

Begged question (Is it dangerous?)

slippery slope

How dangerous is AI? Regulate it before it's too late
BY CYNTHIA RUDIN, OPINION CONTRIBUTOR, THE HILL - 02/08/23 5:00 PM ET

As an Artificial Intelligence researcher, I've always felt the worst feature of AI is its role in the spread of lies. The AI-amplification of lies in Myanmar reportedly contributed to the Rohingya massacre, the spread of COVID-19 and vaccine misinformation likely contributed to hundreds of thousands of preventable deaths and election misinformation has weakened our democracy and played a part in the Jan. 6, 2021 insurrection. This was all possible because humans turned algorithms into weapons, manipulating them to spread noxious information on platforms that claimed to be neutral. These algorithms are all proprietary to companies, and they are unregulated. And so far, none of the companies have admitted any liability. Apparently, no one feels guilty.

If the federal government doesn't start regulating AI companies, it will get a lot worse. Billions of dollars are pouring into AI technology that generates realistic images and text, with essentially no good controls on who generates what. This will make it exponentially easier to generate fake news, fake violence, fake extremist articles, non-consensual fake nudity and even fake "scientific" articles that look real on the surface. Venture capital firms investing in this technology liken it to the early launch of the internet. And as we know, it's much easier to spread outrageous falsehoods than it is to spread the truth. Is this really like the beginning of the internet? Or is this like launching a nuclear bomb on the truth?

AI startups say that by making this technology public, they are "democratizing AI." It's hard to believe that coming from companies that stand to potentially gain billions by getting people to believe it. If they were instead about to be the victim of a massacre stemming from AI-generated misinformation, or even a victim of AI-amplified bullying, perhaps they might feel differently. Misinformation is not innocent — it is a major cause of wars (think of WWII or Vietnam), although most people are unfamiliar with the connection.

There are things we can do right now to address these critical problems. We need regulations around the use and training of specific types of AI technology

Let's start with regulating facial recognition technology (FRT) — that is, unless you don't mind being recognized by AI and then kicked out of Radio City Music Hall because of ongoing litigation involving your employer. FRT users should have to get a license or certification to use it or develop it, which comes with training for all users and developers.

Ⓜ We should figure out how to reduce the spread of particularly harmful misinformation; an easy solution to this is to make social media companies responsible for posted content, like any other publisher. Other countries have such laws, but the U.S. doesn't.

Ⓜ We also should enforce existing laws around monopolistic practice, which will allow users to choose social media platforms. If you cannot easily download your data from your social media platform and upload it into a new one, then the social media company is holding your data hostage, which is arguably monopolistic. More competition will allow users to choose content moderation platforms. We do not all need to be supporting companies that can host and perpetuate real harm online and in the real world, without much effort to combat it. We do not all need to be subject to the same attention-seeking algorithmic behavior.

Ⓜ We should force companies to remove all child abuse content. It is embarrassing that AI can easily find this content but is not enabled to remove it. Even more embarrassing is that the companies apparently don't always remove it when they are notified or delay efforts to do so.

It is extremely important that interpretable (transparent) models are used for high-stakes decisions that deeply affect people's lives. I have written extensively about this, pointing out that for high-stakes decisions, interpretable models have performed just as well as black box models, even on difficult

loaded dictionary, slippery slope

tone changes

personas or a hero in control

prep talk

TA is
in the
middle
EITHER
OR

slippery
slope
↓

unknown
future
↓

"like the beginning of
the internet"
write title here:
(circle loaded words)

"launching a nuclear
bomb on the truth"
runaway train
avalanche of misinformation

😊
"Good Guys"

☹️
"Bad Guys"

WHO: list pos. loaded words *implication* WHO: ... list neg. loaded words *implication*

"we citizens" here are things we
"over democracy" can do

"AI" - "worst feature is spread of lies"
"amplification of lies"
"contributed to Rohingya massacre"
"spread of Covid-19"
"vaccine misinformation"
"weakened out democracy"
"played a part in insurrection"

Regulations
"license"
"certification"
"training"
"laws"
"govt agency"
for AI

"humans" "turned algorithms into weapons"
"spread doxious info"
don't "feel guilty"

companies/
AI startups say there "democratizing AI"
"potentially gain billions"
"can host + perpetuate harm"
don't appear to care

Existential Impact? "spread" "fake" "billions" "should" "high-stakes" "regulate"
Authorizing Tropes? org from authority (researcher) "my job"
Assertive or Accommodating? both.
Other rhetorical tactics (look @ toms list) Absolutes/exagg, "worst" "all" "none" "no one"
neg. loaded diction: "lies" "massacre" "deaths" "billions"
"outrageous falsehoods" "nuclear bomb" "bullying"

Overt Thesis: "we" need to step in + demand regulation or (slippery slope)
Implied Theses: AI "spreads" like a virus + undermines the American
way of life. We will descend into totalitarianism,
if we don't stop it. Only by using the democratic
Target Audience: process + enacting laws will we defeat the
wealthy companies trying to gain power.

TA: Not scientists. Not overly computer savvy. Pro-Rohingya
(leaning liberal) believes Jan 6 was insurrection. Pro-
vaccine. Disgusted by greed (not elite). Feels alone +
powerless. Looking for a hero. Has lost faith in
democracy.