

La rappresentazione dell'invisibile: il sentiment mapping

Mauro Preda

Università Cattolica del Sacro Cuore, sede di Milano, Dipartimento di Scienze statistiche
Tel. 02/72342305, Fax 02/72343064, mauro.preda@unicatt.it

Abstract

Questo lavoro è il risultato di un'attività interdisciplinare: linguistica computazionale ed in particolare le tecniche per l'estrazione automatica di informazioni implicite contenute nei testi (text-mining), la statistica e la geografia. Il fine è la mappatura del sentiment, estratto grazie all'analisi semantica dei documenti distribuiti nella rete: e-mail, sms, post sui blog, ecc. Le applicazioni pratiche possono essere molte, in ambito economico, marketing, scienze sociali, politiche, di intelligence, ecc.

This work is the result of an interdisciplinary activity: computational linguistics and in particular techniques for automatic extraction of implicit information contained in the texts (text-mining), statistics and geography. The aim is the mapping of the sentiment, extracted through semantic analysis of documents distributed in the network: e-mail, text messages, blog posts, etc.. Practical applications may be many, in economics, marketing, social sciences, political, intelligence, etc. ..

Lo spazio geografico, quello che vediamo svelato, materializzato ed immobilizzato sulla carta, si ottiene grazie ad un processo di trasformazione, noto come "proiezione", che deriva da quella prodigiosa "polvere di proiezione" ricercata nell'antica alchimia che aveva la proprietà di trasformare il metallo vile in oro. Con un processo analogo di trasformazione ora si vuole trattare proprio lo spazio nascosto, quello delle reti o Cyberspazio, in modo tale che lo si riduca e lo si blocchi nelle dimensioni e nelle forme comprensibili ai nostri occhi. Il rapporto tra il reale ed il cyberspazio, inizialmente antitetico, può mutare fino a diventare un rapporto dialettico; la parte nascosta, come vi vedrà in seguito, ci può essere di grande aiuto per meglio interpretare il reale stesso. Come ci dice anche Cristiano Giorda, *nella geografia il virtuale era già presente: nelle carte geografiche e nei GIS ma anche nelle mappe mentali, nei modelli matematici ma anche nei legami affettivi che formano il senso dei luoghi. Come ogni rappresentazione geografica, scrive Lévy, "il virtuale scava pozzi di senso al di sotto della piattezza fisica immediata". Tutta la geografia, effettivamente, non è altro che la virtualizzazione dello spazio: lo spostamento nel campo del linguaggio della relazione fra l'uomo e la Terra. Ludwig Wittgenstein racchiuse tutto questo concetto in una formula notissima: "I limiti del mio mondo sono i limiti del mio linguaggio".*

Uno spazio, quindi, dove una fitta rete di relazioni che costituiscono il tessuto connettivo tra le diverse entità che lo compongono entrano in gioco e lì si muovono. Le piste di movimento per lo più sono invisibili o nascoste, pensiamo alle relazioni nello Spazio-Internet al cosiddetto cyberspazio, a quello spazio silenzioso, multidimensionale ed invisibile, in continua espansione che è rappresentato delle reti complesse di comunicazione tra le quali il WEB. A questo punto possiamo azzardare l'intenzione ambiziosa di addentrarci, in quella che potremmo chiamare "l'altra faccia

