

Luca Manassero

## SENTIMENTO, ANALISI ED ELEZIONI

### **Abstract**

*Since search engines and social networks have started being widely used by an enormous amount of users, advertising has always been their primary business model: presenting the “right” ad to the right user, especially focusing on her interests. We assisted therefore to an important and pivotal change from the generically focused traditional advertising industry to a highly specific way to match ads and users, thanks to personal data the same users convey for free to search engines and social networks. Advanced sentiment analysis techniques allow these same service providers to target their users with extremely personal political contents, therefore influencing democracy at a level that should be carefully taken into account.*

Pochi mesi fa Facebook<sup>1</sup> compiva dieci anni: la piattaforma di comunicazione “sociale” più nota e utilizzata al mondo<sup>2</sup> raggiunge ormai una percentuale notevole della popolazione mondiale, provocando effetti di enorme interesse sulle opinioni del pubblico.

La comprensione dell’evoluzione dell’impatto degli strumenti digitali impone un breve esame delle loro strategie comuni.

L’insieme dei protocolli di comunicazione che diedero origine alla rete informatica nota come Internet negli anni ’90 rendono accessibile ai privati la posta elettronica, strumento solo vent’anni prima limitato a un numero ristretto di grandi aziende e di istituzioni accademiche (per lo più statunitensi). Il fenomeno si estende rapidamente al grande pubblico tramite un crescente numero di fornitori gratuiti del servizio, tra i quali spicca evidentemente Google con la propria GMail<sup>3</sup>.

Già otto anni prima di questa importante scelta strategica, la Google, Inc. (dal 2015 proprietà della Alphabet, Inc.) aveva dato vita al celeberrimo motore di ricerca: da allora in poi avrebbe perfezionato costantemente il proprio algoritmo di classificazione dei contenuti<sup>4</sup> sulla base di elementi sempre più sofisticati, anche allo scopo di non cadere

---

<sup>1</sup> Aperto al pubblico il 26 settembre 2006; cfr. <https://en.wikipedia.org/wiki/Facebook>.

<sup>2</sup> 1,86 miliardi di utenti mensili in aprile 2017; cfr. <https://zephoria.com/top-15-valuable-facebook-statistics/>.

<sup>3</sup> Il servizio GMail viene lanciato nel 2004.

<sup>4</sup> Nel linguaggio abitualmente utilizzato da chi opera sul web vengono definiti genericamente “contenuti” i testi, le immagini, le registrazioni video o audio, e la combinazione di tutti questi elementi,

preda delle attività di SEO (*search engine optimization*), che hanno spesso avuto come unico scopo la rimonta di alcuni siti web nella classificazione Google tramite l'uso di tecniche che mirano all'ottimizzazione del posizionamento in sé, a prescindere dalla qualità del sito stesso.

All'utente del motore Google non può sfuggire che l'attività stessa di ricerca di contenuti sul web, così come la classifica di siti in base alle chiavi di ricerca, offre alla stessa Google un'eccezionale opportunità: quella di decidere – in base alle parole acquistate dal proprio cliente<sup>5</sup> – quali indirizzi di siti web debbano essere mostrati per primi all'utente.

Negli anni Google ha perfezionato continuamente i propri algoritmi, così da offrire un servizio obiettivamente di enorme interesse, contemporaneamente collegato a un servizio pubblicitario tra i più innovativi: una pubblicità non più indirizzata genericamente a un settore demografico identificabile tramite la propria frequentazione di luoghi pubblici, l'acquisto di determinate pubblicazioni o la fruizione televisiva in specifici orari, ma tramite i propri interessi.

La creazione e la gestione di un efficientissimo sistema di email gratuito – Gmail – permette a Google (e ad altri fornitori analoghi) di identificare con maggiore precisione l'utenza che da tempo ha creato un proprio account sul motore di ricerca, proprio per poter usufruire del servizio in questione.

Con questo piccolo passo il motore di ricerca ha perciò iniziato a delineare un'identità digitale formata da alcuni dati personali estremamente significativi, uniti a un tracciamento capillare degli interessi che guidano le ricerche dell'utente.

