Read The Sonnet Co-Authored By Shakespeare, An MIT PhD Student & A Machine-Learning Algorithm

*Posted Jan 26, 2014 by*[*Natasha Lomas*](https://techcrunch.com/author/natasha-lomas/)*(*[*@riptari*](https://twitter.com/riptari)*)*



[[Image](http://www.flickr.com/photos/natematias/5644317701/sizes/z/in/set-72157601842647752/) of authoring interface by [natematias](http://www.flickr.com/photos/natematias/%22%20%5Ct%20%22_blank)]

Machine-learning as a technology is, without doubt, the force that will be shaping our digital world for years and years to come, making it smarter and more autonomous, and sometimes taking our breath away in the process with its apparent agency.

This branch of artificial intelligence is already doing that now, whether as the special sauce behind Nest, [the company just acquired by Google for $3.2 billion](https://techcrunch.com/2014/01/13/google-just-bought-connected-device-company-nest-for-3-2b-in-cash/), or the core of a non-profit startup project attempting to [create a better HIV vaccine](http://www.theverge.com/2014/1/23/5338676/the-immunity-project-hopes-to-build-a-better-hiv-vaccine-through-crowdfunding).

The basic premise is this: feed a machine-learning algorithm with particular data-set and its predictive powers can become startling.

Happily, humans aren’t excluded from this process. Machine learning remains a collaboration between man and machine — with the input of each enhancing and extending the other’s powers. So these algorithmic overlords don’t look like the type that want to [harvest us for our organs](https://techcrunch.com/tag/TIDWRTWHUFOO/). Unless you count harvesting the human brain’s ability to make decisions and process selections.

Below is just one example of machine-learning technology that has the power to startle and delight — not least because it involves a (posthumous) collaboration with the greatest writer in the English language: William Shakespeare, to create a new sonnet in the Shakespearean style.

Also involved: U.K. startup [SwiftKey](http://crunchbase.com/company/touchtype)‘s machine-learning powered word prediction engine. And a living human mind with an ear for poetry.

How was the new sonnet composed? SwiftKey’s engine was trained on the sonnets of Shakespeare, and one of its early staff members, [J Nathan Matias](http://natematias.com/portfolio/DesignArt/Swift-SpeareStatisticalP.html) — now doing a PhD at MIT Media Lab — wrote a new sonnet choosing words purely from the next-word suggestions generated by the algorithm.

SwiftKey’s keyboard software can normally be found helping (mostly) Android mobile users type faster by learning their slang, syntax and writing style — and applying that learning to populate tailored three next-word predictions. Give the SwiftKey keyboard enough time to get to know how you write and, provided your writing is not akin to James Joycean streams of consciousness, the algorithm will quickly get very good at guessing what next few words you’re likely reaching for.

But — fed with a particular data-set, and with the addition of a poetically minded human agent — this machine-learning engine can evidently be applied as a creative writing tool capable of creating pastiche writings in the style of the author whose original works you first fed to it.

As well as using SwiftKey’s engine, Matias also built a visual authoring interface (pictured above visualising word suggestions in the style of metaphysical poet John Donne) that extends the core machine-learning technology to specifically aid poetry creation. He called this project ‘[Swift-Speare](http://natematias.com/portfolio/DesignArt/Swift-SpeareStatisticalP.html)‘: aka a set of statistical experiments in “machine-learning-assisted poetry composition”.

“To write good poetry, I needed to know more than what words might come next. I needed to anticipate *future predictions* — what predictions would be made later if I choose this word over that? So I created [this touchscreen interface](http://www.flickr.com/photos/natematias/5644317701/in/set-72157601842647752) to visualize future predictions for poetry writing,” Matias tells TechCrunch.

The result? Multiple new works  (co-)created in the style of various authors — including the following ‘Shakespearean’ sonnet (which depicts a scorned lover struggling with the disconnect between his ongoing love for the outward appearance of the object of his desire, with the knowledge of rejection/betrayal that belies this surface beauty):

When I in dreams behold thy fairest shade

Whose shade in dreams doth wake the sleeping morn

The daytime shadow of my love betray’d

Lends hideous night to dreaming’s faded form

Were painted frowns to gild mere false rebuff

Then shoulds’t my heart be patient as the sands

For nature’s smile is ornament enough

When thy gold lips unloose their drooping bands

As clouds occlude the globe’s enshrouded fears

Which can by no astron’my be assail’d

Thus, thyne appearance tears in atmospheres

No fond perceptions nor no gaze unveils

Disperse the clouds which banish light from thee

For no tears be true, until we truly see

The work has no single author. It’s a collaboration whose only living human agent, the aforementioned Matias, also now a fellow at the Berkman Center for Internet and Society, at Harvard University — whose mind was responsible for the final word selections, and thus also for assembling (and dissembling) the poem’s core meaning — describes as requiring an acknowledged role for each of its different agents (i.e. both human and machine).

“The idea of the author is a well known myth within writing and publishing. Just like startups that promote the myth of the genius founder, we reward individuals for collective projects,” Matias tells TechCrunch, when I ask what sort of authorial ratio he would assign to the work.

