



The Switch

Wanted: The 'perfect babysitter.' Must pass AI scan for respect and attitude.

By [Drew Harwell](#)

November 23 at 11:50 AM

When Jessie Battaglia started looking for a new babysitter for her 1-year-old son, she wanted more information than she could get from a criminal-background check, parent comments and a face-to-face interview.

So she turned to Predictim, an online service that uses “advanced artificial intelligence” to assess a babysitter’s personality, and aimed its scanners at one candidate’s thousands of Facebook, Twitter and Instagram posts.

The system offered an automated “risk rating” of the 24-year-old woman, saying she was at a “very low risk” of being a drug abuser. But it gave a slightly higher risk assessment — a 2 out of 5 — for bullying, harassment, being “disrespectful” and having a “bad attitude.”

The system didn’t explain why it had made that decision. But Battaglia, who had believed the sitter was trustworthy, suddenly felt pangs of doubt.

“Social media shows a person’s character,” said Battaglia, 29, who lives outside Los Angeles. “So why did she come in at a 2 and not a 1?”

Predictim is offering parents the same playbook that dozens of other tech firms are selling to employers around the world: artificial-intelligence systems that analyze a person’s speech, facial expressions and online history with promises of revealing the hidden aspects of their private lives.

The technology is reshaping how some companies approach recruiting, hiring and reviewing workers, offering employers an unrivaled look at job candidates through a new wave of invasive psychological assessment and surveillance.

The tech firm Fama says it uses AI to police workers' social media for “toxic behavior” and alert their bosses. And the recruitment-technology firm HireVue, which works with companies such as Geico, Hilton and Unilever, offers a system that automatically analyzes applicants' tone, word choice and facial movements during video interviews to predict their skill and demeanor on the job. (Candidates are encouraged to smile for best results.)

But critics say Predictim and similar systems present their own dangers by making automated and possibly life-altering decisions virtually unchecked.

The systems depend on black-box algorithms that give little detail about how they reduced the complexities of a person's inner life into a calculation of virtue or harm. And even as Predictim's technology influences parents' thinking, it remains entirely unproven, largely unexplained and vulnerable to quiet biases over how an appropriate babysitter should share, look and speak.

There's this "mad rush to seize the power of AI to make all kinds of decisions without ensuring it's accountable to human beings," said Jeff Chester, the executive director of the Center for Digital Democracy, a tech advocacy group. "It's like people have drunk the digital Kool-Aid and think this is an appropriate way to govern our lives."

Predictim's scans analyze the entire history of a babysitter's social media, which, for many of the youngest sitters, can cover most of their lives. And the sitters are told they will be at a great disadvantage for the competitive jobs if they refuse.

Predictim's chief and co-founder Sal Parsa said the company, launched last month as part of the University of California at Berkeley's SkyDeck tech incubator, takes ethical questions about its use of the technology seriously. Parents, he said, should see the ratings as a companion that "may or may not reflect the sitter's actual attributes."

But the danger of hiring a problematic or violent babysitter, he added, makes the AI a necessary tool for any parent hoping to keep his or her child safe.

"If you search for abusive babysitters on Google, you'll see hundreds of results right now," he said. "There's people out there who either have mental illness or are just born evil. Our goal is to do anything we can to stop them."

A Predictim scan starts at \$24.99 and requires a babysitter's name and email address and her consent to share broad access to her social media accounts. The babysitter can decline, but a parent is notified of her refusal, and in an email the babysitter is told "the interested parent will not be able to hire you until you complete this request."

Predictim's executives say they use language-processing algorithms and an image-recognition software known as "computer vision" to assess babysitters' Facebook, Twitter and Instagram posts for clues about their offline life. The parent is provided the report exclusively and does not have to tell the sitter the results.

Parents could, presumably, look at their sitters' public social media accounts themselves. But the computer-generated reports promise an in-depth inspection of years of online activity, boiled down to a

single digit: an intoxicatingly simple solution to an impractical task.

The risk ratings are divided into several categories, including explicit content and drug abuse. The start-up has also advertised that its system can evaluate babysitters on other personality traits, such as politeness, ability to work with others and “positivity.”

The company hopes to upend the multibillion-dollar “parental outsourcing” industry and has begun advertising through paid sponsorships of parenting and “mommy” blogs. The company’s marketing focuses heavily on its ability to expose hidden secrets and prevent “every parent’s nightmare,” citing criminal cases including that of a Kentucky babysitter charged earlier this year with severely injuring an 8-month-old girl.

“Had the parents of the little girl injured by this babysitter been able to use Predictim as part of their vetting process,” a company marketing document says, “they would never have left her alone with their precious child.”

But tech experts say the system raises red flags of its own, including worries that it is preying on parents’ fears to sell personality scans of untested accuracy.

They also question how the systems are being trained and how vulnerable they might be to misunderstanding the blurred meanings of sitters’ social media use. For all but the highest-risk scans, the parents are given only a suggestion of questionable behavior and no specific phrases, links or details to assess on their own.

When one babysitter’s scan was flagged for possible bullying behavior, the unnerved mother who requested it said she couldn’t tell whether the software had spotted an old movie quote, song lyric or other phrase as opposed to actual bullying language.

