

Smile! Marketing Firms Are Mining Your Selfies

Photo-Sharing Sites Are Being Scanned to Find Brands, Target Ads

Scanning for Clues

Ditto Labs scans photos on sites like Instagram to glean insights for advertisers.



FACES

Human faces can give context. If a logo appears above a face—such as Smith ski goggles in this picture—it can indicate the person is wearing apparel. Smiles help advertisers understand sentiment.



PRODUCTS

Users who post images of things like beer or ski equipment may be flagged for those interests. Marketers also look for correlations between interests, such as which beverages people drink while eating Kraft Macaroni and Cheese.



CLOTHING

Ditto has started classifying objects like purses or sunglasses and can sometimes detect fabrics or patterns in clothing.



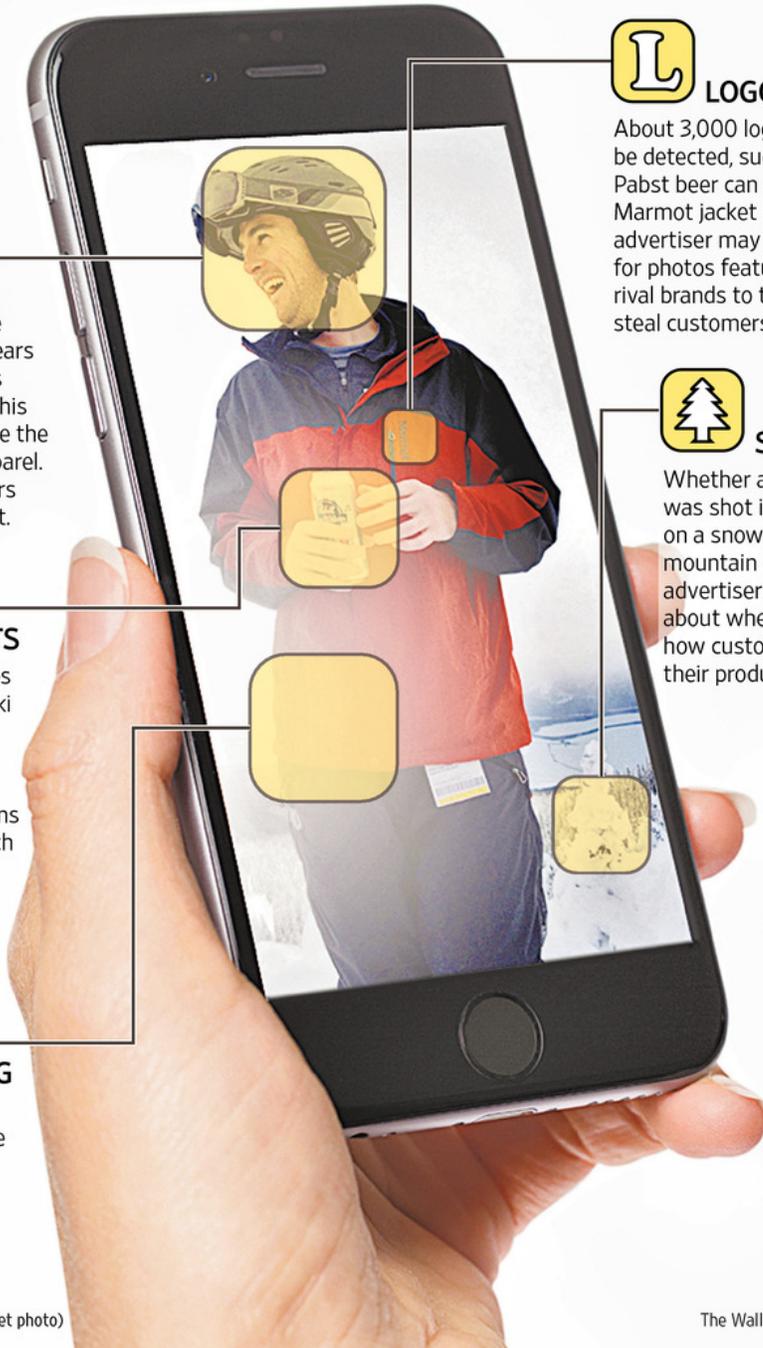
LOGOS

About 3,000 logos can be detected, such as the Pabst beer can and Marmot jacket here. An advertiser may search for photos featuring rival brands to try to steal customers.



SCENES

Whether a photo was shot in a bar or on a snowy mountain can give advertisers clues about where and how customers use their products.



By [DOUGLAS MACMILLAN](#) and [ELIZABETH DWOSKIN](#)

Updated Oct. 9, 2014 9:35 p.m. ET

Most users of popular photo-sharing sites like Instagram, Flickr and Pinterest know that anyone can view their vacation pictures if shared publicly.