Sempre più le pubblicità mirate possono far riferimento a un profilo digitale la cui precisione cresce in relazione a ogni singola attività di ricerca, rendendo il valore della pubblicità stessa enormemente superiore al consueto cartellone nell'espositore della Stazione Centrale o sulla strada di grande traffico, perché proprio la precisa conoscenza degli interessi, delle problematiche (si pensi alle innumerevoli ricerche di informazioni mediche che si svolgono quotidianamente sui motori di ricerca) e dei desideri dell'utenza affida al marketing digitale e alla derivante *marketing intelligence* un valore di svariati ordini di grandezza superiore al marketing e all'*advertising* tradizionali.

Per completare con un ulteriore esempio i passi coinvolti nella creazione di un profilo digitale possiamo citare i servizi di navigazione offerti da Google Maps, che guidandoci a destinazione permette un preciso tracciamento dei luoghi da noi abitualmente frequentati e delle stesse tempistiche di permanenza.

Ma lo scopo di questo breve esame non è mirato alla demonizzazione della Alphabet, Inc. o di altri analoghi fornitori, i cui servizi – apparentemente a costo zero per l'utenza – sono entrati nel tessuto più profondo dell'esperienza quotidiana di miliardi di utenti.

La pervasività crescente di Internet ha reso infatti possibile da oltre un decennio (il primo *smartphone* prodotto da Apple, Inc. viene commercializzato negli USA nel 2007) la fruizione del web tramite dispositivi che ci accompagnano letteralmente passo passo nel

---

pubblicati su un sito web. Il post su un blog, un vostro aggiornamento personale su Facebook, l'immagine di un luogo pubblicata su Instagram sono tutti "contenuti".

<sup>5</sup> Servizio noto come Google AdWords, cfr. <http://www.google.it/adwords/>.

quotidiano e che, attivati tramite un semplice *login*, accrescono ininterrottamente la precisione dell'identità digitale che ci accompagna ormai fino al termine della nostra esistenza, e ben oltre<sup>6</sup>. Gli *smartphones* sono in grado poi, se collegati a dispositivi a contatto con il nostro corpo, di raccogliere i nostri parametri vitali<sup>7</sup>, costruendo una vera e propria cartella clinica di base: servizio consultabile tramite le applicazioni installate sugli stessi *smartphones* e i cui dati sono, ovviamente, a relativa disposizione dei fornitori dei servizi, i cui principali gestori sono al momento Apple e Google. Microsoft segue a strettissimo giro, a partire dall'acquisizione della divisione di telefonia mobile della Nokia e conseguente produzione di dispositivi con marchio Microsoft e sistema operativo Microsoft Windows.

Dovrebbe a questo punto essere chiaro come da alcuni anni il mercato pubblicitario abbia subito un cambiamento epocale, con l'evoluzione relativamente improvvisa di grandi aziende che possono indirizzare un messaggio preciso, per così dire "tagliato su misura" per il singolo utente.

Nonostante però l'uso di un motore di ricerca sia molto frequente (e necessario: si provi a immaginare la navigazione attraverso oltre un miliardo di siti web senza l'aiuto di uno strumento di questo tipo), si può immaginare che le attività svolte su di esso abbiano a che fare con bisogni puntuali e non manifestino necessariamente i desideri e i sentimenti più personali. Cercare dove acquistare un nuovo attrezzo da giardino o dove si trovi la concessionaria auto più vicina può sicuramente fornire indicazioni preziose per la focalizzazione pubblicitaria, ma difficilmente potrà spingersi oltre.

Con la nascita dei servizi di *social networking* (dunque con la comparsa dei *social media*) l'utente inizia a raccontarsi sempre più intimamente: affida alla propria identità digitale i propri amici (spesso catalogati per comodità e privacy in familiari, amici e conoscenti), le fotografie dei momenti più svariati della propria esistenza (immagini che portano con sé la posizione geografica determinata tramite GPS, la data, l'ora e le persone presenti) e soprattutto il proprio umore (il *sentiment* del testo, nell'accezione abitualmente utilizzata dai tecnici dei *social media*), i propri gusti e le proprie preferenze in ogni possibile direzione.