della geografia”. L’analisi geografica si trova sempre di fronte a due realtà ben distinte: una che ci conduce comodamente sulla superficie e che ci consente di vedere e misurare Gaia che brilla e risplende alla luce del sole (**Gé** da cui Gaia). L’altro cammino, ci porta, invece, nelle zone oscure e sotterranee (**Ctòn** contrapposto a **Gé**), ctoniche, appunto, quindi, invisibili. Questi spazi dinamici, continui, senza confini ed inafferrabili sono per esempio quelli culturali, religiosi, linguistici, di pensiero e quindi politici, di sicurezza, ecc. In questo continuo rapporto dialettico tra visibile ed invisibile, materiale ed immateriale, si realizza oggi, grazie anche alle reti informative, il territorio stesso. Certo è che la metrica tradizionale ed i pur sofisticati strumenti informatici che ci troviamo ad usare se non adeguatamente riprogettati rischiano di essere insufficienti a cogliere la realtà multiforme e multidimensionale del territorio. La nuova economia globale ha di fatto accorciato le distanze quasi annullandole e compresso il tempo per la velocità dei trasporti e delle informazioni. La distanza è diventata un’entità, un concetto pure difficile da afferrare e misurare. Una trasformazione, questa, del rapporto tra i sistemi di produzione e la rappresentazione dello spazio, che ha appassionato critici dell’economia politica, urbanisti, geografi e filosofi. Riprendendo il pensiero e le parole del Farinelli, ammettiamo che *“rispetto al passato, il Moderno instaura un’inversione di tendenza: la geografia resta l’imitazione della mappa, ma quest’ultima a sua volta non è più la copia del mondo, bensì è il mondo che diventa la copia della mappa. Lo stato moderno fin dalla sua nascita doveva essere necessariamente, proprio come una tavola, continuo, omogeneo ed isotropico, doveva dunque possedere le tre proprietà che nella geometria euclidea appartengono all’estensione. Quello che si chiama potere è stato, modernamente, nient’altro che la pratica applicazione di tale modello, dunque il predominio della logica cartografica.”* *“L’evento decisivo che segna l’inizio della **globalizzazione** cadde nel 1969 negli Stati Uniti con la prima rete di comunicazione elettronica, proprio mentre eravamo davanti al televisore (come me!) a guardare il primo atterraggio sulla Luna. In silenzio due computer iniziavano a dialogare ed a scambiarsi dati, la materia iniziò così a mutarsi in immateriali unità d’informazione. Preso nella rete, il mondo topografico e spaziale, quello che vediamo e che davvero distingue la superficie della Terra da quella del suo astro, il mondo moderno, iniziò a dissolversi. È l’inizio dell’esaurimento del regno del visibile. Da questo momento la crescita urbana si svincola del tutto dal quadro di riferimento dello stato-nazione per dipendere, secondo la logica dell’universalità, del lavoro Perciò la città viene definitivamente a perdere, nel suo funzionamento, i suoi storici attributi euclidei, diventa discontinua e disomogenea e le sue parti sono funzionalmente voltate in tutte le direzioni, si frantumano. La città più importante non è la città più grande, ma questo ci sorprende soltanto perché siamo ancora abituati a pensare il mondo come una tavola, dove la superiorità di una figura rispetto all’altra dipende dall’estensione, è cioè da una nostra valutazione poggiata ancora sulle proprietà e regole della geometria euclidea.”* (Farinelli, 2006)

Il punto di partenza lo dobbiamo individuare proprio nel linguaggio e questo ci pone nella prima condizione di chiederci ancora: ma dall’idea che ci siamo fatta di territorio, cioè di una entità dinamica, multidimensionale, quasi sfuggente, non è riduttivo utilizzare ancora la parola “territorio”? Certo il termine ci induce subito ed erroneamente a credere ad un suo diretto legame con “terra” e quindi ai concetti di estensione, materialità e solidità. L’analisi etimologica della parola rileva, invece, un’origine ben diversa. Infatti, sempre secondo la lezione del Farinelli, un chiarimento lo troviamo nel Corpus Juris di Giustiniano e precisamente nel Digesto, sotto la rubrica «De significatione verborum», ove il territorio corrisponde all’estensione che ricade sotto la giurisdizione del magistrato, è definito dall’atto di dire giustizia, dell’esercizio del potere: il termine quindi non ha nulla in comune con la terra ma discende dalla stessa base di «**terrore**» (terre). *“Per comprenderlo basta considerare l’ambigua figura dello Spaventapasseri, che non a caso nella media e infima latinità si diceva appunto «territorium» [Du Cange, a.v.]: piantato in mezzo al campo ha il compito di rappresentare il diritto, la proprietà che su di esso vige, ma allo stesso*

