che sentimenti negativi/tristezza (somiglianza)
- Indiani piu' focalizzati sui sentimenti degli americani (differenza)
- Americani sono piu' materialistici degli Indiani (differenza)
- Discussione di questioni specifiche alle due culture

Estrazione dell'informazione e formalizzazione

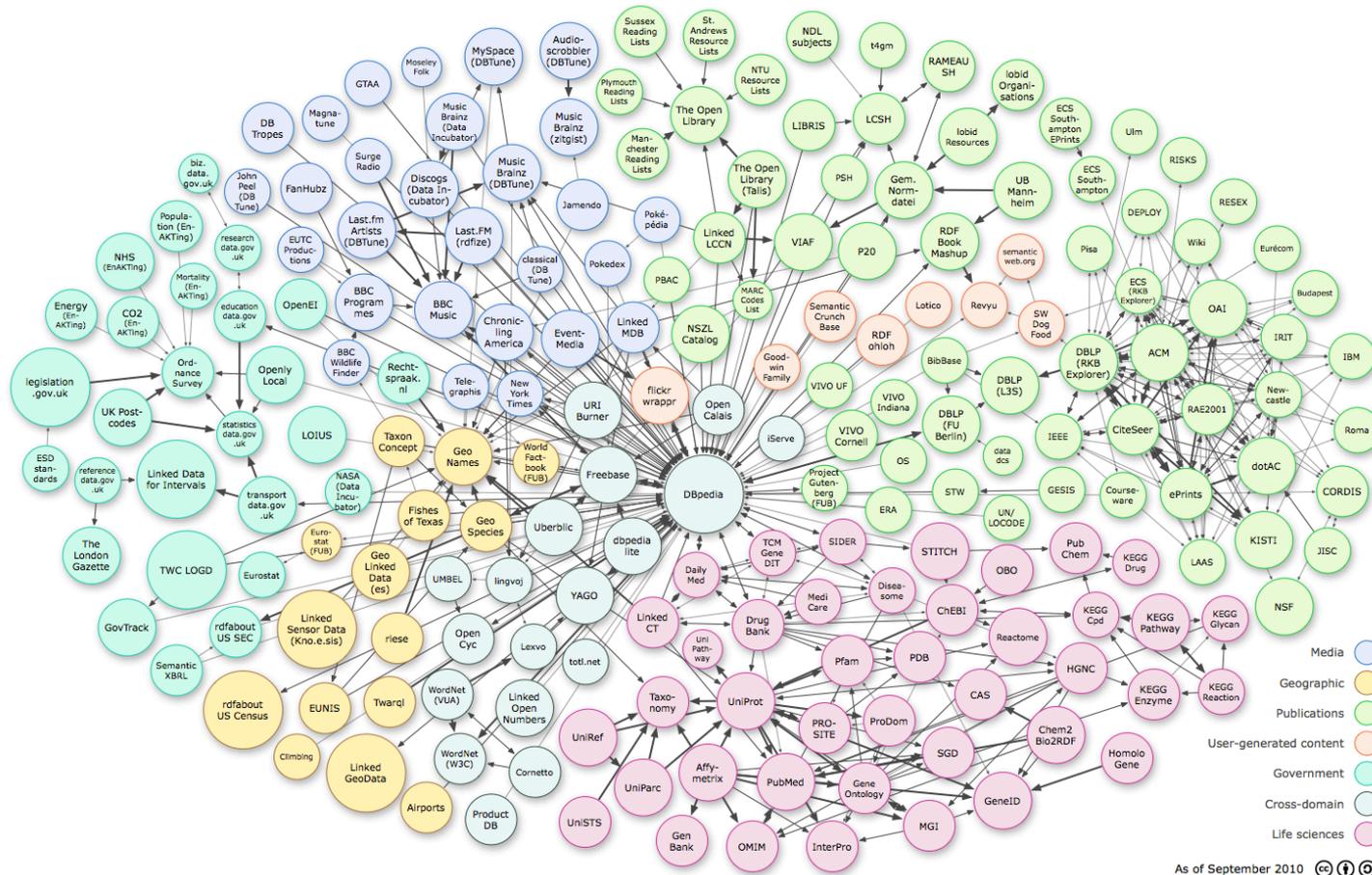
Informazione e conoscenza nelle reti sociali

- Informazione nei social networks e' condivisa con vari pari
- Attenzione e argomenti di interesse giocano un ruolo nella diffusione di informazione (Wu and Huberman 2007, Lerman and Ghosh 2010, Kang et al 2013).
- Analisi di esperti e vocabolario di chi apprende in relazione con la struttura concettuale (Monachesi et al. 2009, Monachesi and Markus 2010).

