The Big Read **Technology sector**

Data brokers: regulators try to rein in the 'privacy deathstars'

Companies that collect consumer information have operated in the shadows. But calls are growing for tougher rules

© FT montage / Tolga Akmen

Aliya Ram and **Madhumita Murgia** in London JANUARY 7 2019

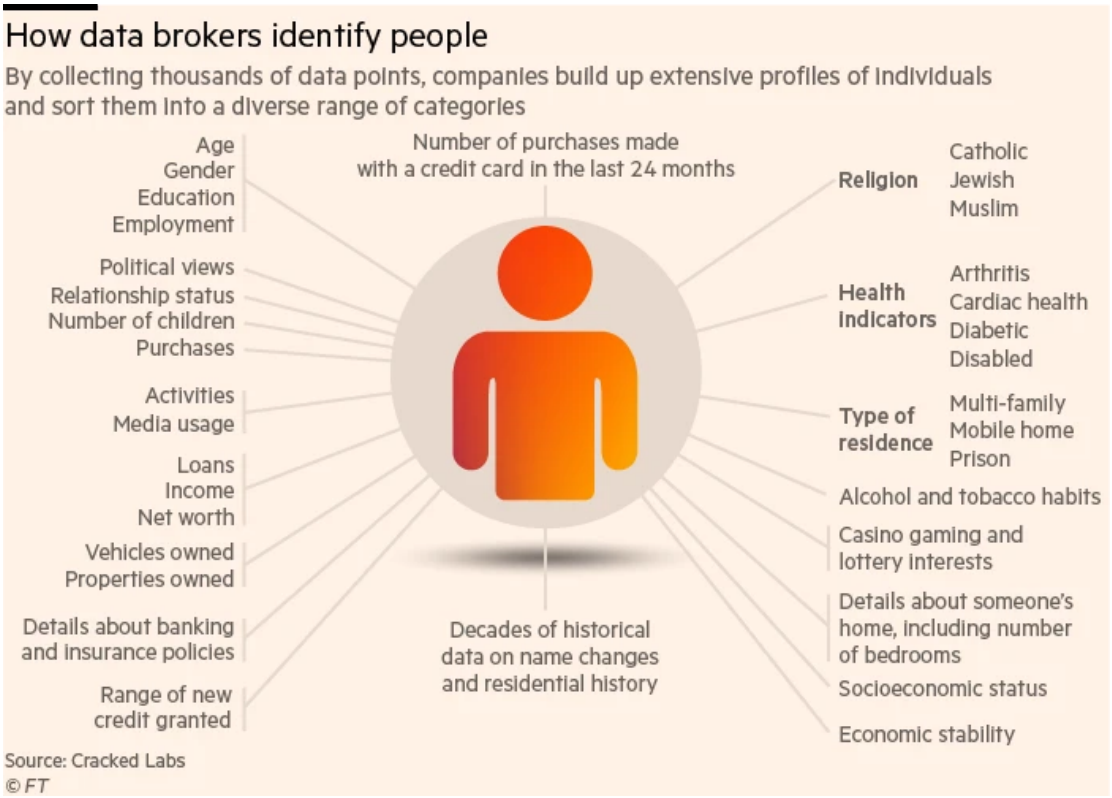
There are many personal details that Paul-Olivier Dehaye is willing to share online, but the behaviour of his bladder is not one of them. Yet when the Belgian privacy campaigner requested his data from advertising technology company Amobee, he found the business had predicted that on June 9 he was "likely to suffer from overactive bladder".

A legal representative for Amobee explained in an email to Mr Dehaye that the data had been licensed from The Weather Company, a business owned by technology group IBM. The Weather Company decided that based on hot weather conditions in Mr Dehaye's area he was likely to have an "overactive bladder" — and buy more drinks — on that day: "The overactive bladder [category] targets a mix of weather conditions that cause symptoms of overactive bladder to flare up, enabling advertisers to message when OAB is most likely to be top-of-mind for sufferers," the email said.

Few internet users will have heard of Amobee, a US company that sells advertising insights to the likes of Airbnb, Publicis and Lexus. But the business is just one of a constellation of adtech groups, data analytics firms and credit reference agencies that make up the rapidly growing data broking industry.

That industry is now very much in the regulatory spotlight in Europe. While the practices of its businesses have been investigated in the US for a number of years, regulators in Europe are now for the first time looking closely into their activities in the wake of the Cambridge Analytica data harvesting scandal last year and the introduction of the General Data Protection Regulation, Europe's new privacy law. Businesses that for years have operated largely in the shadows face the prospect this year of heightened scrutiny as public opinion shifts on questions of privacy.

The regulators have made it clear that they are deeply uneasy about the way the industry has been operating.