tempo di tradurre tale diritto in termini naturali, nella forma di spauracchio nei confronti degli uccelli, quasi che il diritto fosse un accessorio spontaneo e immediato del campo stesso." (Farinelli F., 2009) In questa direzione, quindi, il territorio finisce per superare i limiti imposti della fisicità rappresentando il risultato-processo di una sequenza di fasi di controllo sempre più articolate e complesse (controllo intellettuale, materiale e strutturale): questo processo è definito **territorializzazione** (Turco, 1988). A questo punto, se si guarda il territorio, inteso come risultato-processo di territorializzazione, anche alla luce della straordinaria espansione delle reti di comunicazione globali (prime fra tutte Internet) ed insieme l'"infittimento" dei singoli nodi-punti di connessione, non può sfuggire come in esso sia in atto una sorta di metamorfosi che può essere letta da un lato come uno degli effetti della robustezza strutturale delle reti stesse dall'altro come l'esito della spontanea tensione alla comunicazione globale, vedi per esempio il fenomeno dei social network. Ora, di questo processo che interessa più ambiti, dalla vita sociale, civile, politica, ecc. sta proprio alla geografia pure definita come "scienza delle relazioni nello spazio" (Gottmann, 1966). il compito di fornirci un percorso di analisi ed una possibile chiave di lettura. Proprio in questo periodo, in piena globalizzazione, la geografia potrebbe ritrovare le sue antiche origini e la sua identità tornando ad occuparsi proprio del "ciò che per sua natura è invisibile"¹ e la cartografia ritornerebbe, quindi, ad essere uno strumento più concettuale, il primo modello per un riferimento non tanto metrico e geometrico, quanto, invece, culturale, antropico, politico, economico, sociale, ecc. Prendiamo ora le mosse dall'articolo di Steven Johnson, apparso sul WEB ed intitolato "Mind Share"². *Che cosa succede quando si inizia a vedere il WEB,³ il principale servizio presente in Internet, come una **matrice di menti**, non di documenti? La mente diventa una parte dello spazio puro. Il sito personale diventa un'estensione della nostra memoria, come nella visione di Vannevar Bush del Memex, ma i ricordi fanno parte dell'**intelligenza collettiva del web**. E una rete orientata intorno a menti collegate che rende più facile per quelle menti trovare l'un l'altro nella vita reale.* Il pensiero, molto avvincente e chiaro ha, inoltre, il pregio di introdurre il concetto chiave di azione sinergica e collettiva di mente ed intelligenza che si realizzano nelle reti e quindi direttamente sullo stesso territorio. Normalmente, infatti, quando pensiamo al concetto d'intelligenza lo associamo immediatamente, così come ci siamo abituati, a un qualcosa di intimo, soggettivo, personale, diverso da soggetto a soggetto. Pierrè Lévy e Derrick de Kerckhove, introducendo i termini di **intelligenza collettiva** ed **intelligenza connettiva** intendono, invece, proporci una nuova interpretazione, riferirsi a qualcosa di molto più ampio che va oltre l'ambito della sfera privata e personale. *"L'intelligenza collettiva è un'intelligenza distribuita ovunque, continuamente valorizzata, coordinata in tempo reale, che porta ad una mobilitazione effettiva delle competenze"* (Lévy, 1996, p. 34) Assume, aggiungiamo, una valenza ed una dimensione territoriale, l'intelligenza collettiva ha origine dal basso, nasce, appunto, con le azioni di ogni individuo ed esiste proprio grazie ai legami tra le diverse persone. Come scrive, ancora, Lévy: *"Nel momento in cui l'altro è una fonte di conoscenza, vale anche il contrario. Anch'io, qualunque sia la mia posizione sociale attuale, qualunque sia il giudizio rilasciato dall'istituzione scolastica su di me, rappresento per gli altri un'occasione di apprendimento [...] Anche se sono disoccupato, anche se non ho soldi, anche se non sono diplomato, anche se tiro a campare in periferia, anche se non so leggere, nonostante tutto non sono un nulla. Non sono interscambiabile. Ho un'immagine, una posizione, una dignità, un valore personale e positivo nello spazio del sapere. Tutti gli esseri umani hanno il diritto di vedersi riconoscere un'identità di sapere"* (Lévy, 1966, p. 33).

¹ Ci si deve riferire necessariamente ad Anassimandro di Mileto (610 a.C. circa – 546 a.C. circa), filosofo presocratico e come vuole la tradizione, primo cartografo. Fu proprio lui il primo che "osò" tracciare la prima immagine geografica, cosa che gli costò una grave accusa di empietà, per essersi permesso di rappresentare la terra e il mare dall'alto, come soltanto agli Dei è consentito; ma non solo questo, un altro grave errore fu proprio l'aver costretto il Tutto in continuo divenire immobilizzato e ridotto in sole due dimensioni.

