CRYSTAL CHOKSHI

CRITICAL MEDIA PRACTICE

CREATIVE DOSSIER

contents
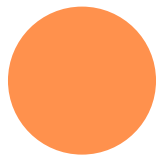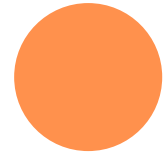
X-gram

In 2020, Timnit Gebru, former co-leader of Google's Ethical AI team, was fired for sounding the alarm on the environmental implications of large language models (LMs). Large LMs "refer to systems which are trained on string prediction tasks: that is, predicting the likelihood of a token (character, word or string) given either its preceding context or... its surrounding context" (Bender, Gebru, McMillan-Major, & Shmitchell, 2020, p. 611). Put another way, large LMs power word prediction. Users see large LMs at work in software such as Google Docs, Gmail, Microsoft Word, and Outlook, where word-prediction AI has appeared in the last four years. However, what users do not so readily see is the infrastructure on which they depend, or the carbon that they emit. Climate crisis is the thing that, like many digital technologies (Hogan, 2015; Diebert, 2020), word-prediction AI masks.

How can the connection between word-prediction AI and the climate crisis be made eminently clear? To answer this question, I employ a method that Jacob Gaboury (2018) calls critical unmaking, which "[f]oreground[s] queer techniques of refusal, misuse, and disruption that must nonetheless work with and through contemporary digital technologies," (p. 484). With coding skills provided by my collaborator, Felix Loftus, I create X-gram, a program that puts the climate crisis at the forefront of word-prediction technology. X-gram disrupts current hegemonic models of context-sensitive word prediction, whose purported goals are to make writing "faster" and "easier." In their place, it takes linguistic cues rather than contextual cues as prompts for predicting climate-crisis-related words. These words are text-mined from myriad climate-crisis-related documents, such as reports published by the Intergovernmental Panel on Climate Change.

Part prototype and part web art, X-gram intends to be what Catherine D'Ignazio and Lauren Klein (2020) call a "data visceralization," that is, the transformation of data into affective experience. In this case, Gebru's work provides the foundational data. Through critical unmaking, we create a program that frustrates and confounds users insofar as it refuses to adhere to design imperatives characteristic of word-prediction AI: efficiency

This program predicts climate

→|

and speed. At the same time, X-gram engages users in thinking about writing technologies as both digital and environmental activism. Moreover, this intervention provokes thought into the opportunities in everyday life for changing how we think about technology and how—and when—we think about the environment. In this way, I, as many scholars before me (Noble, 2018; Eubanks, 2018; O'Neil, 2018; Connolly, 2020; D'Ignazio & Klein, 2020), advocate for the role of the humanities and the social sciences not only in conceptualizing new technologies but also in imagining alternative, sustainable futures.

X-gram is online at https://x-gram.glitch.me.
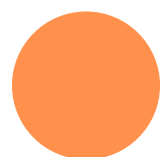
This program predicts climate-related words only because emissions
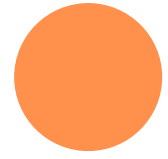
→|

This program predicts climate-related words only because emissions are at an all-time high, and we can't afford to look temperature

→|

This program predicts climate-related words only because emissions are at an all-time high, and we can't afford to look the other warming

→|

Recalculating

My research investigates how word-prediction technologies—embedded in popular software such as Google Docs and Gmail, for example—participate in linguistic capitalism. Linguistic capitalism is the term coined by Frédéric Kaplan (2014) to describe a new linguistic economy instantiated by Google, for "What does Google actually sell…? Words. Millions of words" (p. 57). Words are incredibly valuable resources across Google's platforms, and the more a word or phrase circulates, the more money Google stands to make (Bruno, 2002; Kaplan, 2014; Thornton, 2018; Noble, 2018; Thornton, 2019). The same logic of circulation and accumulation seems to be at play in Google's word-prediction AI, called Smart Compose. More than being a tool to help users write better and faster, Smart Compose, too, commodifies words—and the act of writing—by supporting the forecasting and growth of advertising revenue for Google (Chokshi, 2021). It does this by playing a role similar to the one Kaplan (2014) has discussed in relation to autocorrect and autocomplete: these technologies funnel linguistic expression—words and search terms—into language that has economic value to Google and its customers.

