



Matematica, big data e opacità

Intervista a Paul-Olivier Dehaye, co-fondatore di PersonalData.IO

Poco più di un anno fa le inchieste giornalistiche del *Guardian*, del *New York Times* e dell'*Observer* richiamavano l'attenzione su una società di consulenza fino a quel momento poco o per nulla conosciuta: Cambridge Analytica. Attiva nel campo della consulenza in comunicazione, la società britannica fondata nel 2013 prometteva ai propri clienti di influenzare il comportamento di utenti, elettori o consumatori ricorrendo a un metodo che combinava il *data mining* e la psicomетria. Specializzata nel ricavare profili psicologici dai dati generati dagli utenti dei *social network*, la società si è trovata al centro di uno scandalo che ha coinvolto Facebook, la campagna elettorale del presidente americano Donald Trump e il voto sulla Brexit.

La metodologia di analisi di Cambridge Analytica è stata infatti messa in campo per generare messaggi, *banner* e annunci mirati destinati a influenzare l'elettorato britannico e americano. Finalità discutibili, come discutibile è il procedimento che ha permesso a Cambridge Analytica di accumulare la massa di dati necessaria a sviluppare e rendere operativi i propri algoritmi psicomетrici. Ricorrendo a un'applicazione apparentemente innocua – *ThisIsYourDigitalLife* – la società di comunicazione inglese è in effetti riuscita ad acquisire i dati personali e di navigazione (like, pagine seguite, interessi, profilo psicologico ottenuto attraverso un test validato) di 270 mila utenti di Facebook che avevano volontariamente scaricato l'app. Facendo poi leva sulle policies molto permissive in materia di utilizzo e protezione dei dati del social network, Cambridge Analytica aveva ottenuto anche informazioni sugli 'amici' degli utenti di *ThisIsYourDigitalLife*, senza che questi ultimi fossero al corrente dell'operazione, portando così il numero di profili coinvolti a 50 milioni (lievitati poi, dopo i primi accertamenti, a 87 milioni).

Senza voler approfondire le implicazioni legali o politiche, il caso Cambridge Analytica è significativo in quanto ha messo in luce problemi e contraddizioni della nostra relazione con il mondo digitale e con le tecnologie informatiche: dalla capacità predittiva dei metodi di analisi che ricorrono ad algoritmi e all'intelligenza artificiale, alla capacità dei *social network* di influenzare i nostri comportamenti e le nostre decisioni; dalle potenzialità offerte dalla rete e dai dispositivi mobili, alle possibili derive nel campo della protezione dei dati e della privacy. Seguendo il filo rosso di questo numero dedicato all'uso consapevole dei media e delle tecnologie, la redazione di Scuola ticinese si è rivolta a Paul-Olivier Dehaye. Matematico di formazione (ha studiato a Stanford, è stato ricercatore a Oxford e ha insegnato a Zurigo) Dehaye è fondatore di PersonalData.IO, un'organizzazione no-profit attiva nel campo dei diritti digitali e in particolare della protezione dei dati.

Matematico di formazione e attivista per la protezione dei dati per vocazione. Come si spiega il legame tra questi due centri di interesse a prima vista lontani l'uno dall'altro?

Ad avermi spinto a intraprendere gli studi in matematica è stato uno spiccato interesse per quegli ambiti che riguardano la gestione dell'informazione, quali la teoria dei numeri o la crittografia. Ad appassionarmi però non sono stati i soli aspetti teorici o tecnici. Le implicazioni sociali ed etiche della disciplina mi hanno infatti da sempre interpellato. Credo che oggi i matematici debbano assumersi notevoli responsabilità. Pensate ad esempio alla crisi finanziaria del 2008: i limiti, gli insuccessi e gli eccessi dei modelli matematici applicati ai mercati erano ben noti tra gli specialisti. In pochi però hanno voluto esporsi o assumersi delle responsabilità una volta scoppiata la crisi. È un'attitudine questa che si applica anche a tutte quelle tecniche di *data mining*, *machine learning* o di intelligenza artificiale basate sugli algoritmi. Sono descritte e vendute come fantastiche, anche se in realtà applicano principi matematici piuttosto semplici, sono spesso piene di errori e sono totalmente opache per la maggior parte di chi le utilizza. Chi produce innovazione scientifica e tecnologica non può quindi astenersi dall'interrogarsi criticamente sulle teorie e sugli strumenti che produce, ai quali deve sottoporre un'autocritica che tenga conto anche della dimensione 'umana' o 'umanistica'. Nel mio caso, questo esercizio di autocritica lo svolgo occupandomi di protezione dei dati personali.