Se Facebook incoraggia la pubblicazione da parte dell'utente di un diario digitale per la creazione e la gestione dei propri contenuti, la piattaforma Twitter – inventore dei famigerati *tweets* – nasce invece allo scopo di permettere il *microblogging*, e nei suoi ben noti 140 caratteri dà origine e fama agli *hashtags* (parole chiave identificabili dal segno '#' che le precede), offrendo la possibilità di esprimere e condividere in ogni momento della giornata una brevissima opinione su un evento, la segnalazione di un contenuto qualsiasi, o altro ritenuto significativo per la propria *audience*, i "seguaci" (traduzione mal suonante, ma corretta, dell'inglese "*followers*").

---

<sup>6</sup> La piattaforma Facebook permette agli eredi di trasformare la pagina personale del defunto in una pagina "in memoriam" che non consentirà ulteriori aggiornamenti, ma manterrà nel futuro le informazioni che già contiene.

<sup>7</sup> Gli *smartwatches* e i braccialetti per il *fitness* misurano la frequenza cardiaca, la temperatura, il movimento, la durata dei periodi di sonno e trasmettono le relative informazioni al nostro *smartphone*.

Tramite piattaforme che permettono e incoraggiano l'espressione – e soprattutto la condivisione – del “sentire” immediato si struttura anche una sempre più raffinata composizione dell'identità digitale, ora legata a molto più che semplici ricerche.

Si rende ora necessario introdurre brevissimamente un tema di estrema importanza: l'analisi automatizzata del testo, la sua parziale comprensione e – di conseguenza – l'analisi del sentire espresso da un breve testo, meglio nota tra gli esperti del settore come *sentiment analysis*<sup>8</sup>. Grazie alla crescente potenza delle risorse di elaborazione messe a disposizione da gruppi di *computer servers* accessibili a costi ridotti (la “nuvola”, il *cloud*) e tecnologie di analisi testuale automatizzata sempre più affidabili, la *sentiment analysis* è oggi largamente possibile in tempo reale, rivelando così – ad esempio – il sentire di gruppi numericamente molto significativi di utilizzatori Twitter a proposito di un argomento specifico e durante lo svolgersi di un evento.

Per concretizzare un discorso che può apparire aridamente tecnico, immaginiamo che un noto personaggio politico si stia indirizzando alla nazione tramite un discorso televisivo: è incredibilmente semplice, dati gli strumenti adeguati, seguire in tempo reale il formarsi di un'opinione riguardo la comunicazione in corso, e il tasso di condivisione con altri utenti.

Supponiamo che il personaggio politico possa dunque avere visibilità del formarsi di detta opinione e – data la durata del suo intervento – possa modificare e adattare il contenuto e il tono della comunicazione sulla base dei risultati che la *sentiment analysis* gli sta fornendo.

Si obietterà che Twitter è uno strumento il cui uso è limitato a una fascia ristretta della popolazione: se tuttavia estendiamo tale analisi in tempo reale ad altri strumenti di *social networking*, il campione statistico andrà allargandosi fino a ridurre il margine di incertezza in maniera molto, molto significativa.

Se una semplice soluzione commerciale di *sentiment analysis* permette di valutare in tempo reale l'umore di una frase e di catalogarlo tramite un minimo di cinque o sei valori, dal “fortemente negativo” al “fortemente positivo”, si tratterà perciò di alimentarla con il contenuto di uno o più *social networks* per ottenere uno strumento estremamente potente di valutazione reale del sentire del proprio pubblico, il quale spesso è tanto più emotivamente sincero quanto più ritiene che la conversazione si svolga in uno spazio parzialmente protetto.

Agli elementi di questa breve analisi è opportuno aggiungere che se alcune piattaforme (prima fra tutte Twitter) hanno fatto del pubblico accesso alle conversazioni una caratteristica saliente, rendendo dunque relativamente semplice la *sentiment analysis* delle conversazioni in relazione a un tema, altre piattaforme (Facebook, ad esempio) hanno compiuto scelte diverse e più complesse. Le pagine Facebook dedicate a un marchio sono infatti sempre pubbliche (e le opinioni in esse espresse dunque pubblicamente visibili e analizzabili), mentre le pagine personali degli utenti sono molto raramente pubblicamente accessibili, a meno che l'utente stesso non le dichiari esplicitamente tali.

---

<sup>8</sup> Cfr. [https://en.wikipedia.org/wiki/Sentiment\\_analysis](https://en.wikipedia.org/wiki/Sentiment_analysis).

