

Stroppa: analisi visuale su Instagram, ecco il futuro per le aziende e la politica



L'ex hacker, oggi security researcher, spiega il software che rivoluziona la comprensione di Internet. E come può essere decisivo anche nella lotta ai populismi o nell'antiterrorismo

Secondo i due più completi report sulle operazioni di disinformazione avvenuti attorno alle elezioni americane – ordinati dalla Commissione intelligence del Senato Usa e svolti dal Computational Propaganda Project dell'Università di Oxford (in collaborazione con la società di analisi Graphika), l'altro coordinato da Renee Di Resta e Jonathan Albright di Columbia University – un fatto è certo: non solo le interferenze russe nel voto che ha portato Donald Trump alla Casa Bianca sono state ingenti, ma sono avvenute (oltre che su Facebook) probabilmente ancora di più su Instagram. Facebook aveva comunicato che la campagna di disinformazione russa aveva raggiunto almeno 126 milioni di persone, e altri 20 milioni su Instagram, Ma i numeri per Instagram sembrano

essere stati gravemente sottostimati.

È Instagram, insomma, la nuova frontiera per chi vuole analizzare i dati di Internet. Non solo per quel che riguarda la disinformazione: anche in tantissime altre sfere, l'antiterrorismo, la pubblicità, la profilazione, le analisi dei trend di marketing per le aziende, l'anticontraffazione, lo studio dei comportamenti e delle abitudini. In realtà già da diversi anni le più importanti società di analisi avevano indicato come Instagram avesse superato Facebook nei livelli di engagement: ovvero le persone che fruiscono degli stessi contenuti su Facebook e Instagram, sul secondo generano più interazioni, nonostante Facebook abbia [quasi il doppio degli utenti registrati](#).

Tutto passa da Instagram, insomma. Che oggi vanta un miliardo di utenti attivi ogni mese e che, grazie alla funzione Stories, i contenuti che scompaiono dopo 24 ore, è diventato un diario sempre aggiornato di influencer, brand, politici, aziende e organizzazioni, tanto che alcune campagne di comunicazioni per grandi aziende hanno deciso di creare delle mini-serie destinate soltanto al format delle Stories. In Italia poche persone si occupano principalmente di Instagram e delle analisi complesse dei dati su questo social network: uno di questi, forse il maggiore, è Andrea Stroppa, ex hacker, security researcher, collaboratore del World Economic Forum di Davos, e mente di Ghost Data, un progetto realizzato assieme ad altri esperti provenienti da Russia, Italia e Stati Uniti. Ghost Data è un software che ha focalizzato la sua tecnologia e il suo sviluppo sull'analisi dei dati "a scomparsa", Ghost (come fantasma), e che grazie a diverse tecniche e tecnologie permette di estrarre informazioni fruibili da dati complessi. Stroppa ha accettato di parlare – per la prima volta così diffusamente -di questa nuova frontiera, big data a analisi di dati soprattutto in relazione al mondo dei video e delle immagini.

Possiamo spiegare perché l'analisi di dati testuali è superata

ormai da quella su video e immagini?

«Innanzitutto, quando sentiamo dire “big data” non stiamo parlando di altro che di una grande mole di dati. “Analizzare i dati” però è un’espressione ancora generica, anche in inglese “data analysis” non ci spiega fino in fondo il suo significato. Tuttavia questi termini sono sempre più utilizzati, e allora imparare il loro significato è fondamentale per capire dove sta andando il mondo e lo sviluppo tecnologico. In breve lo sviluppo di potenti algoritmi, sistemi informatici sempre più performanti e accessibili permettono di realizzare analisi che prima erano soltanto fantascienza. I video e le immagini consentono di capire con enorme precisione molte più cose dei testi perché al loro interno hanno una immensa vastità di informazioni».

Andiamo con ordine. Fino a oggi cos'è successo? Le analisi dei dati possono avere tanti clienti, istituzioni, partiti, aziende. Ma finora si sono occupate per lo più di testi, e di social network che avevano al centro i testi, come Facebook, o twitter. Questo ha portato con sé una certa limitazione di prospettive, è così?