Una delle azioni di attivista per la protezione dei dati da lei condotta si è occupata di una piattaforma di e-learning. Se ho ben capito, le è stato possibile dimostrare che l'uso dei dati raccolti dalla società che gestisce la piattaforma non era né trasparente né finalizzato a scopi unicamente didattici. Potrebbe spiegarci meglio il caso?

Il fatto che i dati raccolti dalla piattaforma attraverso i propri utenti fossero utilizzati anche a scopo commerciale era noto a tutti gli attori coinvolti ed esplicitamente dichiarato nei contratti. Il vero problema tuttavia si trovava altrove. All'epoca, nel 2014, insegnavo matematica in un'università che utilizzava questa piattaforma americana di e-learning (basata sui principi del *Massive Open Online Course*, MOOC). L'introduzione dello strumento digitale non era stata pre-

ceduta da nessun dibattito o approfondimento sulle condizioni d'uso delle informazioni raccolte sugli studenti – e sugli insegnanti – oppure sulle possibili ricadute, anche negative, in termini di protezione dei dati personali. Un modo di procedere che mi ha molto preoccupato, a mio parere per nulla corretto e di sicuro poco 'sano'... non solo per l'istituto universitario dove lavoravo, ma anche per l'insieme del sistema educativo. Per me era infatti chiaro che, una volta imboccata la strada di un utilizzo poco consapevole del nuovo strumento digitale di apprendimento, sarebbe stato molto difficile tornare indietro o cambiare direzione. Sulla scorta di queste considerazioni ho quindi deciso di proporre sulla piattaforma un corso online che tematizzasse diversi aspetti legati all'e-learning, dai modelli di business alla questione della protezione dei dati personali.

