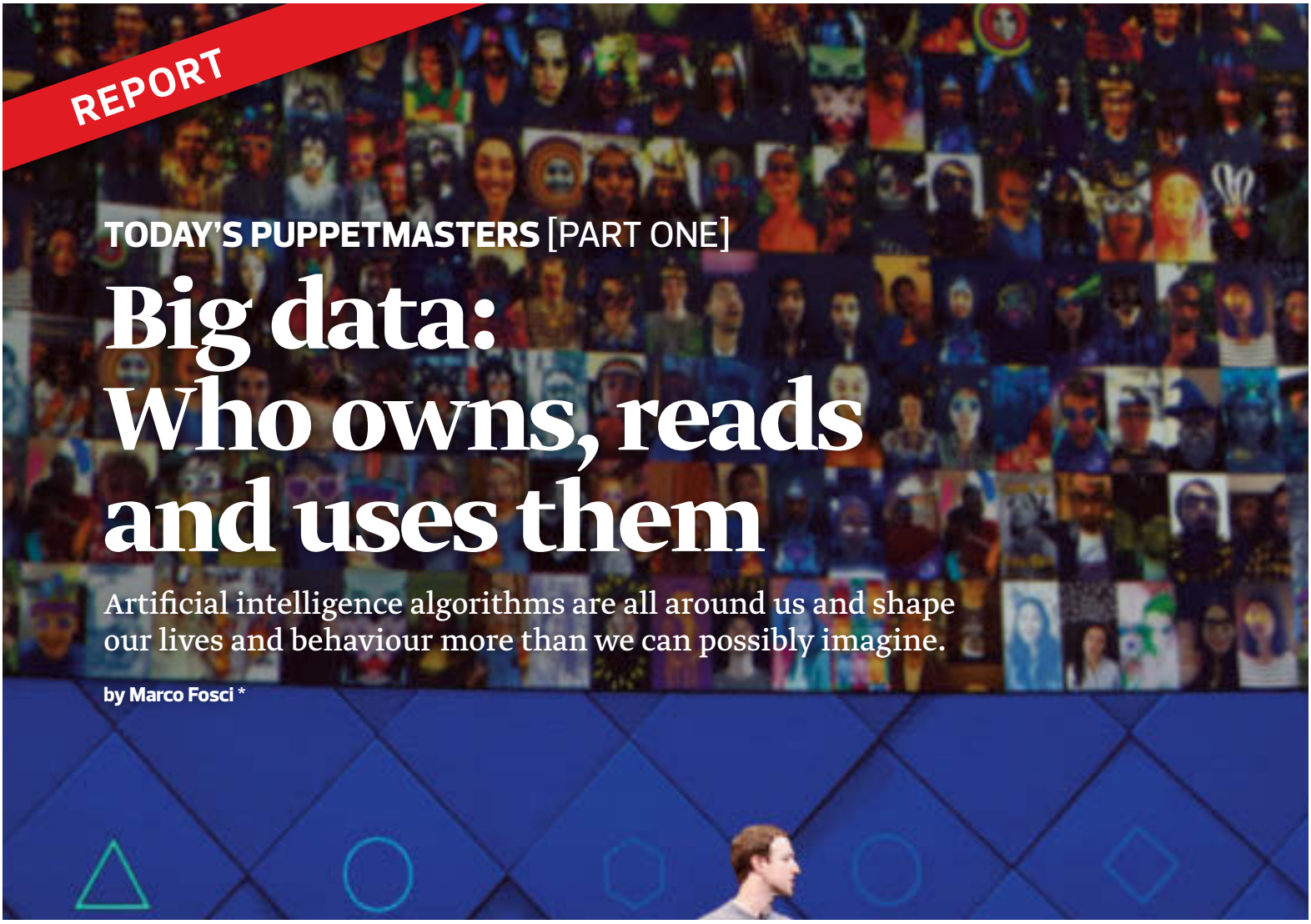
REPORT

TODAY'S PUPPETMASTERS [PART ONE]

Big data: Who owns, reads and uses them

Artificial intelligence algorithms are all around us and shape our lives and behaviour more than we can possibly imagine.