«Le analisi web sono di certo analisi potentissime e diffuse perché contribuiscono a spiegare o anticipare fenomeni sociali, economici, politici, finanziari, geo-politici, tecnologici. La piattaforma per eccellenza in questi anni, per le analisi fatte sul web, è stata Twitter: l’analisi dei tweet, e la profilazione degli utenti e altre informazioni come la geolocalizzazione, i dispositivi utilizzati, eccetera, hanno permesso a ricercatori, accademici, aziende e organizzazioni di studiare le elezioni politiche, trend, analizzare i mercati finanziari e molto altro. Questo genere di analisi viene svolta da circa dieci anni ormai, e è stata poi allargata ad altre piattaforme con tecniche simili».

L'Italia com'è messa?

« Ci sono non poche realtà che analizzano testi, ma pochissime che analizzano video e immagini. Eppure le piattaforme cambiano, come cambiano gli scenari: ci sono molti più utenti collegati a internet con connessioni veloci, smartphone di ultima generazione capaci di realizzare contenuti multimediali di alta qualità, persone sempre più giovani, ma anche tantissimi ultra cinquantenni in grado di comprendere pienamente i nuovi strumenti tecnologici e tecnologie che permettono di analizzare in pochi minuti migliaia di immagini, video e suoni grazie all'intelligenza artificiale, o più specificatamente alla visual recognition. Google, Amazon, IBM, Microsoft e alcune startup – non tantissime, e decisamente poche in Italia – hanno sviluppato software potentissimi in grado, attraverso modelli, di estrarre informazioni incredibili».

Ecco, ma cosa ci possono dire queste informazioni? Perché è tanto utile capirle, e saperle analizzare?

«Perché ci dicono tante cose che possono diventare asset di conoscenza per aziende, istituzioni e via dicendo. Faccio alcuni esempi anche banali e diversi tra loro: quali sono i loghi più ripresi nelle strade dello shopping? Quali sono le categorie di utenti che condividono contenuti di violenza, droghe, nudo o armi? Quali sono i dati demografici (ci sono persone nere, bianche, di quale età stimata) dei partecipanti a un concerto di musica rock? Quali sono i ristoranti di tendenza tra i giovani a Singapore? Quali canzoni e serie tv guardano a Milano le ragazze universitarie? Quali sono le vacanze preferite dai parigini? Tutto questo rappresenta una nuova generazione di analisi che va oltre la semplice analisi di testi dei tweet, ma entra in una dimensione più potente e complessa, che offre informazioni straordinariamente precise. Non sono più i messaggi di utenti che ci dicono “oggi ho ascoltato Vasco Rossi”, oppure “in quel quartiere c'è molta violenza in strada”, o ancora “la mia marca preferita di scarpe è la Nike”. Sono i software e gli algoritmi che, da

soli, interpretano e analizzano i contenuti che le persone condividono spontaneamente. E ovviamente nello stesso video possiamo avere un mix di tutte queste informazioni».

Perché Instagram diventa così centrale, e perché voi analizzate soprattutto video e immagini? Voglio dire, in cosa è più ricco rispetto, poniamo, a twitter?

«Il vero cambiamento è avvenuto con Snapchat, che ha inventato un formato nuovo di contenuti multimediali: i contenuti a tempo. Ovvero immagini e video che scompaiono dopo un tempo massimo di 24 ore. Instagram, di proprietà di Facebook, quasi due anni fa ha intuito che quel modello era vincente e ha creato le Stories. Dopo 24 ore le immagini e video scompaiono: gli utenti si sentono più rilassati nell'usare questo social quasi in modo compulsivo, facendolo ormai quasi coincidere con il quotidiano. Instagram, al di là di ogni considerazione più filosofica che potremmo fare, fa centro e le Stories diventano in meno di due anni la funzionalità più utilizzata, che Facebook inserisce subito su tutti i suoi prodotti. Oggi i trend ci dicono che per chi fa inserzioni, l'interesse commerciale è sempre più rivolto a Instagram, anche a scapito di Facebook. Grazie a questa intuizione Instagram cresce in modo esponenziale e raggiunge un miliardo di utenti attivi al mese. I maggiori analisti tecnologici sia sui media parlano di Instagram come il futuro non solo di Facebook. Ma anche il ceo di Facebook Mark Zuckerberg l'ha confermato, a fine ottobre ha dichiarato che il futuro di Facebook sono i contenuti a tempo, e soprattutto che stanno lavorando a un modo per monetizzare questo formato».