Formalizzazione della conoscenza



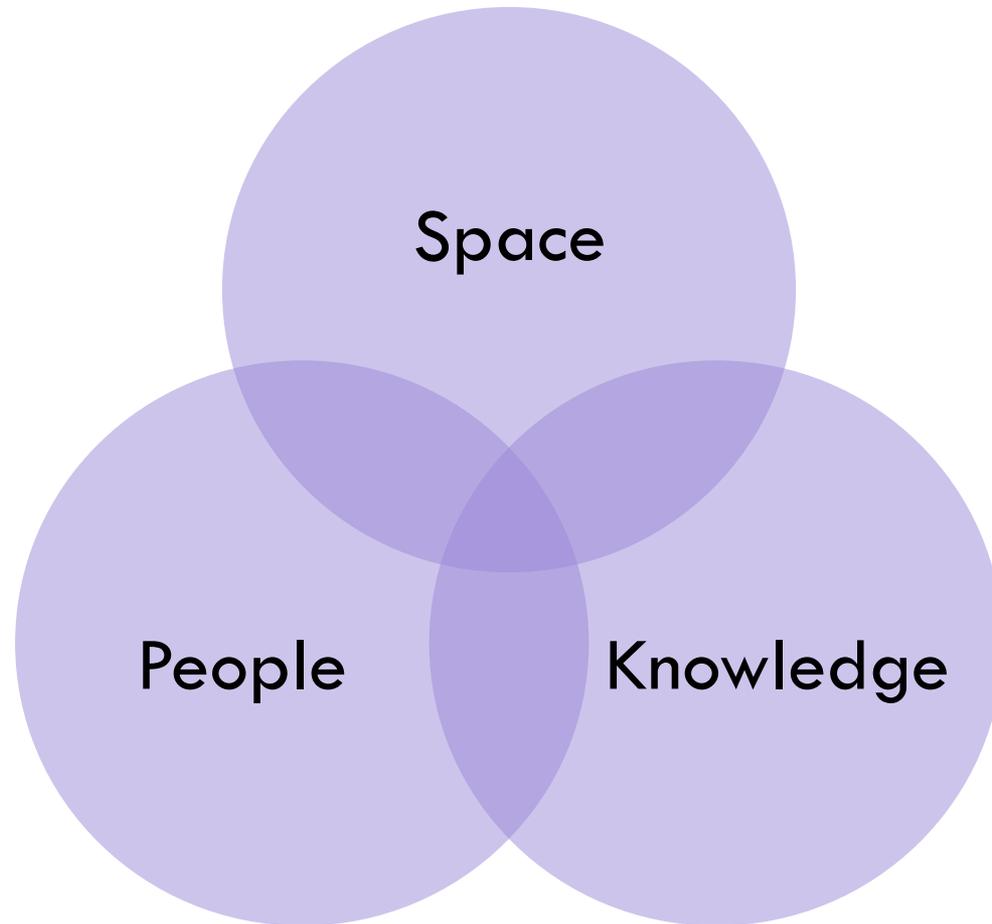
Linked Open Data



Reti

- Qual'è il denominatore comune?