“The economic and social impact on people’s lives is a core motivation of my work at the Media Lab and the Berkman Center. Together with my collaborators, I’ve been doing large scale data analysis and experiments to [measure](https://medium.com/ladybits-on-medium/a16c31e1cdf) and [change](https://www.youtube.com/watch?v=Rt9JLLMcjmE&list=PL055Epbe6d5a2Rulgnt6-ox-ar0ubfSqF&index=3) women’s visibility in online media. I’m also trying to [change how we acknowledge creativity online](http://civic.mit.edu/blog/natematias/designing-acknowledgment-on-the-web).”

“There are two related issues in your question,” he continues. “Who do I acknowledge and who holds copyright. I personally acknowledge all of us. Just like John Kani and Winston Ntshona get recognition with Athol Fugard for [the plays they brainstormed together](https://en.wikipedia.org/wiki/The_Island_%28play%29), I think that all three of us should be acknowledged. I tend to avoid ratios and talk instead about roles. Shakespeare supplied material for inspiration, SwiftKey clustered it, making suggestions. I made the final choices and arrangement.”

Matias concedes that copyright is a “trickier” question — owing to another disconnect between this type of co-creation — and more broadly between language as a shared communication medium rich with intentional and subconscious linguistic resonances echoing down through the ages vs the rigidity of the legal system.

“Bots already hold copyright and legally serve people for copyright infringement. According to [Tim Hwang](https://twitter.com/timhwang) of Robot Robot & Hwang, copyright and patent trolls sometimes use algorithmically generated shell companies to pursue these claims and minimise risk to themselves. Tim, who’s one of the fellows at the new [data & society research institute](http://www.fastcompany.com/3023467/fast-feed/danah-boyds-data-in-society-program-wants-your-help-freeing-the-data), is trying to map out these bots and figure out ways that the legal system can account for and respond to them,” he says.

“Even inside our heads, we write with other people’s words in mind,” he adds. “‘Words belong to each other,’ says Virginia Woolf in [the only surviving recording of her voice](http://news.bbc.co.uk/1/hi/entertainment/arts_and_culture/7684201.stm). She once said that she couldn’t think of the phrase ‘multitudinous’ without also thinking of Shakespeare’s Macbeth, who wonders if trying to wash his hands of guilt might make ‘the multitudinous seas incarnadine.’ (Maria Popova has [posted this clip to SoundCloud](https://soundcloud.com/brainpicker/words-the-only-surviving)).”

On the flip side of Matias’ algorithmically aided poetry, are bots and algorithms that search for found poetry online.

“Now that we have large amounts of human text available on the Internet, we’re also seeing search bots that try to *find* poetry in large datasets. The [Times Haiku](http://haiku.nytimes.com/) finds haikus in New York Times text, and the [Pentametron](http://pentametron.com/%22%20%5Ct%20%22_blank) finds [iambic pentameter in tweet text](https://techcrunch.com/2013/01/13/pentametron-is-a-twitter-poet-that-gives-bots-some-literary-cred/),” he points out, adding: “Algorithms that search for poetry are the reverse of my work — they’re looking among ordinary text for unexpected poetry that has already been written.

“My work with Swift-speare looks among existing poems for probable poetry that has not yet been written.”

Matias says he is currently collaborating on a “stealth art project” that involves another area of computer-assisted creativity known as “human computation”.

“Human computation is a third area of computer assisted creativity. [Michael Bernstein](http://hci.stanford.edu/msb/) at Stanford has pioneered a kind of writing that asks humans to perform writing tasks that we might ordinarily ask an algorithm. It’s a fascinating area,” he adds.

Does Matias believe an algorithm could ever become a poet in its own right? Meaning without any human agent involved in word selections, and without taking a brute force approach to composition — i.e. by writing infinite numbers of poems to stumble, by accident, upon a few good ones?

“I think I’ll see a successful automated poet in my lifetime. It won’t be easy: a poet is more than someone who makes poetry. Yet that doesn’t rule out algorithms,” he says. “It’s true that Western audiences want the stories of writers as much as we want their work. Especially at a time when readings are such an important part of poetry, it would be difficult for an algorithm alone to do everything.

“But people were disappointed when they learned that [@horse\_ebooks](https://twitter.com/Horse_ebooks) wasn’t a bot, and [Hatsune Miku](https://en.wikipedia.org/wiki/Hatsune_Miku%22%20%5Ct%20%22_blank) is popular in Japan despite being a humanoid persona in front of a voice synthesizer.”

As for machines taking the “sweating labour” route to composing poems — writing everything and letting people pick the poems they like — well, why not, argues Matias. Arrange enough words, and some of those configurations will resonate with someone, somewhere.

“This is the Internet; why not generate all the possible poems and see what turns out to be popular? This is how some of the online t-shirt sellers work. When it doesn’t [land them into trouble](http://www.theverge.com/2013/3/2/4057742/amazon-pulls-auto-generated-t-shirts-promoting-rape-and-violence), it seems to work well. Why not poetry?”

[Sonnet reproduced with kind permission of J Nathan Matias, SwiftKey — and, well, we couldn’t ask William Shakespeare but we hope he would approve]

https://techcrunch.com/2014/01/26/swift-speare