² http://www.wired.com/wired/archive/11.06/blog_spc.html

³ WEB acronimo di "World Wide Web" tradotto letteralmente in italiano: "ragnatela grande quanto il mondo".

L'intelligenza connettiva è, inoltre, intersoggettiva, nel senso che la conoscenza si trova tra i soggetti e non solo all'interno dei soggetti stessi. Connessione, quindi, non vuol dire massificazione, omologazione ma sinergia e collaborazione per il raggiungimento di un obiettivo comune: un'occasione di crescita libera⁴ e spontanea.

Prende quindi avvio da queste condizioni di base l'idea di attingere proprio dal WEB quei dati distribuiti che servono alla realizzazione del prototipo di mappa del sentiment. I dati in questione sono di fatto i testi liberi e spontanei es. i blog, un corpus di dati testuali ed informativi che rappresentano un concentrato di informazioni e di emozioni, un corpus che Ong ridefinisce in una felice sintesi come "oralità secondaria". La fase più recente della storia della comunicazione, la fase dell'oralità secondaria, dominata dagli strumenti della tecnologia elettrica ed elettronica (dal telegrafo al Web), che hanno radicalmente trasformato i criteri di scambio delle informazioni producendo eccezionali conseguenze in tutti i settori della società e della cultura, ha, riprendendo le parole di Ong *"sorprendenti somiglianze con quella più antica per la sua mistica partecipatoria, per il senso della comunità, per la concentrazione sul momento presente e persino per l'utilizzazione di formule. Ma si tratta di un'oralità più deliberata e consapevole, permanentemente basata sull'uso della scrittura e della stampa, che sono essenziali per il funzionamento delle attrezzature, nonché per il loro uso. L'oralità secondaria è molto simile, ma anche molto diversa da quella primaria. Come quest'ultima, anche la prima ha generato un forte senso comunitario, poiché chi ascolta le parole parlate si sente un gruppo, un vero e proprio pubblico di ascoltatori, mentre la lettura di un testo scritto o stampato fa ripiegare gli individui su di sé"* (J.W. Ong, 1986). Proprio grazie alla costante espansione della comunicazione globale, si è pensato di rappresentare su carte tematiche quelle informazioni implicite che si possono recuperare dalla grande quantità di testi distribuiti nel WEB. Le applicazioni possono essere molto numerose, solo a titolo di esempio possiamo prevedere applicazioni, in ambito di geomarketing, scienze sociali, economiche, politiche, ecc.

A questo punto gli ingredienti necessari per poter avviare un processo di analisi che ci porti a rappresentare cartograficamente i pensieri distribuiti nel WEB ci sono tutti e gli strumenti necessari sono nell'ordine: l'analisi semantica con il text mining⁵, l'analisi statistica e l'analisi spaziale.

Grazie alla collaborazione di due società, la Expert System, società leader in software semantici, ed il IlGiornale.it, quotidiano on-line nazionale, è stato possibile realizzare un primo prototipo di analisi. La prima ha gentilmente concesso l'utilizzo del software per l'analisi semantica; la seconda ha concesso l'accesso ai dati da analizzare. Quello che è stato sviluppato è quindi un prototipo sperimentale, volto a scoprire la validità e l'utilità pratica di questa idea. Si è cercato di rispondere all'esigenza de Il Giornale.it di conoscere più a fondo i propri lettori, analizzando le opinioni ed i commenti, gli scambi di idee, che gli utenti registrati lasciano sul sito.

Negli ultimi anni c'è stato un boom dell'utilizzo di internet e di conseguenza dello scambio di informazioni tramite la rete. Al giorno d'oggi Internet contiene un'enorme quantità di dati immagazzinati in miliardi di pagine ed è usato per gli scambi di informazione in tutto il globo. Sono numerosissimi i forum dove gli iscritti si scambiano informazioni e ormai ogni sito internet possiede una sezione dove è possibile commentare ed inserire la propria opinione su prodotti o tematiche di ogni genere. Le ricerche basate sul semplice utilizzo di parole chiave spesso inefficienti sono un collo di bottiglia per lo sviluppo del web. Una delle soluzioni proposte per ovviare a questo problema consisterà nel dotare le pagine web di notazioni semantiche, che potranno essere utilizzate per una ricerca più efficiente delle informazioni. Nel prossimo futuro

⁴ Le attuali rivolte contro i regimi "dell'età della globalizzazione" che esplodono prima nella rete sono un chiaro esempio del processo di auto-organizzazione dei popoli oppressi dalle dittature.