"They are all processing personal data, there is absolutely no doubt about that," says Mathias Moulin, director for the protection of rights and sanctions at the French data protection watchdog, CNIL. "They all try to say that it's anonymous to lower the pressure from the public, but that's not true. They know that and we know that."

While much of the attention last year focused on the use of data by tech groups such as Facebook, regulators and policymakers from the UK, France and Ireland who are examining the data-mining industry are turning their attention to the interlocking universe of lesser-known brokers that have also flourished as people spend more time online.

In November, Privacy International, the campaign group, asked European regulators to investigate seven brokers including software company Oracle after accusing them of breaking European data protection laws.

"[We are] concerned about whether or not their practices are compliant with the laws," says Elizabeth Denham, the UK's information commissioner. "We are looking at how they conduct their business and their general compliance with GDPR . . . certainly there is a dynamic tension between the way the businesses are conducted and the principles in the GDPR."

While much of the attention last year focused on the use of data by tech groups such as Facebook, regulators and policymakers are turning their attention to the interlocking universe of brokers © AFP

Data brokers mine a treasure trove of personal, locational and transactional data to paint a picture of an individual's life. Tastes in books or music, hobbies, dating preferences, political or religious leanings, and personality traits are all packaged and sold by data brokers to a range of industries, chiefly banks and insurers, retailers, telecoms, media companies and even governments. The European Commission forecasts the data market in Europe could be worth as much as €106.8bn by 2020.

"The explosive growth of online data has led to the emergence of the super data broker — the 'privacy deathstars', such as Oracle, Nielsen and Salesforce, that provide one-stop shopping for hundreds of different data points which can be added into a single person's file," says Jeffrey Chester, executive director of the Center for Digital Democracy based in Washington. "As a result, everyone now is invisibly attached to a living, breathing database that tracks their every move."

Over the past five years, the data broker industry expanded aggressively in what amounted to a virtual regulatory vacuum. The rise of internet-connected devices has fuelled an enhanced industry of "cross-device tracking" that matches people's data collected from across their smartphones, tablets, televisions and other connected devices. It can also connect people's behaviours in the real world with what they are doing online.

"The dream for the industry is to be able to connect the online and offline worlds to have a 360-degree view of the customer," says Gabriel Voisin, partner in international privacy and data protection at Bird & Bird, the law firm.

In November, Privacy International, the campaign group, asked European regulators to investigate seven brokers including software company Oracle after accusing them of breaking European data protection laws © Bloomberg

While brokers do not ever buy data directly from consumers, they are central to the data market. Even consumer data leaders such as Facebook, Google, Twitter and Snapchat have signed up as customers of brokers such as Acxiom, Oracle, Experian and others, because of the wealth and granularity of offline and cross-device data they have accumulated. For instance, if you went into a supermarket and bought baby wipes or nappies, that information could land you on a list for showing targeted ads to new parents.

According to IDC analyst Karsten Weide, growing demand should cause data vendor sales to more than triple to $10.1bn by 2022, compared with $3.1bn in 2017.

“The large platform giants, Google, Facebook and a few others — they are the major nodes in today’s personal data economy,” says Wolfie Christl, a researcher at Cracked Labs, a non-profit group based in Austria. “At the same time there is a kind of distributed surveillance economy . . . [that is] also collaborating with each other and large old-school data brokers like Acxiom.”

Facebook chief Mark Zuckerberg at a US Senate hearing. Consumer data leaders such as Facebook have signed up as customers of brokers such as Acxiom, Oracle, Experian and others, because of the wealth and granularity of offline and cross-device data they have accumulated © AFP

One of the largest data marketplaces is Oracle, the computer software company based in California. Oracle owns and works with more than 80 data brokers who funnel in an ocean of data from their own range of sources, including consumer shopping behaviour at brick-and-mortar stores, financial transactions, social media behaviours and demographic information. The company claims to sell data on more than 300m people globally, with 30,000 data attributes per individual, covering "over 80 per cent of the entire US internet population at your fingertips".

Richard Petley, head of Oracle in the UK, told the FT in August that there was "lots of opportunity" in data analytics as people spend more time online. Oracle declined to comment for this article.

Others, such as credit rating agency Experian and Acxiom use demographic, sociographic, lifestyle, cultural, mortgage and property data to categorise individuals. Experian uses the "Asian heritage" label for targeting "large extended families in neighbourhoods with a strong South Asian tradition", while Bank of Mum and Dad describes households where a grown-up child still lives at home.

Best known for its consumer credit scores, Experian's business model has changed significantly since it went public over a decade ago. Its data business comprises 55 per cent of its revenues, with the rest coming from other services such as identifying fraud or helping customers make credit decisions.