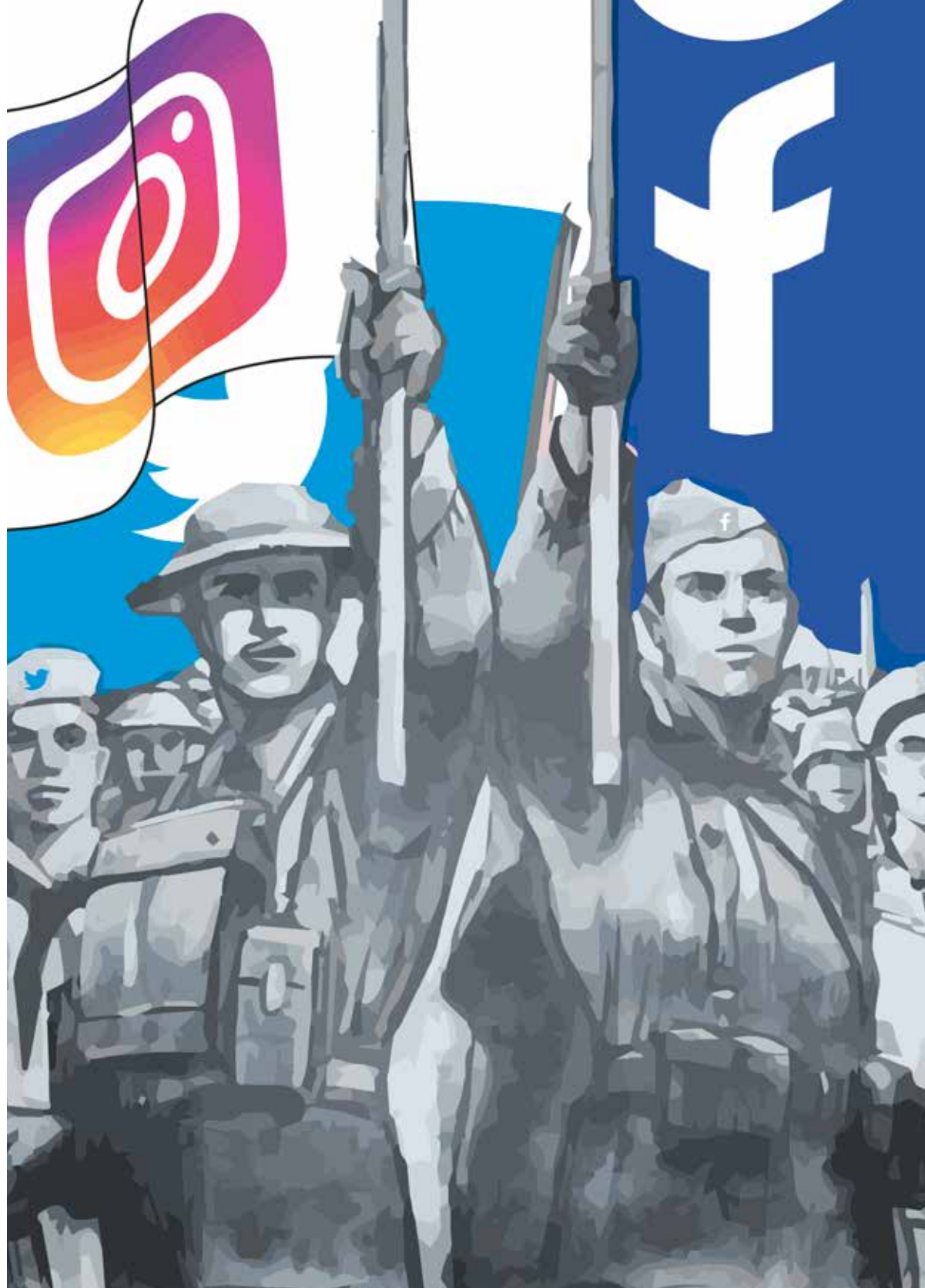
Avevo ad esempio l'intenzione di mostrare che la piattaforma MOOC eludeva le leggi europee sulla privacy. Il che mi ha causato non pochi problemi: i contenuti del mio corso sono stati censurati, sono stato citato dalla società, ho subito dei tentativi di diffamazione... e sono stato anche indagato dalla mia università. Quando però ho cercato di accedere ai miei dati personali per chiarire i miei intenti e difendermi, la società in questione si è rifiutata di rispettare i suoi obblighi contrattuali nonostante i miei ripetuti tentativi di mediazione e il ricorso alle vie legali.

Come detto, questo avveniva cinque anni fa e devo dire che il grado di consapevolezza attorno alla questione era ancora molto basso. Mi sono quindi ritrovato ad affrontare praticamente solo una riflessione critica che, ripeto, non coinvolgeva unicamente la piattaforma di e-learning, ma voleva spingere la comunità universitaria e la società civile a confrontarsi criticamente con questi nuovi strumenti educativi digitali.

Prima di entrare nel vivo degli aspetti educativi, una piccola provocazione: Who cares... in che misura un cittadino medio che non ha nulla da nascondere deve occuparsi dell'uso (commerciale o meno) dei suoi dati personali?

I suggerimenti di Google, YouTube o Booking sono molto comodi... Perché dunque dovremmo preoccuparci o addirittura mobilitarci?

La vera posta in gioco è la libertà. La nostra, ma anche quella delle generazioni future. Sono d'accordo con lei quando dice che le funzionalità appena evocate sono



Kiya Chiaratto
4° anno di grafica – CSIA

molto pratiche. Credo tuttavia che a noi spetti la responsabilità di decidere se le attuali condizioni di utilizzo dei *social network*, dei motori di ricerca, ecc. corrispondono a quelle che vorremmo dare ai nostri figli. Sono poi convinto che il fatto di trasmettere i nostri dati a dei giganti commerciali indebolisca, in un certo senso, la facoltà di disporre liberamente del nostro destino individuale e collettivo. Mi spingerei perfino ad affermare che è la nostra sovranità a essere messa in

discussione. È accettabile che i nostri dati siano registrati su dei server impiantati ai quattro angoli del pianeta? È accettabile che in sostanza non ci sia alcun modo di appropriarsene, così da verificare chi li usa e come sono utilizzati? In un paese come la Svizzera, molto attento all'esercizio dei diritti democratici, questa perdita di sovranità dovrebbe interrogare ogni cittadina e ogni cittadino. Anche perché accade rapidamente e impercettibilmente.