Perché le aziende, e le istituzioni, dovrebbero essere così tanto interessate all'analisi dei dati visuali?

«Perché gli utenti attraverso le Stories raccontano ogni istante della loro vita quotidiana offrendo delle informazioni nuove, uniche, fresche, multimediali, altamente profilabili e genuine. Dove si trovano, cosa acquistano, cosa indossano,

cosa hanno mangiato, chi sono le persone che incontrano, come guidano l'automobile, che musica ascoltano, il loro umore, i comportamenti, le abitudini e molte altre informazioni. Un vero tesoro per chi analizza dati e è in grado di estrarre informazioni. Una nuova frontiera delle analisi dei dati».

Sembra che anche l'intelligence possa essere rivoluzionata, da queste analisi: l'antiterrorismo, per esempio. È così?

«Faccio un esempio su un lavoro abbastanza fortunato che ho fatto anni fa. Associated Press a settembre 2017 scrisse, sulla base di un mio report, che l'Islamic State utilizzava la funzione Stories di Instagram per comunicare e condividere messaggi di propaganda. Era la prima volta al mondo che veniva documentato attraverso un report che lo Stato Islamico utilizzava il social dei filtri per eccellenza per comunicare messaggi in tutto il mondo con sostenitori e lupi solitari. Fino ad allora soltanto Twitter e Telegram erano conosciuti dal pubblico come mezzo favorito dal sedicente stato islamico. Molte delle Stories contenevano anche dati di geolocalizzazione, altre degli screenshot degli smartphone. Potete immaginare quanto può essere utile per un analista di intelligence avere dati di geolocalizzazione o vedere che operatore utilizza, quanto campo ha, di quale connessione internet dispone, quale sistema operativo e app sono installate sul cellulare di un simpatizzante o affiliato a un'organizzazione terroristica».

Il vostro team fu contattato anche da esperti di intelligence, in quel caso?

«Sì, è successo».

Cos'è, esattamente, "Ghost Data"?

«È un software in grado di profilare, estrarre e analizzare le Instagram Stories e Live (oltre che ai semplici post di IG) utilizzando differenti tecniche: la network analysis, la data analysis e l'ausilio dell'intelligenza artificiale che

prendiamo dai migliori: Amazon Rekognition, IBM Watson, Google Vision, Microsoft Azure e della startup americana Clarifai. Integrando differenti conoscenze e tecniche lo scopo è generare report o costruire cluster di dati che possono realmente aiutare aziende e non rispondere a delle esigenze ben precise».

È mai successo che qualche azienda vi abbia proposto di vendere il progetto, di comprare il vostro software?

«Sì, in generale le aziende del settore assicurativo, per esempio, sono molto interessate a comprendere i comportamenti dei loro clienti o potenziali. Non vogliono sapere nello specifico il comportamento del singolo, vogliono avere una panoramica più chiara su gruppi di persone».

Qual è la situazione in Italia e fuori?

«Per rispondere si può pensare al libro del cinese Kai-Fu Lee dal titolo AI Superpowers: China, Silicon Valley, and the New World Order. L'autore, un esperto del settore e anche un investitore in società di intelligenza artificiale, racconta come la Cina e gli Stati Uniti stanno costruendo sistemi di intelligenza artificiale straordinariamente avanzati e che soprattutto la Cina sta facendo dei passi da gigante soprattutto negli ultimi anni. Gli Stati Uniti ovviamente rispondono con le loro super università o i super laboratori delle aziende tech, come quello di Google, partito però dall'acquisizione di una società inglese alcuni anni fa. In Italia ci sono accademici di altissimo profilo nel settore, ma ho l'impressione ci siano scarse risorse finanziarie – per tutta una serie di ragioni – per avviare progetti che possono competere con i più bravi»