⁵ Il text-mining consiste in un processo automatico di analisi software che consente l'estrazione di informazioni implicite contenute in grandi quantità di testi e documenti scritti. L'informazione acquista più valore quando è collocabile pure spazialmente.

questo potrà divenire realtà, con l'avvento del web 3.0. Uno dei motivi per cui questo non è ancora stato attuato consiste in un circolo vizioso da cui è difficile uscire. Infatti creare un web semantico senza avere gli strumenti software adatti a sfruttarne le potenzialità risulta inutile; lo stesso vale al contrario, infatti implementare software che utilizzano dati semantici, che non esistono ancora e non si sa se mai ci saranno, è una spesa di risorse ingiustificata. In attesa di questa evoluzione, per rispondere all'esigenza di estrarre informazioni da testi scritti in linguaggio naturale, sono state sviluppate sofisticate tecniche statistico-linguistiche. Per estrarre informazioni da un testo è necessario prima di tutto fare in modo che il computer capisca il senso delle parole scritte. Il processo che permette di distinguere l'accezione con cui viene utilizzata una parola viene chiamato **disambiguazione**. La disambiguazione semantica al calcolatore può essere utilizzata dunque per estrarre informazioni dai testi, per classificare i testi in diverse categorie. La disambiguazione semantica è il processo di identificazione del significato di ogni parola di una frase secondo il contesto in cui viene utilizzata. In letteratura vengono descritti diversi metodi per arrivare alla disambiguazione del senso delle parole, utilizzando basi informative come il WordNet che è una rete semantica di libero utilizzo per la lingua inglese. WordNet è nata e viene mantenuta presso l'Università di Princeton sotto la direzione del professor George A. Miller. Ora esiste anche una versione di WordNet in italiano, ItalWordNet, nata all'interno del progetto EuroWordNet.

Come ogni rete, la rete semantica è formata da nodi e dai legami instaurati tra di essi. I nodi della rete semantica sono concetti o classi di sinonimi detti *synset* (dall'inglese "set of synonyms"), legati l'uno all'altro tramite diverse relazioni semantiche. Quindi la rete semantica è un grande database della lingua utilizzato come banca dati organizzata gerarchicamente ad albero per cercare il significato di una parola.

I *synset* sono divisi in nomi, aggettivi, verbi e avverbi. Naturalmente ogni singolo lemma, cioè parola nella sua forma base (infinito per i verbi, singolare per sostantivi e singolare maschile per aggettivi), può apparire in più di un *synset*, poiché significati diversi di una stessa parola sono assegnati a *synset* differenti. Ogni *synset* è definito da una glossa che spiega il significato del concetto rappresentato e contiene le parole o le espressioni che sono generalmente utilizzate per esprimere quel concetto. Infine ad ogni *synset* è associato anche un numero che indica la frequenza con cui una parola appare con quello specifico significato. Infatti solitamente per una parola con più significati, ce ne sono alcuni che sono molto più comunemente usati di altri. I legami che uniscono i nodi della rete sono relazioni semantiche di vario tipo. Le più importanti sono le relazioni di iponimia e iperonimia per i sostantivi e aggettivi, troponimia ed iperonimia per i verbi. Un sostantivo è iperonimo di un altro se il suo significato è incluso o è una specificazione del significato dell'altro. Per esempio cacciavite è un iperonimo di attrezzo. Le relazioni di iponimia per i sostantivi e troponimia per i verbi indicano la relazione inversa.

Un'altra relazione importante è quella di olonimia e meronimia. Una parola è olonima di un'altra se quest'ultima fa parte della prima. Edificio è olonimo di finestra. Come prima, la relazione di meronimia è la relazione che va nel senso contrario ad olonimia.