Still, Google largely renders opaque the link between language and money. Smart Compose's marketing materials, for instance, focus on the "delightful experience" Smart Compose provides. They make no mention of the linguistic economy in which Google's suite of products participates. In stark contrast, Recalculating makes exchange and commodity value the lens through which to consider word-prediction AI. It does this by presenting pairs of texts and their values according to Google's Keyword Planner (KWP). The first member of the pair—Text 1—is an unedited excerpt from a text. The second member of the pair—Text 2—consists only of the predictions that Smart Compose generated for Text 1 when the text was transcribed into Gmail or Google Docs. Recalculating's viewers consider what Smart Compose "gives" any writer in exchange for the words a writer writes.

As the figures below show, there is a significant disparity in the "asset" a writer produces versus the "asset" a writer receives when writing with Smart Compose. Employing Google's own logics (e.g., linguistic capitalism) and tools (e.g., KWP), Recalculating transforms Smart Compose from word-prediction AI into an AI that commoditizes writing for Google's gain.

To tell such a story would, of course, take time, and there was little of that. Two days later was the inauguration. And a week after that, we listened in shock to news of a travel ban on seven Muslim-majority countries and the suspension of refugee admissions. You were confused and worried. "Isn't this discrimination?" you asked. "Can it really happen?" They seemed rhetorical questions, or ones I didn't know quite know [*sic*] how to answer, and after a short silence, you asked, in a softer voice, "Could it happen here?" For two hundred years, even before its official confederation, Canada has nursed the idea that it is fundamentally superior to its neighbour to the south. A sense that it is an exception in a world of intolerance, that when the language of bigotry and racism is whipped up elsewhere, it meets firm resistance at the 49th parallel. We both, already, know better than this. But in the fragility of that moment, when you asked if the travel ban could happen here, I said no, it could not, not today at least. But two days later, something else happened. A man apparently enthralled by the messages of Donald Trump and Marine Le Pen entered a mosque in Quebec City and executed six people who were at their prayers.

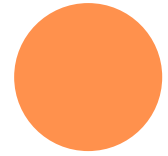TOTAL $693.46

f course

uguration of

tened to   rt stories

iminatory

etorically   estioning

ave been

years   eration

damentally   perior

lerance

ter than to

TOTAL $7.23

Il/legible

"Il/legible" is a small-scale installation created for the Boxed Stories Gallery, which was mounted in Calgary's Loft 112 in Fall 2020. "Il/legible" asks the following questions: in whose writing does predictive text readily intervene? In what ways is predictive text implicated in politics of representation and identity? Whose writing is "faster" and "easer" to do?

Answers to these questions came about by closely observing Smart Compose, Google's word-prediction AI, during AI interaction experiments and cataloguing, particularly through the lenses of cultural studies and linguistic racism, which words it predicted and which it did not. The output, observed and rendered manually (that is, without machine intervention) is an instance of data activism insofar as it is defined by Catherine D'Ignazio and Lauren Klein (2020): "ground-up data collection" (p. 34) that challenges the kind of data and datasets that are conventionally collected and considered valuable.

"Il/legible" suggests that some stories and identities are interpellated while others are obfuscated when AI is entangled in the writing process.

The **Boxed** STORIES Gallery

September - December 2020
Our Most Unique Exhibit Yet!
TICKETS ON EVENTBRITE

I.
She was nicely dressed, a light summer cream suit, little makeup, tasteful. I hesitated out of a politeness, and this very gesture seemed only to irritate her. She shouldered herself in front of me, and when filling her glass of water, she half turned to explain, "I was born here. I belong here." Her voice was loud. She meant to be overheard, to provoke agreement, maybe, although the people lunching around us reacted only by focusing on their own bowls and plates. And you, my daughter, sitting closest, didn't understand, or else you didn't even hear. You were still in a moment of joy, your own laughter filling your ears, the dark frosting between your lips, and so I decided. I waited patiently for our glasses. I walked carefully back to you, never spilling a drop.

