

'Crowdworking' provides the humans who train artificial intelligence

31 July 2019, by Melissa Hellmann



Credit: CC0 Public Domain

Eager to make extra money on the side, Washington, D.C., resident Paula Alves Silva turned to a gig emblematic of the digital age: She recorded sentences read aloud in the comfort of her home to help train artificial intelligence (AI) software.

Silva completed the tasks in her native Portuguese tongue for Seattle-based startup DefinedCrowd, which develops machine learning algorithms that power products for businesses including heavyweights MasterCard and BMW. Such recordings could be used in voice recognition products introduced in new countries, or to train existing systems to recognize non-native speakers or regional accents, the company says.

Silva earned \$20—from 8 to 33 cents per sentence—and considered that satisfactory given the short amount of time it took to complete the tasks. The knowledge that her [task](#) would contribute to a new artificial intelligence system was a bonus, she said.

When voice-activated software such as Amazon's

Alexa responds to the simple command of calling Mom, thousands of global workers have helped train the software to ensure that, say, Tom from work isn't dialed instead. The workers transcribe and annotate recordings that are fed back into the software to improve Alexa's human speech recognition (sometimes using recordings from unaware consumers, according to Bloomberg).

The rise of AI in the age of the gig economy has ushered in an invisible workforce in which ordinary people, like Silva, train technology to be smarter.

Created in 2015 by CEO and founder Daniela Braga, DefinedCrowd is one of many companies that use so-called crowdworking to teach tech devices how to follow commands. Others including Amazon Mechanical Turk, and Figure Eight, formerly known as CrowdFlower, were established over a decade ago.

In the past two years, DefinedCrowd has grown from 20 employees in 2017 to 110 globally. That's not counting the 130,000 people across 60 countries that DefinedCrowd says have worked on its tasks. It calls them the Neevo community and says they can work in 50 different languages.

Not all their tasks involve speech: For instance, if the makers of an autonomous vacuum want their machines to clean a living room without running into objects, crowdworkers would annotate the different objects in images of living rooms. Those annotated images would then be used as training data to teach the machines precision, said DefinedCrowd spokeswoman Catarina Salteiro.

"You can think of us as the gasoline of a car," said Salteiro. "You'll put good gas in there to make sure that it's working properly."

To complete tasks for DefinedCrowd, workers register with their email or social media account on the company's web platform or the recently

launched app. After passing a test based on a skill set such as native French fluency, the worker is sent email notifications of available tasks and paid through Paypal at the end of a gig.

Another DefinedCrowd crowdworker, Rakesh Kumar of Delhi, India, said in an email that the recordings he makes in English and Hindi provide a necessary extra income of nearly \$10 per month for about six [hours of work](#) altogether. It's helpful, although he noted the payment "is quite less than other freelance work I do."

A user gets paid only if their work is high quality and matches the requirements. Several members of the Neevo community tweeted at DefinedCrowd that they hadn't gotten paid yet, nor received notifications of new tasks in a while. "You guys have not paid many completed tasks and the site seems to be down," wrote one user in December.

The availability of jobs and the skills required for tasks are dependent upon the projects that clients develop, said Salteiro. Certain languages are in a higher demand for tasks than others. "As we grow our client base, the offer of work on our Neevo platform will increase and more tasks will become available for different languages and at a higher frequency," she said.

A 2018 International Labour Organization report found that crowdworkers who complete tasks for sites such as Amazon Mechanical Turk are paid low wages. Based on two surveys of 3,500 crowdworkers in 75 countries, the report found a third of them relied on the tasks as their main source of income. The report concluded that across five online global platforms—Amazon Mechanical Turk, Microworkers, CrowdFlower, Prolific and Clickworker—the average pay per hour amounted to U.S. \$4.43 for work considered payable. When accounting for work that was rejected, pay that wasn't received, or the amount of time it took to search for tasks, respondents averaged \$3.31 per hour. Similar to Kumar, nearly 90% of the surveys' respondents said they wanted more work than was available, with workers averaging about 25 hours of crowdwork per week.

Labor law professor Charlotte Garden of Seattle

University's School of Law considers crowdworking a form of outsourcing work that was once done by company employees. Such arrangements "can make workers more vulnerable" by preventing them from advancing into higher roles or enjoying the labor protections that regular employees have, said Garden.

Given the absence of clear guidelines on the treatment of crowdworkers, last week the Allen Institute for Artificial Intelligence (AI2) issued a set of ethical recommendations to AI companies on proper pay, privacy and transparency for crowdworkers. In a blog post, it said U.S. crowdworker companies should pay at least the U.S. average minimum wage of \$8.50 per hour. Minimum hourly wage in developing countries should be around \$4, given the lower prevailing incomes, AI2 recommended. Companies that employ crowdworkers also should be transparent about how long a task will likely take and conditions that may lead to the rejection of work, AI2 urged.

DefinedCrowd is already following the guideline of calculating payment based on the minimum hourly rate of the country where the job is aimed and the estimated time it will take to complete a task, said Salteiro. She wouldn't be more specific about the pay ranges.

The lack of governance mechanisms in crowdworking has St. Louis University employment law professor Miriam Cherry wondering about the end goal in relying on temporary workers to train AI data systems. "Is it just efficiency for the sake of efficiency ... or is it something that really could help people?"

One answer to Cherry's question could be found in DefinedCrowd's work with Portugal's biggest electricity company, EDP, according to company spokesperson Jorge Simões. Last year, DefinedCrowd's crowdworkers helped the company determine which electricity poles need to be repaired, a once expensive and time-consuming process that required specialists to survey the poles from a helicopter. DefinedCrowd instead devised a machine-learning algorithm to detect defects in poles from images captured by drones and from helicopters.

To create it, crowdworkers identified the type of damage in 900 images of electric poles that were then used to train and test a damage-detection AI system that monitors the state of poles.

EDP was so satisfied with the work that it became an investor in DefinedCrowd last year. The company is now working with DefinedCrowd to create a voice transcription algorithm that works in industrial environments with loud background noise and the use of technical language by service technicians.

To Simões, the work demonstrated that the most difficult tasks could be "successfully automated, optimizing the final result of the inspection by minimizing human error and visual limitations," even if it took human intelligence to get there.

©2019 The Seattle Times

Distributed by Tribune Content Agency, LLC.

APA citation: 'Crowdworking' provides the humans who train artificial intelligence (2019, July 31) retrieved 30 September 2020 from <https://techxplore.com/news/2019-07-crowdworking-humans-artificial-intelligence.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.