Ci sono poi altri e numerosi tipi di legami semantici tra i *synset*. Uno dei più significativi è la relazione di *synset/geografia*, dove i *synset* che definiscono luoghi geografici sono legati tra loro in base alla loro inclusione geografica. Per esempio Europa è legata ad Italia che a sua volta include il *synset* Milano. Altri legami uniscono *synset* di tipo diverso, come link tra sostantivi e verbi di cui spesso sono soggetto, oppure aggettivi e sostantivi che generalmente sono legati nelle frasi.

Percorrendo la rete a partire da un dato concetto si può ricavare la distanza semantica che intercorre tra due parole e si definisce così una gerarchia, che verrà sfruttata proprio per la disambiguazione.

Uno dei metodi più comuni utilizzati per la disambiguazione del significato delle parole (WSD - Word Sense Disambiguation) a partire da una rete semantica è quello che utilizza la distanza concettuale tra i *synset*. La distanza concettuale è una misura che serve a quantificare la relazione che intercorre tra il senso delle parole ed il contesto in cui si trovano. Questa determina la vicinanza nello spazio dei significati tra coppie di parole, prendendo come riferimento una rete semantica gerarchica. La misura della distanza concettuale dovrebbe essere sensibile a:

- la lunghezza del percorso più breve che unisce i concetti considerati;
- la profondità nella rete gerarchica: i concetti che si trovano più in profondità nella rete gerarchica dovrebbero essere classificati come più vicini ed aver quindi più peso, poiché più specifici.
- la misura dovrebbe essere indipendente dal numero di concetti che si stanno misurando.

La densità concettuale CD del concetto c è definita dal rapporto tra i punti pesati della sotto-gerarchia di c e il numero di synset discendenti dal concetto c . Quindi h è l'altezza della sotto-gerarchia di c , m il numero di punti (synset) contenuti in essa delle parole trovate nel contesto e n_{hyp} il numero medio pesato di iponimi per nodo della sotto-gerarchia.

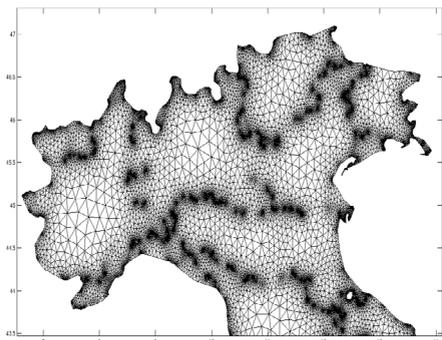


Figura 1 – Raffittimento della rete di contatti.

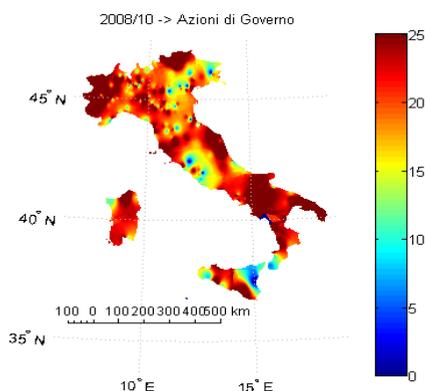
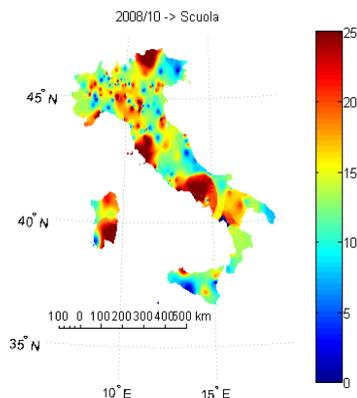
Utilizzando quindi la rete semantica e la densità concettuale è possibile stabilire l'accezione con cui viene utilizzata una parola all'interno di un testo e comprendere in modo automatico il significato di un testo. Attorno alla disambiguazione semantica nascono così dei software che a partire dal risultato della disambiguazione classificano i testi e ne estraggono informazioni. Uno di questi è COGITO, sviluppato dalla Expert System.