But they may be surprised to learn that a new crop of digital marketing companies are searching, scanning, storing and repurposing these images to draw insights for big-brand advertisers.

Some companies, such as Ditto Labs Inc., [use software to scan photos](#)—the image of someone holding a [Coca-Cola](#) can, for example—to identify logos, whether the person in the image is smiling, and the scene’s context. The data allow marketers to send targeted ads or conduct market research.

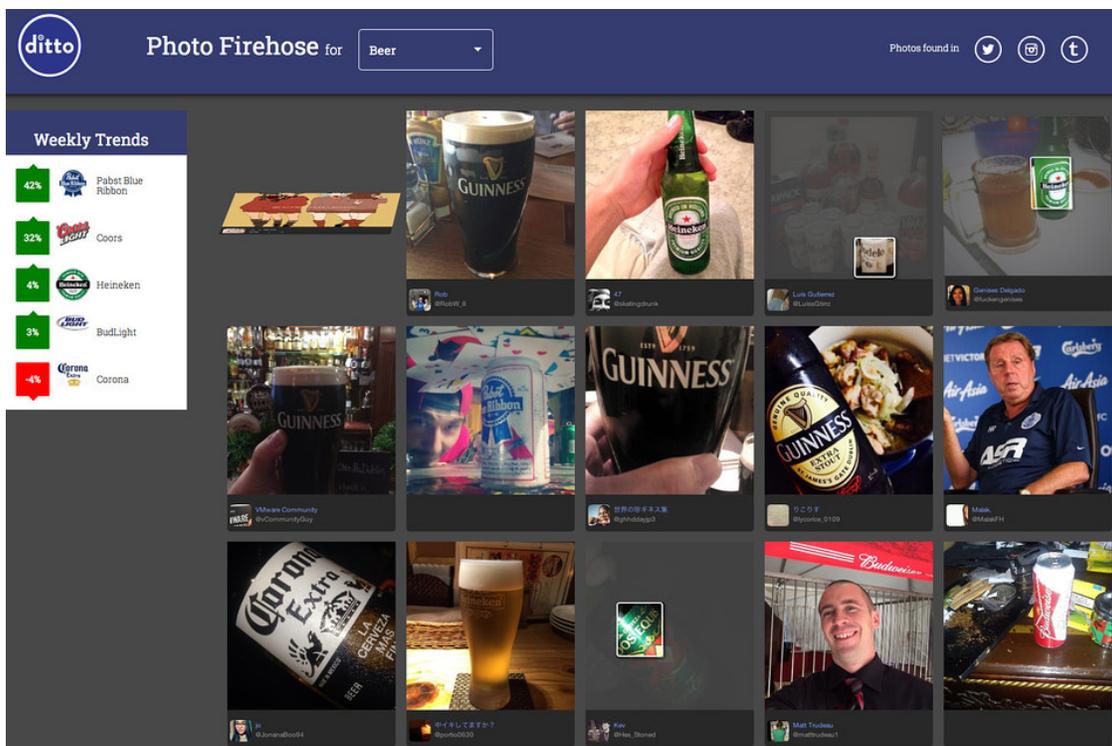
Others, such as Piqora Inc., store images for months on their own servers to show marketers what is trending in popularity. Some have run afoul of the loose rules on image-storing that the services have in place.

The startups’ efforts are raising fresh privacy concerns about how photo-sharing sites convey the collection of personal data to users. The trove is startling: Instagram says 20 billion photos have already been shared on its service, and users are adding about 60

million a day.

The digital marketers gain access to photos publicly shared on services like Instagram or Pinterest through software code called an application programming interface, or API. The photo-sharing services, in turn, hope the brands will eventually spend money to advertise on their sites.

Watchdogs contend these sites aren't clearly communicating to users that their images could be scanned in bulk or downloaded for marketing purposes. Many users may not intend to promote, say, a pair of jeans they are wearing in a photo or a bottle of beer on the table next to them, the privacy experts say.



A screenshot of the Ditto Labs site shows the fire hose of photos that it scans for brands. The site filters photos by categories such as beer. *DITTO LABS*

“This is an area that could be ripe for commercial exploitation and predatory marketing,” said Joni Lupovitz, vice president at children’s privacy advocacy group Common Sense Media. “Just because you happen to be in a certain place or captured an image, you might not understand that could be used to build a profile of you online.”

In recent years, startups have begun mining text in tweets or social-media posts for keywords that indicate trends or sentiment toward brands. The market for image-mining is newer and potentially more invasive because photos inspire more emotions in people and are sometimes open to more interpretation than text.

Instagram, Flickr and Pinterest Inc.— among the largest photo-sharing sites— say they adequately inform users that publicly posted content might be shared with partners and take action when their rules are violated by outside developers. Photos that are marked as private by users or not shared wouldn’t be available to marketers.