Le conversazioni e le informazioni che in esse si svolgono non sono comunemente raggiungibili: da circa tre anni la Facebook, Inc. ha ritenuto più opportuno selezionare un ridottissimo numero di concessionari dei propri dati. Tali aziende partner hanno il compito di rendere anonimi i profili degli utenti Facebook, mantenendone però le caratteristiche demografiche salienti. Il risultato di tali aggregazioni anonime, ma demograficamente classificate, dei dati personali degli utenti Facebook (quegli utenti, ricordiamolo, che non hanno reso pubbliche le proprie conversazioni) può essere acquistato a costi variabili, sebbene decisamente non popolari, presso i suddetti concessionari<sup>9</sup>.

Se dunque nel caso succitato non sarà possibile analizzare direttamente le conversazioni del singolo utente, sarà però certamente lecito lo stesso procedimento su un utente anonimo, ma portatore delle caratteristiche salienti dell'utente originario. In pratica non si risalirà a dati che permettano di identificare (con nome e cognome) il singolo utente, ma si avrà comunque accesso ai dati consolidati della persona.

Diventa conseguentemente più semplice e concreto chiarire l'enorme valore della mole di dati raccolti tramite l'utilizzo dei profili digitali (in grandissima parte ormai derivanti da dispositivi mobili) e metterli in relazione ai dati disponibili in forma pubblica o anonima provenienti dai *social media*.

Qualunque marchio può, tramite l'accesso a questi dati, costruire un profilo estremamente preciso del proprio cliente e adattare il proprio messaggio non solo al gruppo demografico di interesse, ma al singolo cliente, a me e a Voi.

Se il marketing ha avuto bisogno di ben poco tempo per iniziare a sfruttare le opportunità di queste tecnologie – oggi comunemente raggruppate sotto l'etichetta di *marketing intelligence*<sup>10</sup> – la politica non è stata da meno.

Dalla nascita di queste possibilità le singole personalità politiche (anche nell'ottica del *personal branding*) hanno iniziato a commissionare rapporti mensili ad agenzie specializzate nell'elaborazione di dati provenienti da fonti del marketing digitale. Tali *reports* sono fonte diretta di ispirazione per la misurazione del messaggio politico e possono, se ulteriormente sofisticati, fornire risultati in tempo reale che confrontino una campagna di comunicazione con i suoi risultati diretti sui *social media*, tramite la crescita dei *followers*, l'aumento o la contrazione delle espressioni di interesse (i *like* della piattaforma Facebook) e la variazione delle espressioni di condivisione dei contenuti.

Le cosiddette “viralizzazioni” di una registrazione video o di un messaggio, episodi a tutti noti, sono ampiamente misurabili e riconducibili a cause analizzabili nel dettaglio e soprattutto riproducibili.

Il percorso fin qui riassunto ha tentato di mostrare come le tecnologie digitali abbiano reso possibile, tramite una costruzione sempre più precisa di profili digitali, un'esposizione straordinariamente mirata della propria utenza a messaggi pubblicitari sempre più ritagliati su misura.

---

<sup>9</sup> Per un esempio si veda la DataSift, Inc. e il proprio servizio Pylon, cfr. <http://datasift.com/products/pylon-for-facebook-topic-data/>.

<sup>10</sup> Cfr. [https://en.wikipedia.org/wiki/Marketing\\_intelligence](https://en.wikipedia.org/wiki/Marketing_intelligence).

La definitiva apertura di Facebook all'utilizzo pubblicitario – iniziata nel 2011<sup>11</sup> – guida la piattaforma verso capacità tecnicamente sempre più evolute di presentare messaggi altamente specifici al singolo utente, sulla base delle proprie preferenze, espresse tramite i contenuti che ha creato e condiviso.

In ambito strettamente politico, supponendo che siano disponibili gli importi estremamente significativi necessari per raggiungere un grandissimo numero di utenti, l'utilizzo e il controllo di questi strumenti può significare che messaggi anche tra loro differenti, ma provenienti dal medesimo personaggio pubblico, possono essere presentati a utenti diversi, parametrati sulla base del linguaggio e delle opinioni espresse dall'utente *target* sulla piattaforma Facebook.

