On the Web's Cutting Edge, Anonymity in Name Only

By EMILY STEEL and JULIA ANGWIN

You may not know a company called [x+1] Inc., but it may well know a lot about you.

From a single click on a web site, [x+1] correctly identified Carrie Isaac as a young Colorado Springs parent who lives on about $50,000 a year, shops at Wal-Mart and rents kids' videos. The company deduced that Paul Boulifard, a Nashville architect, is childless, likes to travel and buys used cars. And [x+1] determined that Thomas Burney, a Colorado building contractor, is a skier with a college degree and looks like he has good credit.

The company didn't get every detail correct. But its ability to make snap assessments of individuals is accurate enough that Capital One Financial Corp. uses [x+1]'s calculations to instantly decide which credit cards to show first-time visitors to its website.

In short: Websites are gaining the ability to decide whether or not you'd be a good customer, before you tell them a single thing about yourself.

The technology reaches beyond the personalization familiar on sites like Amazon.com, which uses its own in-house data on its customers to show them new items they might like.

In an interview with WSJ's Alan Murray, WPP CEO Sir Martin Sorrell conceded that advertisers must do better to inform customers about the tracking and mapping of online behavior. On the U.S. economy, he characterized the last 6-7 months as "America Bites Back" but wonders how long the recovery will last.
that data with records of home ownership, family income, marital status and favorite restaurants, among other things. Then, using statistical analysis, they start to make assumptions about the proclivities of individual Web surfers.

"We never don’t know anything about someone," says John Nardone, [x+1]'s chief executive.

Capital One says it doesn’t use the full array of [x+1]'s targeting technology, and it doesn’t prevent people from applying for any card they want. "While we suggest products that we believe will be of interest to our visitors, we do not limit their ability to easily explore all products available," spokeswoman Pam Girardo says.
investigation into online privacy has found that the analytical skill of data handlers like [x+1] is transforming the Internet into a place where people are becoming anonymous in name only. The findings offer an early glimpse of a new, personalized Internet where sites have the ability to adjust many things—look, content, prices—based on the kind of person they think you are.

New York-based Demdex Inc., for instance, helps websites build "behavioral data banks" that tap sources including online-browsing records, retail purchases and a database predicting a person's spot in a corporate hierarchy. It crunches the data to help retailers customize their sites to target the person they think is visiting.

"If we've identified a visitor as a midlife-crisis male," says Demdex CEO Randy Nicolau, a client, such as an auto retailer, can "give him a different experience than a young mother with a new family." The guy sees a red convertible, the mom a minivan.

The technology raises the prospect that different visitors to a website could see different prices as well. Price discrimination is generally legal, so long as it's not based on race, gender or geography, which can be deemed "redlining."

In financial services, fair-lending laws prohibit discrimination based on race, color, religion, national origin, gender, receipt of public assistance or marital status. The laws also require that borrowers have access to any data used to evaluate their creditworthiness.

But the law doesn't specifically bar using web-browsing history to make lending decisions. That means, in theory, a bank could deny a loan based on knowledge of the applicant's visits to, say, gambling sites. In such a case, however, the bank would be required to let the applicant see the browsing data and correct it if inaccurate.

Capital One says it doesn't use [x+1] or browsing history in lending decisions. Rather, it uses [x+1] to suggest products to individuals.

The regulators who monitor fair lending at the Federal Trade Commission say suggesting offers isn't illegal. But it could violate the law if the suggestions result in protected groups such as minorities being steered into paying higher credit-card rates despite having solid credit.

"Steering can be a law violation depending on how they do it," says Alice Hrdy, an assistant director at the FTC. "Credit decisions have to be based on the customer's creditworthiness."

Capital One spokeswoman Ms. Girardo says, "Our practices are fully compliant with banking regulations and privacy laws."

[x+1] says none of its credit-card services use gender, ethnicity or age data. It adds that the company doesn't
The idea of using data about website visitors’ offline lives was controversial in 1999 when [x+1] first started a website-personalization business. The FTC was investigating the privacy implications of online-advertising network DoubleClick Inc.’s acquisition of Abacus Direct, which tracked people's "offline" purchases at traditional retailers.