Jamie L. Williams, a staff attorney at the civil-liberties group Electronic Frontier Foundation, said most algorithms deployed now to assess the meaning of words and images online are widely known to lack a human reader’s context and common sense. Even tech giants such as Facebook have struggled to build algorithms that can tell the difference between a harmless comment and hate speech.

“Running this system on teenagers: I mean, they’re kids!” Williams said. “Kids have inside jokes. They’re notoriously sarcastic. Something that could sound like a ‘bad attitude’ to the algorithm could sound to someone else like a political statement or valid criticism.”

And when the system gets it wrong — suggesting, for instance, that a babysitter abuses drugs — it can be impossible for a parent to know. The system’s clear-cut ratings and assertions of confidence might lead parents to expect it to be far more accurate or authoritative than a human could be, steering parents

toward sitters they otherwise would have avoided or away from people who had already earned their trust.

“There are no metrics yet to really make it clear whether these tools are effective in predicting what they say they are,” said Miranda Bogen, a senior policy analyst at Upturn, a Washington think tank that researches how algorithms are used in automated decision-making and criminal justice. “The pull of these technologies is very likely outpacing their actual capacity.”

Malissa Nielsen, Battaglia’s 24-year-old babysitter, gave her approval recently to two separate families who asked her to hand over social media access to Predictim. She said she has always been careful on social media and figured sharing more about herself couldn’t hurt: She goes to church once a week, doesn’t curse and is finishing a degree in early-childhood education, with which she hopes to open a preschool.

But after she learned that the system had given her imperfect grades for bullying and disrespect, she was stunned. She had believed she was allowing the parents to review her social media, not consenting to having an algorithm dissect her personality. She also hadn’t been told of the results for a test that could cripple her only source of income.

“I would have wanted to investigate a little. Why would it think that about me?” Nielsen said. “A computer doesn’t have feelings. It can’t determine all that stuff.”

Americans still harbor a lingering distrust over algorithms whose decisions could affect their daily life. In a Pew Research Center [survey](#) released this month, 57 percent of respondents said they thought automated résumé screening of job applicants was “unacceptable.”

But Predictim nevertheless says it is gearing up for a nationwide expansion. Executives at Sittercity, an online babysitter marketplace visited by millions of parents, said they are launching a pilot program early next year that will fold in Predictim’s automated ratings into the site’s current array of sitter screenings and background checks.

“Finding a sitter can come with a lot of uncertainty,” said Sandra Dainora, Sittercity’s head of product, who believes tools like these could soon become “standard currency” for finding caregivers online. “Parents are always seeking the best solution, the most research, the best facts.”

Predictim’s leaders also believe they can greatly expand the system’s capabilities to offer even more intimate measurements of a babysitter’s private life. Joel Simonoff, the company’s chief technology officer, said the team is interested in gaining “useful psychometric data” from babysitters’ social media by running their histories through personality tests, such as Myers Briggs, and offering to sell parents the results.

Predictim's social media mining and interest in mass psychological analysis mirror the ambitions of Cambridge Analytica, the political consultancy that worked for the Trump campaign and wrenched Facebook into a global privacy scandal. But Predictim's leaders say they have set up internal safeguards and work to protect babysitters' personal data. "If we ever leaked a babysitter's info, that would not be cool," Simonoff said.

Experts worry AI rating systems such as Predictim's portend a future where every job, not just in child care, is decided by a machine. A number of firms in hiring and recruiting are already building or investing in systems that can analyze candidates' résumés on a massive scale and provide an automated assessment of how each might perform. Similar AI systems — including from Jigsaw, a tech incubator created by Google — are used to patrol online comments across the Web for harassment, threats and abuse.

But hiring and recruiting algorithms have routinely been shown to hide the kinds of subtle biases that could derail a person's career. Amazon.com stopped developing a recruiting algorithm after learning that it had been unfairly penalizing female candidates, sources [told](#) Reuters — because the company's history of hires in the male-dominated tech industry had taught the system that male attributes were preferred. The company has said the tool was never used to evaluate candidates. (Amazon founder and chief executive Jeffrey P. Bezos owns The Washington Post.)

Some AI experts believe that systems like these have the potential to supercharge the biases of age or racial profiling, including flagging words or images from certain groups more often than others. They also worry that Predictim could coerce young babysitters into handing over intimate data just to get a job.

But Diana Werner, a mother of two living just north of San Francisco, said she believes babysitters should be willing to share their personal information to help with parents' peace of mind.

"A background check is nice, but Predictim goes into depth, really dissecting a person — their social and mental status," she said.

Where she lives, "100 percent of the parents are going to want to use this," she added. "We all want the perfect babysitter."

**Drew Harwell**

Drew Harwell is a national technology reporter for The Washington Post specializing in artificial intelligence. He previously covered national business and the Trump companies. [Follow](#)

The story must be told.

Your subscription supports journalism that matters.

Try 1 month for \$1

Be the first to know.

Our award-winning journalists are there when the news breaks.

Try 1 month for ~~\$10~~ \$1

Send me this offer

Market Watch

Dow 24,369.58

Today  0.39%

S&P 2,641.1

Today  0.33%

NASDAQ 6,960.99

Today  0.16%

Last Updated:12:45 PM 11/23/2018