by Marco Fosci *



A long article in the 14 December edition of *The New York Times Magazine* described the different stages of a technological revolution that's going on almost unnoticed. In recent months Google added an artificial neural network to Google Translate, the free service used by millions of people to convert text into over 100 different languages. As a result the system is now capable of digesting a document by reading and processing entire sentences at a time instead of just word sequences. The quality levels achieved are now such that the Japanese translation of some passages of Francis Scott Fitzgerald's novel *The Great Gatsby* were comparable to the human translation by writer Haruki Murakami. *The Times* article explained how this achievement was made possible thanks to the deployment of a special class of artificial intelligence (AI) algorithms called "deep learning".

We are talking about a form of machine-learning architecture based on theories that have been developed since the 1960s. This architecture attempts to mimic the human brain's neural networks and focus on the idea of learning by experience. Its practical applications have only become genuinely effective since 2010 thanks to the greater computational capacity of processors and extensive use of big data. Software deploying these algorithms is capable of absorbing a considerable amount of raw information and identifying patterns and repetitious structures that it then transforms into useful and functional knowledge. The neural network can therefore improve its learning capability as it receives more data. In the case of Translate, these data are the texts of speeches and documents along with their accurate translations.

Similar algorithms are at the core of the vocal recognition systems used by virtual assistants. Apple's Siri, Microsoft's Cortana and Amazon's Alexa have been trained to recognise modulations in the human voice and to respond to smartphone users' requests. The same technology allows Facebook to identify faces from the photographs we publish on our profiles; in the United States, where privacy restrictions are more relaxed than in Europe, social networks even provide suggestions as to the identity of the faces in the photographs. A report on the possible impact of AI compiled by the Obama administration indicated that



REUTERS/CONTRASTO/STEPHEN LAM

between 2011 and 2015, the capacity for image recognition by algorithms showed a drastic improvement, reducing its margin of error from an initial 26% down to just 3.5%; on the same scale, human capacity is 5%.

Robotics too seem to be moving ahead rapidly in the wake of these developments. Car manufacturers and technology firms are currently investing heavily in AI managed driving systems. Important milestones have been reached by companies such as Tesla, Waymo and Uber, all of which have been testing driverless cars on the roads of five US states. The cars are equipped with sensors that enable them to trigger an immediate response to external stimuli which include road signs, the presence of pedestrians, traffic, the road surface or weather conditions. By the end of the year Tesla aims to be producing a vehicle that under certain conditions will be completely autonomous.

But systems that operate without direct human intervention are also of interest to other sectors. For some time now, AI algorithms are being used by banks for trading, by insurance companies to calculate risk, by marketing and communications departments to keep up with (and in some cases forecast and influence) consumer choices, in pharmaceutics and scientific institutes to accelerate research processes, and also in education.

In November, the Humanitas Research Hospital on the outskirts of Milan announced the introduction of Watson, a virtual cognitive tutor for the institute's medical students. The IBM supercomputer will devise customised study plans, simulate case reports and provide comments and in-depth analysis based on the knowledge level of each student. In some clinics in the US and elsewhere, Watson is already being used to help doctors make diagnoses. One such case that made headlines in August 2016 involved Ayaco Yamashita, a 66-year-old patient at the Imsut Hospital in Tokyo. After analysing her medical records, Watson identified a genetic mutation and was thus able to diagnose her type of leukaemia and provide recommendations for her treatment. The computer was able to complete this complex task thanks to the millions of clinical studies fed to it during its learning phase. The medical consultation provided, making the due allowances, was comparable to that of an authority on oncological disease.

This capacity to support and in some cases replace human intervention poses important challenges to governments in the field of professional training and requalification. The aforementioned report by the US administration envisions an increase in productivity and the creation of new professional skills, the potential downside being a dramatic impact on employment and income inequality. "AI has the potential to eliminate or drive down wages of some jobs", the report noted, "especially low- and medium-skill jobs". One emblematic example is that of taxi drivers currently seeking protection from Uber. In a not too distant future, Uber will be running a worldwide fleet of driverless cars. Without corrective measures, taxi companies, and presumably haulage firms, will be forced to do away with human drivers and merge with other firms in order to acquire the means to invest in the new technologies.