Partendo da dati estratti dall'analisi semantica geo-referenziati, è possibile rappresentare delle mappe geografiche tematiche che permettono di avere una visione immediata di come questi dati si distribuiscano sul territorio. Ciò diventa ancora più interessante se si inserisce anche la dimensione temporale. Per disegnare

le carte sono in gioco diverse variabili. Innanzitutto in base alla numerosità del campione di dati da rappresentare si dovrà scegliere la discretizzazione temporale in modo da avere sottoinsiemi statisticamente significativi. Infatti se la discretizzazione temporale viene fatta con intervalli di dimensione troppo ridotta, i dati potrebbero perfino essere insufficienti a riempire la superficie di una mappa. Al contrario, se si utilizzano intervalli troppo grandi, si perdono informazioni, poiché le mappe tendono ad uniformarsi, essendo il valore rappresentato il valore medio calcolato sull'intervallo temporale, e quindi diventano indipendenti dal tempo.

I dati geo-referenziati sono distribuiti in modo discreto. Le singole informazioni ma anche i concetti generali espressi nei testi, sono riconducibili per lo più a punti sulla mappa. I punti possono rappresentare la localizzazione approssimata del soggetto scrivente ma anche la localizzazione dell'oggetto cui il suo messaggio testuale fa riferimento. Per esempio può emergere che n individui concentrati in una determinata area si interessano a temi relativi alla scuola, alle azioni di governo, ecc. ma anche quali sono e dove sono i gli oggetti fisici cui i testi fanno riferimento es. quella scuola, quel Comune ecc. Ogni punto viene poi relazionato al più vicino e/o simile, in modo da ricostruire teoricamente la rete di tutte le più probabili relazioni spaziali (v. figura 1). A questo punto occorre riempire gli spazi tra un punto e l'altro, creando una superficie continua di valori. Anche in questo caso ci sono diverse possibilità di risoluzione del problema. Un primo metodo consiste nel definire delle aree per le quali viene calcolato un valore singolo e quindi tutti i dati presenti in quell'area verranno riassunti in uno, tipicamente il valore medio. Il metodo più idoneo individuato è stata l'interpolazione dei dati puntiformi. In questo caso si deve trovare un buon metodo di interpolazione. Nelle elaborazioni sono stati utilizzati metodi diversi di interpolazione: l'algoritmo nearest neighbour, il kriging, il metodo IDW (Inverse Distance Weighted interpolation), ma ve ne sono molti altri, come l'interpolazione polinomiale o l'interpolazione spline. Soprattutto l'ultima è molto utilizzata, ma nel nostro caso questo metodo introdurrebbe caratteristiche nella superficie non presenti nelle osservazioni, un effetto indesiderato.

Qui di seguito si riportano due esempi di sentiment mapping riferiti al mese di ottobre dell'anno 2008 e relativi al tema scuola ed azioni di governo, ove le tonalità di colore rappresentano le aree con i diversi gradi di attenzione all'argomento discusso sul WEB



Bibliografia

- Barber, C. B., Dobkin, D. P., and Huhdanpaa, H., (Dec.1996), *The quickhull algorithm for convex hulls*. ACM Transactions on Mathematical Software 22, 4, 469–483.
- Cipriani L. (2008), *Processamento automatico dei testi: tecnologie, algoritmi matematici a supporto dell'analisi e rappresentazione dello spazio semantico*, Tesi di Laurea Specialistica, AA 2008/'09, Politecnico di Milano, Relatore: Prof. Luigi Mussio, Correlatore: Prof. Mauro Preda
- ExpertSystem. *Cogito studio discover, user manual*. Version: CSD 2.7, CRules 11.15, Semantic Disambiguator 11.1 .
- Dematteis G., Ferlaino F. (2003), *Il Mondo e i Luoghi: geografie delle identità e del cambiamento*. Torino, Ires Piemonte
- Farinelli F. (2009), *Crisi della ragione cartografica*, Einaudi, Torino
- Farinelli F. (2006), *Il globo moderno, copia della mappa*, intervista a cura di Roberto Ciccarelli
- Giorda C. (2000), *Cybergeografia. Estensione, rappresentazione e percezione dello spazio nell'epoca dell'informazione*, Torino, Tirrenia Stampatori, 2000.
- Lévy P. (1996), *L'intelligenza collettiva: per un antropologia del cyberspazio*, Feltrinelli, Milano
- Ong, W.J. (1986), *Oralità e scrittura*, Milano, il Mulino
- Turco A. (1988), *Verso una teoria geografica della complessità*, Unicopli, Milano
- Gottmann J. (1966), *Essais sur l'aménagement de l'espace habité*, Mouton, Paris