Ritornando agli aspetti educativi, e lasciando da parte il caso disfunzionale di cui parlavamo poco fa, in che misura big data, intelligenza artificiale, machine learning, ecc. possono essere integrati ai processi di insegnamento e di apprendimento? Quali opportunità offrono ai sistemi educativi e alla scuola?

La rete, l'informatica e gli strumenti digitali offrono senza dubbio molteplici opportunità. Il modo con il quale queste opportunità sono sfruttate in campo educativo è però molto rigido, direi quasi 'robotico' o 'meccanico': si ricorre allo strumento digitale per inquadrare gli utenti in un dato sistema, il che appare molto riduttivo rispetto alla 'persona' che sta al centro dei processi educativi. Ci si limita dunque a sviluppare strumenti per automatizzare la correzione di compiti o esercitazioni, per ottimizzare i flussi di lavoro o per profilare minuziosamente l'allievo ricorrendo ai suoi dati personali e ai suoi risultati scolastici. Il mondo digitale offre tuttavia potenzialità che possono contribuire a incoraggiare la creatività oppure lo spirito critico. Un'attitudine quest'ultima che dovrebbe poi essere rivolta verso lo stesso mondo digitale. Sono difatti un grande sostenitore della necessità di applicare una profonda riflessione critica, perfino introspettiva, a quelle che chiamiamo nuove tecnologie (anche se in realtà appartengono ormai alla nostra quotidianità).

Vorrei cambiare per un attimo prospettiva. Spesso sentiamo dire o leggiamo sui giornali che l'educazione è l'unico antidoto alle fake news. Cosa ne pensa di quest'affermazione?

È illusorio pensare che esista un'unica soluzione al problema delle *fake news*. Allo stesso tempo però, ritengo che la soluzione passi inevitabilmente dall'educazione. Bisogna capire che cosa c'è di nuovo nel fenomeno. Di sicuro, la novità non consiste nell'esistenza di informazioni false o nella circolazione di voci non confermate. Nuova invece è la sensazione di una perdita totale di controllo provata a fronte dello 'tsunami' di false informazioni che ci investe. Questa massa di *fake news* cancella la nostra capacità di adottare uno sguardo critico su quello che ci è detto e raccontato: come collettività siamo sempre più in difficoltà nell'identificare e smantellare le informazioni false. Accade perché siamo isolati, perché le *fake news* sono indirizzate a pubblici diversi utilizzando metodi

