

February 4, 2015 6:46 pm

Big Data: Credit where credit's due

Tracy Alloway

Companies use technology to create 'proxy' credit profiles, raising concerns about the tactics



In late 2008, Kevin Johnson returned to his home in Atlanta, Georgia, after a honeymoon abroad to find the financial system imploding and a letter from his credit card company in his postbox.

American Express was cutting his credit limit from \$10,800 to \$3,800. The reason, according to the letter, was that Mr Johnson had been shopping at stores frequented by people deemed by the credit card company to have a poor repayment history.

[Sign up now](#)

FirstFT is our new essential daily email briefing of the best stories from across the web

For Mr Johnson, who prides himself on being a media and internet entrepreneur, the notion that he would be financially limited because of the supposed behaviour of his fellow shoppers did not sit well at all. At the time, says Mr Johnson, his FICO score — the standard measure of creditworthiness in the US — was about 760, solidly within good credit territory.

"It was a sort of wake-up call," Mr Johnson says. "The financial crisis created a wonderful opportunity for companies to apply some of these insidious algorithms to deny credit to people who are deserving of credit."

Mr Johnson, an African-American, went public with the letter, prompting a debate about whether such credit profiling was fair. Amex abandoned it, and a provision mandating further study of the practice was inserted into the Credit Card Act passed by President Barack Obama in 2009.

Six years later, financial companies have the option of using data-guzzling technologies that make the observation of shopping habits look downright primitive. A plethora of information gathered from social media, digital data brokers and online trails can be used to mathematically determine the creditworthiness of individuals, or to market products specifically targeted to them.

The degree to which such algorithms are utilised by mainstream banks and credit card companies is unclear, as are their inputs, calculations and the resulting scores. While many types of data-driven algorithms have been criticised for opacity and intrusiveness, the use of digital scorecards in finance raises additional issues of fairness. Using such information to make predictions about borrowers can, critics say, become self-fulfilling, hardening the lines between the wealthy and poor by denying credit to those who are already associated with not having access to it.

“You can get in a death spiral simply by making one wrong move, when algorithms amplify a bad data point and cause cascading effects,” says Frank Pasquale, a professor of law at University of Maryland and author of a book on algorithms called *The Black Box Society*.

Proponents of such technology argue that the ability to get a comprehensive view of prospective customers — based on their friends, their employer and even their exercise habits — can ultimately help expand the availability of credit to those who struggle to open bank accounts or obtain fair loans.

“When consumers are spending in a pattern that leads them to lose a little bit of control of their finances, are they able to get back on track and how? That’s the kind of data we can get from the big data world,” says Alex Sion, president of Moven, which provides debit accounts to customers who want to keep better track of their spending habits.

Both sides agree that new data-centric underwriting methods and algorithmic scoring pose legal questions at a time when the gap between rich and poor has become a hot-button political issue.

“It’s just such a big grey area right now,” says Galen Buckwalter, the scientist who helped develop the matching engine of dating website eHarmony and who is now at Payoff, an alternative lender. “The genie is out of the bottle, there’s no going back in terms of pulling this information or telling companies they can’t be analysing click times, even typing patterns and things like that.”

The 1935 map of Greater Atlanta in the National Archives comes in shades of blue, yellow and red. The hand-scrawled legend that accompanies it reads “light blue — best, dark blue — still desirable, yellow — definitely declining, red — hazardous”.

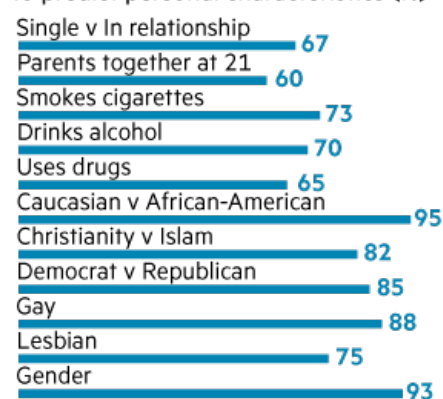
It is a stark reminder of a dark period in US history, when potential borrowers were classified not by their individual credit characteristics but by the neighbourhoods in which they lived.

The “redlined” neighbourhoods were often poor and dominated by a racial or ethnic group. In the case of Atlanta and other cities in the segregated South, the aim of the practice was often to keep African-Americans from moving into neighbourhoods dominated by whites.