There are no laws forbidding publicly available photos from being analyzed in bulk, because the images were posted by the user for anyone to see and download. The U.S. Federal Trade Commission does require that websites be transparent about how they share user data with third parties, but that rule is open to interpretation, particularly as new business models arise. Authorities have

charged companies that omit the scope of their data-sharing from privacy policies with misleading consumers.

“Our API only provides public information to a handful of partners intended to help their clients understand the performance of their content on Pinterest.”

—Pinterest

The FTC declined to comment.

The photo sites’ privacy policies—the legal document enforced by law as promises to consumers—vary in wording but none of them clearly convey how third-party services treat user-posted photos.

For example, the privacy policy of Instagram, which is owned by [Facebook](#) Inc., directs its more than 200 million users to a separate document that explains rules for developers. Pinterest and Flickr, owned by [Yahoo](#) Inc., have no explicit mention of third-party developers in their privacy policies. Other popular sites for photos, including [Twitter](#) Inc. and another Yahoo-owned site, Tumblr, warn users they may share nonprivate content with third parties.

While Facebook is one of the largest photo-sharing sites, the fact that most of its users restrict their photos’ access with privacy controls has deterred outside developers from mining those images. Developers commonly use Facebook’s API to pull in

profile photos of its members but not for marketing purposes.

An Instagram spokesman said its partnerships with developers don't "change anything about who owns photos, or the protections we have in place to keep our community a safe place." Flickr said it takes steps to prevent outside developers from scanning photos on its site in bulk.

Pinterest said "our API only provides public information to a handful of partners intended to help their clients understand the performance of their content on Pinterest."

Spokeswomen for Tumblr and Twitter declined to comment.

Jules Polonetsky, the director of Future of Privacy Forum, an advocacy group funded by Facebook and other tech companies, said users should assume that companies are scanning sites for market research if their photos are publicly viewable.

But the boom in image-scanning technologies could lead to a world in which people's offline behavior, caught in unsuspecting images, increasingly becomes fodder for more personalized forms of marketing, said Peter Eckersley, technology-projects director for the Electronic Frontier Foundation.

Moreover, the use of software to scan faces or objects in photos is so new that most sites don't mention the technology in their privacy policies.

Advertisers such as [Kraft Foods Group](#) Inc. pay Ditto Labs to find

their products' logos in photos on Tumblr and Instagram. The Cambridge, Mass., company's software can detect patterns in consumer behavior, such as which kinds of beverages people like to drink with macaroni and cheese, and whether or not they are smiling in those images. Ditto Labs places users into categories, such as "sports fans" and "foodies" based on the context of their images.

Kraft might use those insights to cross-promote certain products in stores or ads, or to better target customers online. David Rose, who founded Ditto Labs in 2012, said one day his image-recognition software will enable consumers to "shop" their friends' selfies, he said. Kraft didn't respond to a request for comment.

Ditto Labs also offers advertisers a way to target specific users based on their photos posted on Twitter, though Mr. Rose said most advertisers are reluctant to do so because users might find it "creepy."

Mr. Rose acknowledges that most people who upload photos don't understand they could be scanned for marketing insights. He said photo-sharing services should do more to educate users and give them finer controls over how companies like his treat photos.

Beyond image recognition, some API partners employ a process called "caching," meaning they download photos to their own servers. One of the more common uses of caching is to build a

marketing campaign around photos uploaded by users and tagged with a specific hashtag.

The companies don't mention caching in their privacy policies and they vary in how long developers can store photos on their servers.

Tumblr, for example, restricts caching to three days while

Instagram says "reasonable periods."

Some developers have already overstepped the rules set forth by photo-sharing sites. Last month, Pinterest learned from a Wall Street Journal inquiry that Piqora, one of seven partners in its business API program, launched in May, was violating its image-use policy.

Piqora, a San Mateo, Calif., marketing analytics startup, collects photos into a graphical dashboard that help companies such as clothing and accessories maker Fossil Inc. track which of its own products and those of competing brands are most popular. This violated Pinterest's rules, which restrict partners from using images from the site that were posted by anyone except their own clients.

After Pinterest learned about the violation, the company asked Piqora to discontinue the practice and plans to begin performing regular audits of its business partners, a spokesman for Pinterest said. Fossil didn't respond to a request for comment.

Piqora co-founder and Chief Executive Sharad Verma says he has

removed the ability to view competitors' images in the dashboard. He also clarified his company's cached photos policy from Instagram. Rather than keeping photos for an indefinite period of time, Mr. Verna said he will now delete photos from his servers within 120 days.

“We might be looking at doing away with caching and figuring out a new way to optimize our software,” Mr. Verma said.

[Lisa Fleisher](#) contributed to this article.

Write to Douglas MacMillan at douglas.macmillan@wsj.com and Elizabeth Dwoskin at elizabeth.dwoskin@wsj.com