After a flood of negative publicity, DoubleClick agreed not to combine its online data with Abacus’s offline data. For years, DoubleClick’s experience deterred other companies from merging online and offline data.

[x+1] struggled to stay afloat. It cycled through six CEOs and three names, including Poindexter Systems (after a nerd scientist in "Felix the Cat" cartoons).

In 2008, Mr. Nardone took the helm just as things changed. Online ad spending rebounded and marketplaces for online data sprang up, letting companies like his tap data about people’s Web browsing. Traditional data brokers (mostly serving direct-mail and catalog companies) began making data available online, too, but with names stripped out to address privacy worries.

Mr. Nardone saw opportunity. He revived the company’s decade-old data-crunching patent for a "predictive optimization engine," now turbocharged with newly available data. "I discovered very quickly that we were going back to the original roots of the company,” he says.

He found a receptive audience for site customization in the credit-card business, nabbing Capital One. [x+1] says clients pay $30,000 to $200,000 a month for its technology.

"You don’t get information on everybody, but there are ways of doing analysis that you can fill out the gaps” says Ted Shergalis, [x+1]'s co-founder and chief strategy officer. "That is the whole science of this.”

Its technology works like this: A visitor lands on Capital One’s credit-card page, and [x+1] instantly scans the information passed between the person’s computer and the web page, which can be thousands of lines of code containing details on the user's computer. [x+1] also uses a new service from Digital Envoy Inc. that can determine the ZIP code where that computer is physically located. For some clients (but not Capital One), [x+1] also taps additional databases of web-browsing history.

Armed with its data, [x+1] taps consumer researcher Nielsen Co. to assign the visitor to one of 66 demographic groups.

In a fifth of a second, [x+1] says it can access and analyze thousands of pieces of information about a single user. It quickly scans for similar types of Capital One customers to make an educated guess about which credit cards to show the visitor.

To gauge the system's accuracy, the Journal asked eight people to visit the credit-card page of Capital One's site and note the credit cards they were shown. The Journal also analyzed the computer code that zipped back and forth between the testers’ computers and Capital One.

Separately, the Journal asked its testers to click on a custom website that [x+1] built to demonstrate its technology. After the testers clicked on that site, [x+1] described to the Journal what it knew about each person.

Throughout both of these processes, the testers didn’t reveal any personal information.

[x+1]'s assessments of the testers were generally accurate, though some specific details missed the mark. For instance, [x+1] correctly placed Ms. Isaac, the Colorado Springs mom, in a Nielsen demographic segment called "White Picket Fences." People in this group live in small cities, have a median household...
income of $53,901, are 25 to 44 years old with kids, work in white-collar or service jobs, generally own their own home, and have some college education.

All of those points were correct for Ms. Isaac—to her surprise. "They pinpointed my income more accurately than I remembered it," she says.

But the "White Picket Fence" category wasn't 100% accurate. It suggested Ms. Isaac might read People en Espanol, watch Toon Disney and drive a Nissan Frontier truck. In fact, she doesn't speak Spanish, doesn't subscribe to cable TV and doesn't drive a truck.

Nielsen says its segments are intended to provide a broad framework to help marketers understand their customers, rather than an exact template.

[x+1] says its analysis isn't meant to be pinpoint accurate, either. "It is just saying, 'Do I have better than 50-50 odds of guessing what this anonymous user is going to want?'" says Mr. Shergalis, the company co-founder.

The Journal also captured and analyzed 5,219 lines of code that passed between Ms. Isaac's computer and Capital One. That code contained some of the results of [x+1]'s analysis, which was generally accurate, putting her in "Colorado Springs" with a "midscale" income and saying she was either a college grad or had "some college."

As a result of the analysis, Capital One showed her some of its least generous cards, which it describes as being for people with "average" credit. The bank defines "average" as people who might not currently have a card, or whose credit limits are below $5,000 or who "may have been late" on a loan or card in the past six months.

Ms. Isaac says that category fits. She and her husband use only debit cards after using credit cards in the past.

Capital One's Ms. Girardo says: "Like every marketer, online and offline, we're making an educated guess about what we think consumers will like and they are free to choose another product of their liking."