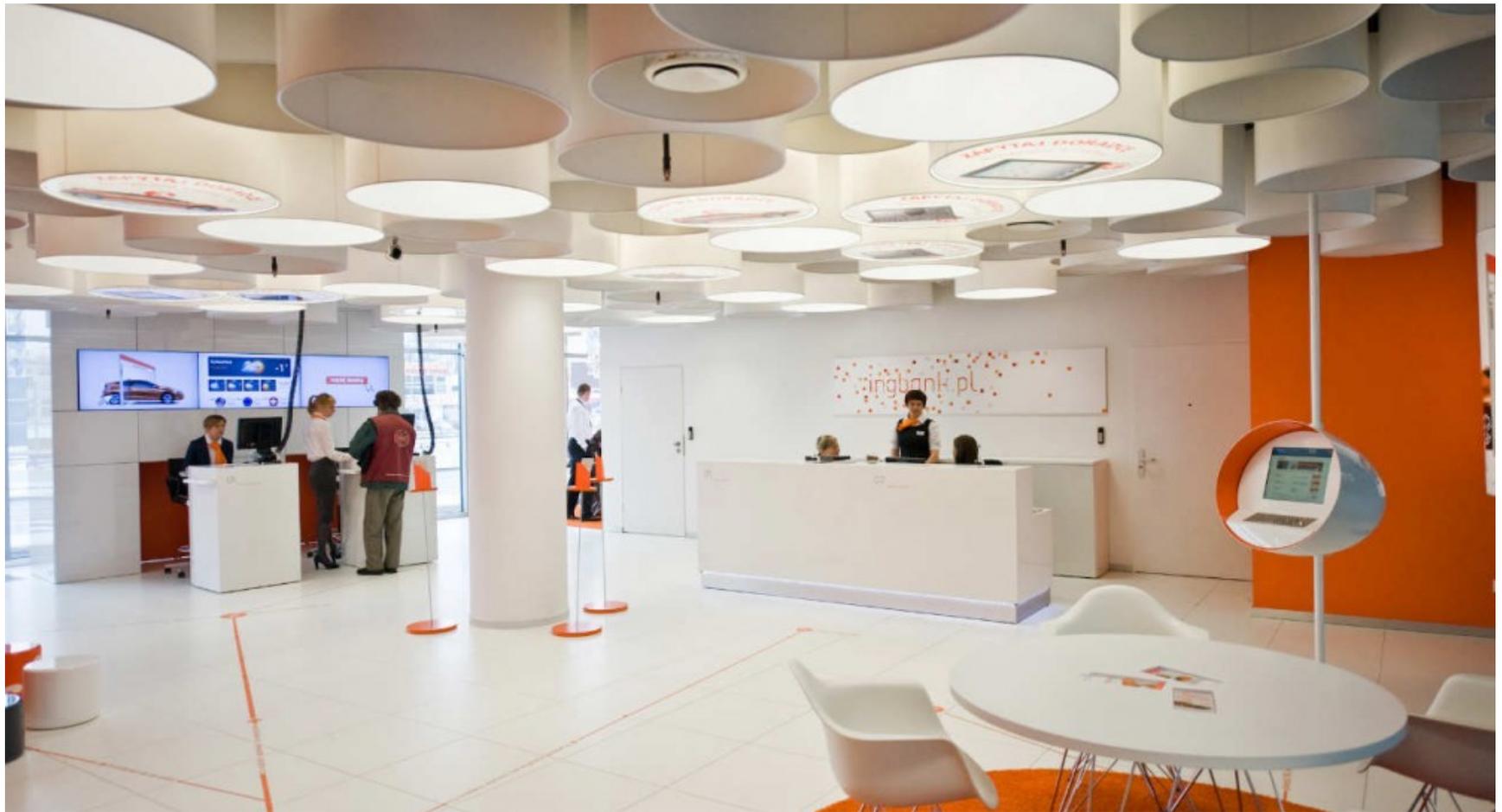
Reti



Fusione spazio fisico e spazio virtuale

- Permette di collezionare grandi quantità di dati
- Analizzare comportamento umano
- Cambia gli spazi fisici

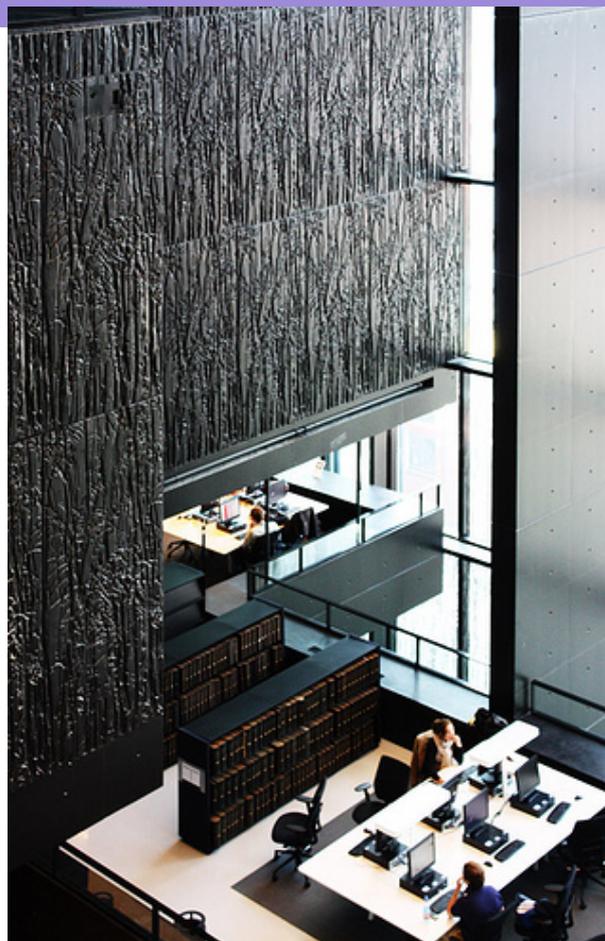
Che cosa e'?



Questo e' come era...



Spazi educativi



Spazi di co-working



Spazi pubblici



Intersituativity in Central Park, New York. Source: Ed Yourdon, Creative Commons

Spazi pubblici