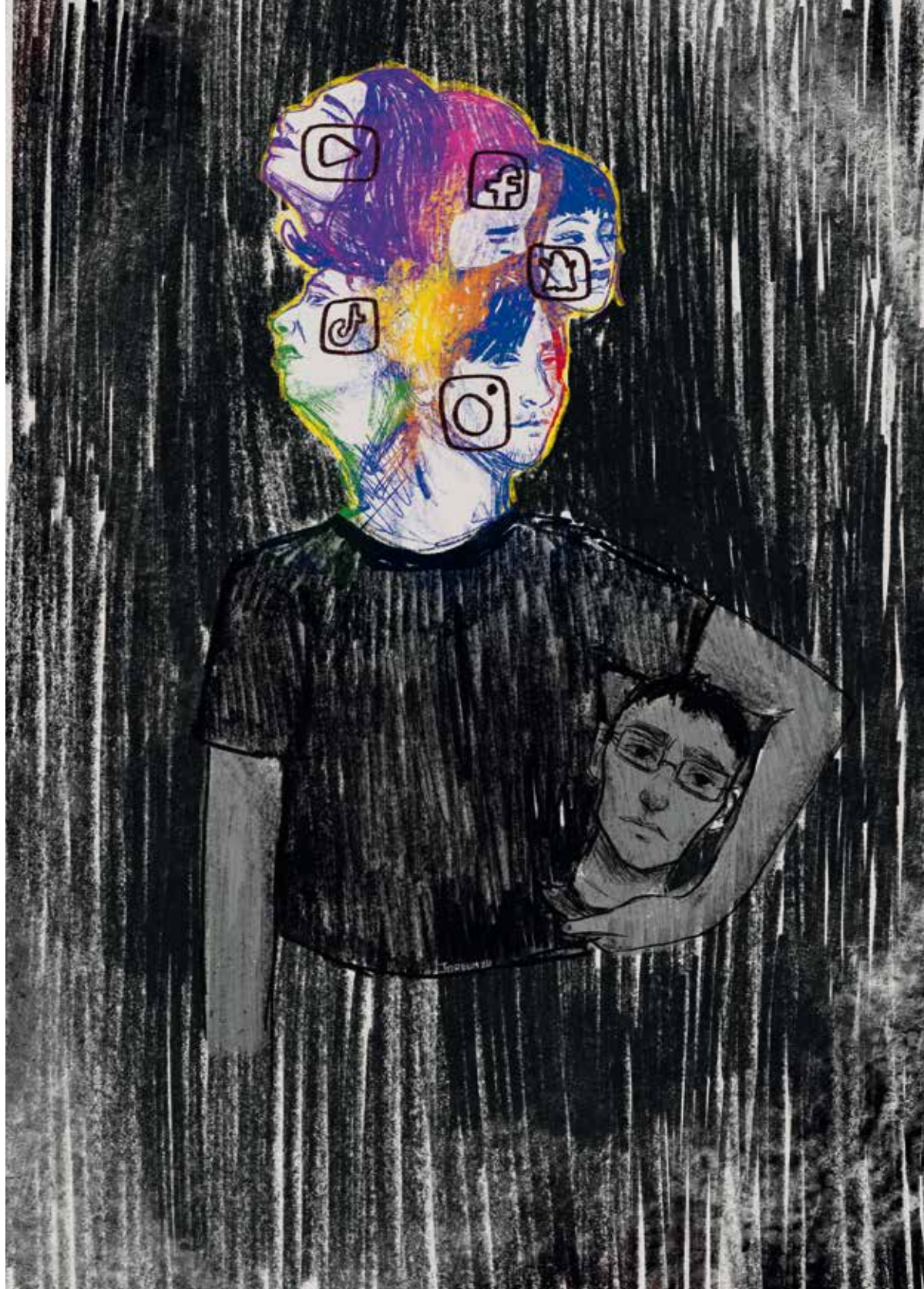
molto precisi, che ricorrono al *data mining* e alla psicologia. Per riallacciarmi alla sua domanda, credo che l'educazione sia fondamentale poiché ci permette di decostruire questo ecosistema poco o per nulla trasparente, di capirne le logiche interne e i meccanismi di funzionamento.

Un intervento educativo svolto attraverso la comprensione e la decostruzione di un ecosistema molto complesso e poco trasparente non rischia di diventare, per così dire, 'astratto' e quindi poco efficace?

Il rischio è reale. Capire il modo di funzionamento del sistema è necessario, ma non sufficiente. Occorre anche confrontarsi con dei casi concreti utilizzando come punto di entrata, ad esempio, i dati personali prodotti e trasmessi ai motori di ricerca piuttosto che ai *social network*. Bisogna mostrare quali sono le informazioni che generiamo volontariamente o involontariamente, mostrare come questi dati sono utilizzati, spiegare perché un certo contenuto ci appare sullo schermo di una pagina web senza che lo abbiamo richiesto. In altre parole: dobbiamo riuscire a portare in primo piano quanto avviene invece sotto la superficie del mondo digitale. Dobbiamo confrontarci concretamente con questa realtà opaca, così da rendere l'approccio educativo più efficace.

Si tratta allora di sviluppare nuove competenze in campo digitale? Oppure la scuola dispone già di strumenti educativi adeguati, derivati dalle sue finalità tradizionali?