[x+1] zeroed in on Mr. Boulifard's love of travel. The Nashville architect was shown only one card on the Capital One website: a "VentureOne Rewards" card, shown floating above a beach scene, with a headline: "Still searching? Get double miles with Venture."

Mr. Boulifard is a member of several frequent-flier clubs and is saving up for a Europe trip. He routinely shops for hotels online and uses his American Express card to rack up travel points. However, Capital One's suggestion didn't tempt him to switch. "I have 90,000 points on my American Express, so I'm going to Europe on that," he says.

With Teresa Britton, [x+1]'s algorithms bumped into the limits of their ability to capture some of the complexities of modern America. The company correctly pegged her as a Greensboro, N.C., resident and assigned her to the "Young Influentials" Nielsen segment, a group that lives in suburbs and earns about $50,000 a year.

That underestimates her income, Ms. Britton says. But she really bristled at Nielsen's suggestions that she might buy rap music or read Vibe, the hip-hop magazine. Ms. Britton is white, and her husband is black.

"I don't know if they somehow got me and my husband mixed," she says. That said, her husband doesn't...
read Vibe or buy much rap music. "That is so stereotypical," she says. [x+1] says its Nielsen data are from broad segments and "don't apply at the individual level."

The technology was better at sizing up Mr. Burney, the Colorado contractor who builds ski-vacation homes. He saw only one credit card, the Capital One Prestige Platinum, under a headline: "Our best rewards at a glance." The pitch included an initial 0% interest rate and no annual fee.

That card made sense, Mr. Burney says. Not only does he tend to charge a lot of expenses for his construction business, but "I have a wallet full of platinum credit cards," he says. "My credit is sparkling."

Based on Mr. Burney's visit to [x+1]'s test site, the company pegged him as a member of a Nielsen segment called "God's Country." People in that group live in small towns or rural areas, have median household income of $86,724, are 35 to 54 years old with no kids, work in management, mostly own their homes and are college graduates.

Mr. Burney is, in fact, a homeowner, a college grad and a manager, and has no kids. At 28, he's younger than predicted, and his income is less than predicted.

When he saw the 3,748 lines of code that passed in an instant between his computer and Capital One's website, Mr. Burney said: "There's a shocking amount of information there." Buried in the code were references to his income level ('uppermid'), education ('college%252b graduate') and his town ('avon').

In fact, [x+1]'s assessment of Mr. Burney's location and Nielsen demographic segment are specific enough that it comes extremely close to identifying him as an individual—that is, "de- anonymizing" him —according to Peter Eckersley, staff scientist at the Electronic Frontier Foundation, a privacy-advocacy group.

Mr. Eckersley does research in the field of de-anonymization, the mathematics of identifying individuals based on a few specific details from their life. In the jargon of the field, Mr. Eckersley says, all that's needed to uniquely identify one person is a total of 33 "bits" of information about him or her.

Calculating "bits" gets complex, as some facts about a person are more valuable—and thus have more "bits"—than others. ZIP codes and birthdates, for instance, are extremely valuable when zeroing in on individuals.

Bottom line: Mr. Eckersley determined Mr. Burney's location (the small town of Avon, Colo.) and his Nielsen demographic segment ("God's Country") together offered about 26.5 bits of information that could be used to identify Mr. Burney individually.

That's enough to narrow him down to one of just 64 or so people world-wide.

With one more piece of information about him, such as his age, Mr. Eckersley says, it's likely that Mr. Burney could be de- anonymized. "You're starting to look very close to identified."

Mr. Nardone of [x+1] acknowledges the possibility of de-anonymization, but says it isn't worth the effort: The company already has enough information to sell. "It would be a massive undertaking," he says, "and it is hard enough to make money."
Corrections & Amplifications

Online-marketing company[x+1] Inc. helps some companies estimate the "lifetime value" of their customers. An earlier version of this article incorrectly said [x+1] provided this service to Capital One Financial Corp. as part of [x+1]'s overall website-customization services to the bank.

Write to Emily Steel at emily.steel@wsj.com and Julia Angwin at julia.angwin@wsj.com