If the transformation of the employment market is a point of concern, there are even more complex questions to be answered. A May 2016 article in *ProPublica* denounced racial discrimination against black people due to the widespread use of algorithms in US courts to evaluate the risks of repeat offending. Created by private companies, these algorithms produce reports by processing collections of vast amounts of data related to prison populations which might easily be biased or distorted. "They may exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society", Attorney General Eric Holder explained to a Federal Commission. As it turns out, judges have been consulting this technology before deciding whether to hand down a prison sentence or rehabilitation.

Facebook founder and CEO Mark Zuckerberg speaking at the annual software developers conference in San Jose, California. Currently the social network has approximately 1.94 billion monthly users.

A similar problem occurs with the abuse of predictive analysis in the financial sector. In this case, the problem is not the accuracy with which AI determines economic value, but rather the danger of managers feeling relieved of responsibility when following suggestions made by their computers.

These kinds of risk are outlined in an investigation of the algorithm age, published by the Pew Research Center in February. "Algorithms have the capability to shape individuals' decisions without them even knowing it, giving those who have control of the algorithms an unfair position of power", claimed journalist Mary K. Pratt, one of the 1,300 researchers, technology experts, information technology professionals and government managers interviewed for the study. The document voiced concerns about the increase in the digital divide, the risk of creating a society governed exclusively by logic (or the pursuit of profit), the possibility of individuals' choices being limited or, worse still, the danger of societies being manipulated. "What concerns me", explained philosopher Peter Levine, "is the ▸▸



REUTERS/CONTRASTO/STEPHEN LAM

Craig Federighi, Apple's senior vice president of software engineering, during a conference. In an artificial intelligence report drafted by the Obama administration, it was announced that thanks to algorithm developments the margin of error in image recognition has now dropped from 26% to 3.5%.

ability of governments and big companies to aggregate information and gain insight into individuals that they can use to influence those individuals in ways that are too subtle to be noticed or countered".

One example is fake news or the leaks artfully assembled in order to discredit or favour a particular political group. This issue has been discussed recently in relation to the Brexit referendum and the US presidential election campaign, but less has been said about robot profiles (commonly known as bots) that are employed in an effort to spread these stories. Such bots are actually automatic algorithms, controlled by a few individuals but clearly capable of influencing the judgement of millions of people. It is not yet clear if the leaks against En Marche!, spread in a clumsy attempt to influence the French presidential elections, were another instance of this phenomenon. Whether those behind the leaks were skulking in the shadows of the American far right or were Russian hackers, the leaks themselves should provoke serious reflection about the capacity of cyber criminals, or cyber spies, to use small armies of bots to amplify untruths and propaganda.

There is also a more general problem linked to the potential concentration of power provided by algorithms. What would the implications be, for example, if Facebook CEO Mark Zuckerberg decided to run for political office? This point was raised by the media following the election of Donald Trump and reformulates the question about the control of information. Facebook is extremely active in the development of AI, which it uses in particular for marketing purposes, employing automatic profiling that generates tailor-targeted content for its 1.94 billion monthly users.

Apple, Amazon, Baidu in China and the other multinationals that boast hundreds of millions of users do more or less the same thing. *CB Insights* pointed out in March that a kind of arms race is underway: 200 takeovers have taken place in the sector over the last six years and as many as 34 during the first quarter of 2017 alone. Almost all of these acquisitions have been made by large US tech companies, fully aware of the key role already played by algorithms.

Google is the most dynamic of these large firms and is flourishing through its holding Alphabet. The long list of acquired companies includes the following: in 2013 the neural network of Canadian DNN Research; in 2014 the British DeepMind specialized in machine learning; in 2016 the software company Apigee and API.AI experienced in developing robot profiles that use natural language; and lastly in March the predictive analysis platform Kaggle. Google's old motto "Don't be evil" seems to have gone by the wayside. Nowadays, the corridors of its Mountain View headquarters are ringing with the words: "AI first". e

✱ **Marco Fosci [Milan]** has worked for *Il Sole 24 Ore* and Class Editori. He was also editor-in-chief of the information website Corrispondenti.net.