Sono convinto che la vera sfida consista nell'adattare e nel trasporre nel mondo digitale principi, valori e pratiche educative che già esistono. Sono d'accordo con lei quando sembra suggerire che le finalità perseguite dalla scuola siano adeguate ad affrontare la sfida... penso allo spirito critico, alla conoscenza della storia, alla trasmissione dei valori democratici, ecc. La rete e le cosiddette nuove tecnologie agiscono tuttavia in un ambiente terribilmente opaco. Se siamo in grado di costruire dei percorsi educativi molto efficaci nel mondo offline, lo stesso non si può dire del mondo online. Questa incapacità è data proprio dalla mancanza di trasparenza. Possiamo sforzarci di spiegare e di educare, ma fino a quando non romperemo il muro di opacità che caratterizza il mondo digitale, avremo mani e piedi legati.



Giona Sciaroni
4° anno di grafica – CSIA

Potrebbe spiegarci meglio questa idea?

Prendo come esempio quello dei dati, con i quali sono confrontato quotidianamente. Un approccio educativo deve permetterci di scoprire cosa accade ai dati personali che affidiamo ai *social network* e ad altre piattaforme o servizi digitali. E il primo passo consiste nel potervi accedere, cosa che al momento non è ancora possibile. Ho di recente assistito una giornalista che ha voluto confrontarsi con i dati personali as-

sociati al suo profilo di Tinder. L'ho aiutata nel compito, non facile, di ottenere tutti o perlomeno una buona parte dei dati registrati dall'applicazione. La giornalista, da parte sua, si è confrontata con l'enorme massa di informazioni personali, spesso riguardanti la sfera intima e affettiva, consegnate – a volte inconsiamente – a Tinder (dati di movimento, *likes*, liste di contatti, ecc.). Il racconto giornalistico nato da questo esercizio concreto e autoriflessivo ci fornisce un esempio su

come avvicinarci con spirito critico alle raccolte di dati che ci concernono. Il che rappresenta già di per sé un atto educativo. Naturalmente il presupposto fondamentale è quella trasparenza che, come dicevo, per ora non esiste.

Adottando un orizzonte temporale di cinque-dieci anni, come immagina l'evoluzione nel campo dell'utilizzo e, rispettivamente, della protezione dei dati personali?

Inizio con l'osservare che negli ultimi anni sono stati fatti notevoli progressi. In materia di leggi e regolamentazioni l'Unione europea ha introdotto nel 2018 il nuovo regolamento europeo sulla protezione dei dati (GDPR) che influenzerà anche la Svizzera. La consapevolezza dell'opinione pubblica si è poi rafforzata, grazie anche ad alcuni scandali molto mediatizzati. A mio parere la società civile, i regolatori e i legislatori hanno quindi fatto dei passi avanti. Occorre tuttavia che questi attori trovino un modo per agire insieme e, quando lo avranno trovato, saranno molto potenti e potranno sostenere cittadine e cittadini nell'esercizio dei loro diritti in materia di protezione dei dati personali. Nella mia visione a medio termine, assisteremo a un'accelerazione di questa tendenza nei prossimi cinque anni, che produrrà effetti concreti nei successivi cinque, quando ci saremo riappropriati degli strumenti attualmente nelle mani delle società che gestiscono i motori di ricerca, i *social network* o altri servizi digitali.

Ha menzionato poco fa i 'grandi scandali'. È trascorso poco più di un anno dal caso Cambridge Analytica, di cui lei si è occupato da vicino. Senza entrare nel merito delle responsabilità penali e delle implicazioni politiche, cosa ci ha rivelato Cambridge Analytica?

Credo che nel 'dopo Cambridge Analytica' non sia più possibile qualificare le piattaforme digitali come inoffensive o considerarle come un semplice spazio di gioco. La situazione è ben più seria. Le rivelazioni sull'utilizzo abusivo e illegale di dati personali raccolti tramite applicazioni e *social network* ci hanno resi più ricettivi... soprattutto la politica lo è diventata. I politici hanno preso coscienza dei rischi in termini di perdita di controllo e di sovranità. Certamente non tutto è risolto, anzi. Cambridge Analytica ha però aperto una porta che ci permetterà di fare notevoli passi avanti.