The practice of redlining has since been outlawed with a slew of regulation, including fair lending and equal credit opportunity laws. But the worry is that redlining in the 21st century may not be done overtly with ink and paper but with the information that flows through computers and across the internet.

Social media data can be used to predict personal traits

Success rate of using Facebook ‘likes’ to predict personal characteristics (%)



Source: Kosinski, Stillwell and Graepel for the National Academy of Sciences of the USA

While US law forbids the discrimination of borrowers based on factors such as gender or race, parsing publicly available information on social networks such as Facebook and Twitter has been shown capable of accurately predicting everything from users’ political inclination to ethnicity and sexual orientation.

Chief among critics’ concerns is the ability to use new types of data and computerised algorithms to build “proxies” that do not overtly discriminate on the basis of factors such as race or gender, but may use correlated information to build an in-depth profile of a particular customer.

“Now you have data, you can predict anything,” says Max Gasner, founder of One Financial, a retail financial services start-up. “You can use your Facebook and Twitter data to tell what race you are, whether you’re gay or straight. You can predict a lot of stuff that’s illegal to use for lending and you can also more finely discriminate.”

Sifting through “clickstream” data or using “web beacons” that allow companies to track the movement of users across the internet can offer valuable clues about a person’s identity or socioeconomic status simply by observing how they access and navigate the web. For example, studies from the Pew Research Center and the Federal Reserve found

that blacks and Latinos were far more likely to access bank accounts using a mobile phone, giving an indication of racial identity simply by looking at how people browse the web.

“Because big data scores use undisclosed algorithms, it is impossible to analyse the algorithm for potential racial discriminatory impact,” the National Consumer Law Center wrote in a recent paper on big data.

In addition to the algorithms themselves, the laws that govern big data’s use in lending remain murky.

“When it comes to big data, there is no clear prohibition on using data for underwriting,” says Anand Raman, partner at law firm Skadden, Arps. “It can become a fair lending issue, however, if the use of that data results in disproportionate negative outcomes for members of a protected class.”

It is this fear of regulatory scrutiny that has left many of the big banks and credit card companies believed to be experimenting with new data reluctant to dive wholeheartedly into the world of non-traditional credit information.

“All the credit card companies have really been both attracted to, and afraid of, social data,” said a former data miner at one of the largest credit card companies. Financial groups that do overtly use big data to inform their underwriting may face “headline risk or could be construed as discriminatory”, in a manner similar to the fallout caused by Mr Johnson’s letter in 2008, he said.

Instead, use of non-traditional information has so far been publicly undertaken mostly by a crop of start-ups that say their goal is to use the smorgasbord of information now available to extend credit in a more efficient way, or to those lacking traditional credit information.

Peer-to-peer lenders that use the internet to directly connect borrowers with lenders are a prominent example; many use social media and other types of non-traditional data to augment their underwriting processes or prevent fraud — for instance, automatically verifying a loan applicant’s employment against their LinkedIn profile.

FT series

Democratising finance



An FT analysis of the impact of new technology and alternative financial models on the market

Further reading

Analysing social networks may allow lenders to guess not only a loan applicant’s race and socioeconomic status but also their comparative customer loyalty. Likewise, examining the way applicants click through web pages can give hints about character traits such as impulsiveness. The use of wearable technologies, which can track everything from exercise habits to heart rate, is also opening up an other realm of information for data-hungry lenders.

“If you are an organised, conscientious person who has good habits you’ll probably do well in credit,” said Andy Wen, vice-president of behavioural sciences at Payoff. “These things tend to correlate and if you can see organised behaviour in one area of life, you would think that would carry over to credit.”

Such companies are adamant that using behavioural data will expand credit to those who might otherwise be denied it. Moreover, many say using such information can remove subjectivity from an underwriting process that historically involved being sized up in person by a loan officer at a bricks-and-mortar bank branch.

“If you look at what big data is, it’s essentially the use of objective information,” says Mr Raman. “One can certainly argue that it’s better from a fair-lending risk perspective to use that sort of objective information than to rely on judgment or discretion.”