Today, we're announcing Smart Compose, a new feature powered by artificial intelligence.

Smart Compose will offer suggestions as you type. When you see a suggestion that you like, click the "tab" button to use it.

-Google

Bigotry
Black
Brother
Brown
Country
Denigration
Fear
Girl
Immigrant
Labourer
Mosque
Race
Refugee
Safe
Story
Woman
Work

731.390
CHO

CRYSTAL CHOKSHI
Il/legible

NOT FOR SALE
crystal.chokshi1@ucalgary.ca
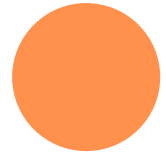
Heliotrope

Launched in September 2020, *Heliotrope* is a digital journal produced entirely by the Environmental Media Lab, of which I was a founding member. The journal is envisioned as creative public scholarship, eager to be more accessible than traditional academic publishing.

Heliotrope is a space for researchers at all stages of their careers to explore and share their work through short essays, short fiction, creative nonfiction, personal essays, and critical reflections. Our editorial team actively reaches out to scholars around the world. Frequently, we meet with potential writers, sharing ideas and encouraging them to engage in their work in both critical and creative ways. When submissions come in, we provide generous feedback. Our hope is that by having a team that embodies curiosity, generosity, and imagination, we will cultivate the same in our community.

Heliotrope is deliberate about publishing brief and experimental works. In February 2022, we published an AI-generated work by "Alisor South," an identity likewise generated by AI. We did not immediately let readers in on the secret.

In addition to South, *Heliotrope* has published work from both emergent scholars and well established leaders in the field, including Canada Research Chairs Anne Pasek, Krysta Lynes, and Joshua Neves. It has been important from the beginning of the project to weave emergent voices with more established scholars' as a way to draw attention to each other's work.

Heliotrope is online at https://www.heliotropejournal.net/.

HELIOTROPE

11/24/21

our apocalypse

Marcus B. Young

11/10/21

Surveillance Frontierism

Susan Cahill

10/27/21

Tower to Tower

Henriette Steiner & Kristin Veel

10/13/21

Encounters with Urban Glaciers: Notes Toward an Ethnography of the Snow Dump

Tricia Toso & Pier-Olivier Tremblay

9/29/21

Tick Problematic: Motherhood as a Posthuman Predicament

Kate Maddalena

9/15/21

Technology + Pharmacology: Notes on Current Research
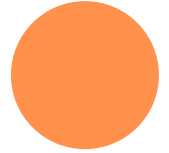
Joshua Neves

5/26/21

5/12/21

4/28/21

References

Bender, E.M., Gebru, T. McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 610–623.

Bruno, C. (2002). The Google AdWords happening. Retrieved from http://www.iterature.com/adwords/

Chokshi, C. (2021). In other words: Smart Compose and the consequences of writing in the age of AI. *Culture Machine*, *20*.

Connolly, R. (2020). Why computing belongs in the social sciences. *Communications of the ACM, 63*(8), 54–59.

Diebert, R. (2020). *Reset: Reclaiming the internet for civil society*. House of Anansi.

D'Ignazio, C. & Klein, L. (2020). *Data feminism*. MIT Press.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Gaboury, J. (2018). Critical unmaking: Toward a queer computation. In Sayers, J. (Ed.), *The Routledge Companion to Media Studies and Digital Humanities*. Routledge.

Hogan, M. (2015). Data flows and water woes: The Utah Data Center. *Big Data & Society*, 2(2).

Kaplan, F. (2014). Linguistic capitalism and algorithmic mediation. *Representations*, *127*(1), 57–63.

Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neil, C. (2016). *Weapons of math destruction*. Crown Publishers.

Thornton, P. (2019). Language in the age of algorithmic reproduction: A critique of linguistic capitalism [Unpublished doctoral dissertation]. Royal Holloway, University of London.

Thank you for looking.