Under the company's "One Experian" transformation plan announced last year, it has sought to "connect different data sources", according to its annual report. By law, the company cannot sell data collected for credit decisions to advertising customers,but according to chief executive Brian Cassin, the businesses overlap "to a degree" as the company seeks "to build much more precise products" that use data to build more accurate tools for predicting credit worthiness, affordability and the likelihood of fraud.

Despite the sensitive nature of the data that brokers gather, acquiring the information they compile can be surprisingly easy. In October, Spanish researcher Joana Moll was able to buy the online dating profiles of 1m people for €136 from data broker USDate. The profiles of unsuspecting customers, garnered from online dating app Plenty Of Fish, included 5m photographs and details like their date of birth, zip code and gender as well as intimate information like sexuality, religion, marital status and whether they smoke, drink or have children. Plenty of Fish says it does not sell user data to USDate, and was unclear about the provenance of this data set.

"It's really easy, it was like buying a T-shirt on Amazon, and you can buy it anywhere in the world," Ms Moll says. "We acquired a second batch two weeks ago to see if anything had changed after GDPR but nothing had, we got the same number of profiles."

Alexander Nix, CEO of Cambridge Analytica © Reuters

Data brokers in the US have been scrutinised by lawmakers for more than two decades, but they have never been subject to any federal oversight. The last time the industry faced close inspection was in 2014 when the US Federal Trade Commission produced a 110-page report compiled over two years on nine of the biggest data brokers, including Acxiom and Datalogix, bought by Oracle at the end of that year. The commission recommended strongly that Congress introduce legislation to limit the reach of brokers, but versions of this draft legislation are still being kicked around on Capitol Hill today.

In Europe, pressure has built as investigations were launched into the industry. In the aftermath of the Cambridge Analytica scandal, the UK Information Commissioner's Office issued assessment notices to Acxiom, Data Locator Group and GB Group as well as Experian, Equifax and Callcredit, allowing it to carry out compulsory audits. CNIL, the French data protection authority, has carried out more than 50 inspections of data brokers and adtech companies in the past two years, including Paris-based Criteo.

Data brokers insist they comply with local laws by keeping consumers' identities anonymous; instead, they compile information on people's locations, shopping habits and browsing behaviour using pseudonymous identifiers or aggregated information. According to Amobee and IBM, for example, the overactive bladder category is not based on health information, but uses the weather in a particular area to predict whether people might be more likely to buy beer or water. Amobee added that it never used the data to send targeted ads.

IBM did not reply to requests about whether the predictions were based on location data but said the category "allows an advertiser to know when the weather in a certain location is potentially suitable for overactive bladder to occur". The Weather Company is currently facing a lawsuit from the city attorney of Los Angeles for "deceptively [using] its Weather channel app to amass its users' private, personal geolocation data". IBM has said The Weather Company has always been transparent about its use of location data.

Elizabeth Denham, the UK's information commissioner: 'We are looking at how [data brokers] conduct their business and their general compliance with GDPR' © Jon Super/FT

Critics say brokers are misleading people by claiming the data are truly anonymous. “None of these actors are processing anonymous data; they are processing personal data,” says Mr Moulin at the CNIL. “Data on location is very sensitive, with data on location you can identify a natural [real] person.”

Other regulators say businesses could fall foul of GDPR if sensitive data can be inferred from these audience categories. The European rules set higher standards for any data revealing categories such as “racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership”.

“We will be asking organisations to justify if they have [audience] names that suggest a special category [of] data,” says a senior official at one European data protection regulator, which is examining the industry.

Data brokers are already starting to make organisational changes. “When GDPR came in, people were forced to look at the legislation and realised the tech they were using was right at the boundary and limit of the existing [law],” says John Mitchison, head of policy and compliance at the Data & Marketing Association, the trade body for data-driven businesses.

“One of the most drastic things I’ve seen happen is all of these companies have radically reduced the number of third-party companies they will accept data from. You now need evidence that data were collected properly so they’ve weeded out a lot of suppliers that don’t meet those standards.”

CallCredit, one of the major credit reference agencies, which also had a big marketing data file, took a product off the market completely, and was subsequently taken over by TransUnion, a US company. Meanwhile Acxiom sold off its business, now called LiveRamp, that offers more controversial “identity resolution” services that link disparate atoms of data to create a profile of an individual, although it still accesses some of these services as a customer of LiveRamp.

Industry executives are hoping that these measures will fend off a much tougher assault on their business model from anxious regulators.

Additional reporting by Camilla Hodgson

Letters in response to this article:

GDPR should help expose violations of consent / From Stephen Wright, London, UK

Data harvesting in the days of direct mail / From Marco Bueninck, Mexico City, Mexico

---