Some say big data and algorithm-driven consumer scoring may find its greatest use in marketing financial products, rather than underwriting them. Nielsen Prizm looks at postcodes to classify customers into categories ranging from “upper crust” to “low-rise living” — a group it describes as a “transient” world of “ethnically diverse” singles and single parents. The use of such precise targeting means that those lumped in with an undesirable group may never get a chance at a certain loan or financial product simply because they may never see them advertised.

The NCLC report found that credit products sold on the basis of non-traditional data-led processes offered annual percentage rates of 134 per cent to 748 per cent — triple-digit figures more typical of payday loans than credit cards.

“These technologies always get billed as, ‘we’ve found a new way to bring people with thin credit histories into the system’,” says Mr Pasquale. “I have suspicions about that. My sense is you’re just going to see a hardening of existing credit castes.”

Mr Sion at Moven says that using information from social networks and other non-traditional data sources may personalise algorithms that otherwise rely on more detached data to judge the behaviour of an individual borrower. “The social data brings it back to the person,” he says. “At the end of the day if you’re a bad risk, you’re a bad risk.”

In Atlanta, Mr Johnson says he is ultimately hopeful that new technology can be used to create opportunities for borrowers without exacerbating biases against them.

“There’s an upside and there’s a downside,” Mr Johnson says. “My experience is a very real story of how there is a downside and how, often, a lot of people don’t even know there is a downside.”

Algorithmic data: New scoring systems in need of legal definition

In the mid-1950s, William Fair and Earl Isaac began selling the first data-driven consumer scoring system to an investment company based in St Louis, Missouri. Their “FICO” scores — an abbreviation of the Fair Isaac Corporation name — would grow in tandem with the powers of computers to become ubiquitous in US lending.

By 1995 Freddie Mac and Fannie Mae, the US government’s housing finance agencies, endorsed the use of credit scores as part of the

mortgage underwriting process, embedding them in the fabric of the American financial system. By 2000, FICO scores were used in more than 75 per cent of home mortgage originations. In 2015, Fair Isaac says its scores are used in more than 90 per cent of lending decisions.

For decades, the scores were shrouded in secrecy and their details largely unknown to the borrowers whose credit they often dictated. It was not until 2004 that the US government required consumer reporting agencies to provide extensive credit score information at customers’ request.

In a report commissioned by the World Privacy Forum, Pam Dixon and Bob Gellman argue that new algorithm-driven consumer scores occupy a murky legal position similar to the one occupied by FICO scores in the 1950s.

“Data brokers, merchants, government entities and others can create or use a consumer score without notice to consumers,” they wrote.

“Laws governing credit scores do not typically extend protection to the new consumer scores.”

For many lenders, gathering non-traditional data is simply an attempt to augment or supplant traditional credit scoring models that still rely largely on historical information, which critics view as an imperfect indicator of creditworthiness. The biggest component of a FICO score is payment history. Such historical measures are considered ripe for disruption by new methodologies.

“The FICO score is creating false positives and false negatives,” says Mitch Jacobs, founder of OnDeck, an alternative lender that uses its own data-fuelled system to make small business loans. “A business might have been around for 10 years and have lots of customers and good cash flow but it looks like a bad credit.”

RELATED TOPICS [Big data](#) [United States of America](#)

Content recommended for you

Related articles

- | | |
|--|--|
| Integrated reporting can be the ally of business | ECP is a completely separate entity |
| Endeavour is the standout quality | Fox cuts 2015-16 forecast to 'mid-\$7bn' |
| US should be taking the higher ground on TPP | Bonds / buybacks: too popular to be smart |
| Economy is not robust but recovery is on track | FCC chief lays out new rules for web regulation |
| Getting borrowings off the balance sheet | Bitcoin continues to evolve into a worse version of the current system |

Printed from: <http://www.ft.com/cms/s/0/7933792e-a2e6-11e4-9c06-00144feab7de.html>

Print a single copy of this article for personal use. Contact us if you wish to print more to distribute to others.

© **THE FINANCIAL TIMES LTD 2015** FT and 'Financial Times' are trademarks of The Financial Times Ltd.



To read the full article please do one of the following:

Answer a question to continue reading this page

Have you ever visited a single product restaurant?

No

Yes

[Show me a different question](#)

or

Register for free and gain access to 8 articles per month and unlimited blogs

[Register for free](#)

If you already have an FT.com account please [sign in](#)

[Learn more – Privacy](#)