Non si tratterà più dunque di partecipare unicamente a *talk shows* a sfondo politico od organizzare una campagna di comunicazione tramite mass media tradizionali: i risultati più efficaci sono ampiamente raggiungibili tramite campagne di comunicazione personalizzate sul singolo profilo digitale, aggiornate in tempo reale sulla base della *sentiment analysis* e misurate con estrema precisione, anche in concomitanza con campagne e risultati ottenuti dalla propria concorrenza che opera utilizzando strumenti analoghi, anch'essi misurabili.

I media statunitensi hanno ampiamente valutato l'andamento delle ultime elezioni presidenziali americane anche sulla base della potenza comunicativa espressa dai candidati tramite l'utilizzo delle tecnologie di *marketing intelligence*: la superiorità dello staff del Presidente Donald Trump<sup>12</sup> nello sfruttamento degli strumenti di *digital marketing* può essere considerata una delle ragioni della sua vittoria, come pure costituire un messaggio importante per avviare una riflessione al presente sui valori della democrazia nell'era della comunicazione digitale.

Il quotidiano “USA Today”, nel suo formato digitale, già nell'agosto 2016 forniva una dettagliata comparazione<sup>13</sup> dei risultati ottenuti sui social networks dai due candidati alla Presidenza USA: i dati mostravano chiaramente stili differenti di utilizzo, ma una netta supremazia di Donald Trump nel produrre un «flusso di coscienza che operava un richiamo diretto agli elettori di Donald Trump».

Mentre Hillary Clinton dominava chiaramente la scena legata ai video (la piattaforma YouTube assegnava sedici milioni di visualizzazioni a Hillary Clinton e otto milioni a Donald Trump), l'avversario primeggiava su tutte le altre piattaforme di tipologia *social*.

---

<sup>11</sup> Per una più precisa cronistoria della pubblicità in Facebook, aggiornata al 2013, si veda L. TONER, *The History of Facebook Advertising* (<https://blog.hubspot.com/marketing/history-facebook-adtips-slideshare>).

<sup>12</sup> Per un'ulteriore informazione sul confronto tra i due candidati nell'utilizzo dei social media: *Trump credits Facebook and Twitter with his dominant win over Clinton* (<https://thenextweb.com/socialmedia/2016/11/14/trump-win-facebook-twitter-clinton>) e I. LAPOWSKY, *Here's How Facebook Actually Won Trump the Presidency*, in “WIRED”, 15 novembre 2016 (<https://www.wired.com/2016/11/facebook-won-trump-election-not-just-fake-news/>).

<sup>13</sup> J. GRAHAM, *Trump vs. Clinton: how the rivals rank on Twitter, Facebook, more*, in “USA TODAY”, 4 agosto 2016 (<https://www.usatoday.com/story/tech/news/2016/08/04/trump-clinton-social-media-twitter-facebook-youtube-snapchat/87974630/>).

Le tecniche di *sentiment analysis* offrivano però agli esperti valutazioni sempre nettamente positive del candidato rivelatosi in conclusione vincitore.

Numerose misurazioni condotte in prima persona dall'autore tramite gli strumenti messi cortesemente a disposizione dalla tedesca SAP SE<sup>14</sup> hanno mostrato analoghi risultati durante i *talk shows* televisivi cui partecipavano noti esponenti del mondo politico italiano: risultava relativamente semplice intuire chi tra i partecipanti avrebbe avuto successo (o sperimentato un doloroso fallimento) in occasione del successivo appuntamento referendario.

Si può naturalmente obiettare che gli strumenti di *sentiment analysis* automatica possono essere applicati unicamente ai contenuti affidati dai molti utenti ai *social networks*, ma tali analisi diventano eccezionalmente significative proprio nei casi in cui i candidati stessi abbiano considerato questi stessi *social networks* come l'arena privilegiata per i propri messaggi a contenuto politico.

Appare dunque rilevante e urgente aprire un fronte di riflessione dedicato alle nuove e potenti capacità di influenzare il pensiero del singolo elettore offerte dagli strumenti brevemente esaminati e – a oltre settantacinque anni dall'uscita nelle sale del film *Quarto Potere* diretto da Orson Welles – non cessare di ripensare le condizioni stesse della democrazia, anche a partire da una concreta comprensione delle nuove possibilità qui brevemente delineate.

---

<sup>14</sup> Un esempio tra moltissimi: <https://experience.sap.com/skillup/feel-sentiment-analysis-cheap/>.