

The New York Times | <http://nyti.ms/1FrObpk>

SundayReview | OPINION

If an Algorithm Wrote This, How Would You Even Know?

By **SHELLEY PODOLNY** MARCH 7, 2015

LET me hazard a guess that you think a real person has written what you're reading. Maybe you're right. Maybe not. Perhaps you should ask me to confirm it the way your computer does when it demands that you type those letters and numbers crammed like abstract art into that annoying little box.

Because, these days, a shocking amount of what we're reading is created not by humans, but by computer algorithms. We probably should have suspected that the information assaulting us 24/7 couldn't all have been created by people bent over their laptops.

It's understandable. The multitude of digital avenues now available to us demand content with an appetite that human effort can no longer satisfy. This demand, paired with ever more sophisticated technology, is spawning an industry of "automated narrative generation."

Companies in this business aim to relieve humans from the burden of the writing process by using algorithms and natural language generators to create written content. Feed their platforms some data — financial earnings statistics, let's say — and poof! In seconds, out comes a narrative that tells whatever story needs to be told.

These robo-writers don't just regurgitate data, either; they create human-sounding stories in whatever voice — from staid to sassy — befits the intended audience. Or different audiences. They're that smart. And when you read the output, you'd never guess the writer doesn't have a heartbeat.

Consider the opening sentences of these two sports pieces:

“Things looked bleak for the Angels when they trailed by two runs in the ninth inning, but Los Angeles recovered thanks to a key single from Vladimir Guerrero to pull out a 7-6 victory over the Boston Red Sox at Fenway Park on Sunday.”

“The University of Michigan baseball team used a four-run fifth inning to salvage the final game in its three-game weekend series with Iowa, winning 7-5 on Saturday afternoon (April 24) at the Wilpon Baseball Complex, home of historic Ray Fisher Stadium.”

If you can't tell which was written by a human, you're not alone. According to a study conducted by Christer Clerwall of Karlstad University in Sweden and published in *Journalism Practice*, when presented with sports stories not unlike these, study respondents couldn't tell the difference. (Machine first, human second, in our example, by the way.)

Algorithms and natural language generators have been around for a while, but they're getting better and faster as the demand for them spurs investment and innovation. The sheer volume and complexity of the Big Data we generate, too much for mere mortals to tackle, calls for artificial rather than human intelligence to derive meaning from it all.

Set loose on the mother lode — especially stats-rich domains like finance, sports and merchandising — the new software platforms apply advanced metrics to identify patterns, trends and data anomalies. They then rapidly craft the explanatory narrative, stepping in as robo-journalists to replace humans.

The Associated Press uses Automated Insights' Wordsmith platform to create more than 3,000 financial reports per quarter. It published a story on Apple's latest record-busting earnings within minutes of their release. Forbes uses Narrative Science's Quill platform for similar efforts and refers to the firm as a partner.

Then we have Quakebot, the algorithm The Los Angeles Times uses to analyze geological data. It was the "author" of the first news report of the 4.7 magnitude earthquake that hit Southern California last year, published on the newspaper's website just moments after the event. The newspaper also uses algorithms to enhance its homicide reporting.

But we should be forgiven a sense of unease. These software processes, which are, after all, a black box to us, might skew to some predicated norm, or contain biases that we can't possibly discern. Not to mention that we may be missing out on the insights a curious and fertile human mind could impart when considering the same information.

The mantra around all of this carries the usual liberation theme: Robo-journalism will free humans to do more reporting and less data processing.

That would be nice, but Kristian Hammond, Narrative Science's co-founder, estimates that 90 percent of news could be algorithmically generated by the mid-2020s, much of it without human intervention. If this projection is anywhere near accurate, we're on a slippery slope.

It's mainly robo-journalism now, but it doesn't stop there. As software stealthily replaces us as communicators, algorithmic content is rapidly permeating the nooks and crannies of our culture, from government affairs to fantasy football to reviews of your next pair of shoes.

Automated Insights states that its software created one billion stories last year, many with no human intervention; its home page, as well as Narrative Science's, displays logos of customers all of us would recognize: Samsung,

Comcast, The A.P., Edmunds.com and Yahoo. What are the chances that you *haven't* consumed such content without realizing it?

Books are robo-written, too. Consider the works of Philip M. Parker, a management science professor at the French business school Insead: His patented algorithmic system has generated more than a million books, more than 100,000 of which are available on Amazon. Give him a technical or arcane subject and his system will mine data and write a book or report, mimicking the thought process, he says, of a person who might write on the topic. Et voilà, “The Official Patient’s Sourcebook on Acne Rosacea.”

Narrative Science claims it can create “a narrative that is indistinguishable from a human-written one,” and Automated Insights says it specializes in writing “just like a human would,” but that’s precisely what gives me pause. The phrase is becoming a de facto parenthetical — not just for content creation, but where most technology is concerned.

Our phones can speak to us (just as a human would). Our home appliances can take commands (just as a human would). Our cars will be able to drive themselves (just as a human would). What does “human” even mean?

With technology, the next evolutionary step always seems logical. That’s the danger. As it seduces us again and again, we relinquish a little part of ourselves. We rarely step back to reflect on whether, ultimately, we’re giving up more than we’re getting.

Then again, who has time to think about that when there’s so much information to absorb every day? After all, we’re only human.

Related: Interactive Quiz: Did a Human or a Computer Write This? A shocking amount of what we’re reading is created not by humans, but by computer algorithms. Can you tell the difference? Take the quiz.

Shelley Podolny is a director at H5, a company specializing in search, and a writer

on issues related to electronic information.

A version of this op-ed appears in print on March 8, 2015, on page SR6 of the New York edition with the headline: If an Algorithm Wrote This, How Would You Even Know?.

© 2015 The New York Times Company