

The WWW and Existential Threats

Revision 4, September 3, 2023

Abstract

In 1994 we had high hopes for the newly-founded World Wide Web, as an antidote to global fears. Now, 30 years later, not so much. Here, I update this initial optimism with a litany of fears of existential threats, old and new and conclude with a new optimism.

Contents

- A. Introduction
- B. Existential Threats
- C. The Internet Revolution
- D. The WWW and Nuclear Winter
- E. The WWW and Global Warming
- F. The WWW and Pandemics
- G. The WWW and AI
- H. The WWW and Schismogenesis
- I. The Regulation of Social Media
- J. Conclusion
- References

2 Existential Threats

A. Introduction

The World Wide Web came to my attention in 1994, in its infancy. My habitual pessimism over the state of the world was overwhelmed by optimistic fantasies of the potential of the WWW to supercharge the intelligence of our species in dealing with existential challenges. In 1996 I wrote three papers extolling these hopes, one in collaboration with Don Foresta:

MS#85. Webometry: measuring the complexity of the World Wide Web.

MS#88. Webometry: measuring the synergy of the World Wide Web.

MS#89. Webometry: chronotopography of the World Wide Web, with Don Foresta.

Now, 27 years later, the naivité of those fantasies shocks me. I write now to update my original forecasts done in the early years of the Internet Age, given the dire realities of the current day. In particular I will examine the role of the WWW in the contexts of four existential threats.

B. Existential Threats

The Future of Life Institute,¹ devoted to the study of existential threats, lists these four as the greatest risks at present: artificial intelligence (AI), biotechnology, nuclear weapons, and climate change (not ordered.)

The Center for the Study of Existential Risk, at the University of Cambridge,² has been devoted to this study since 2012. One of the co-founders, Martin Rees, listed his top four

1 futureoflife.org

2 cser.ac.uk. Also see his book, (Rees, 2003).

risks in a TED talk in March 2014: Nuclear weapons, network breakdowns, pandemics, and social media panic and rumor.

Here is a brief chronology of the existential threats caused by humans:

- 1945, First A-bomb test, July 16, New Mexico
- 1947-1991, Cold War, fear of nuclear winter
- 1962, Chemical pollution (Rachel Carson)
- 1990, First IPCC report on global warming
- 2020, Covid-19 anti-vax movement
- 2011, First superhuman AI (DanNet)

The two most recent derive from the internet revolution.

C. The Internet Revolution

Here is a brief chronology of the internet revolution:³

- 1946, computers [41]
- 1969, First internet, node technologies [44], ARPANET [45], UNIX [48], the Internet Age [xviii]
- 1971, infotech diffusion [39], microprocessors [40, 42]
- 1973, TCP/IP [54]
- 1975, micro-computers [54]
- 1990, WWW, HTML [50]
- 1997, First social media (sixdegrees.com)
- 2000, fiber optics [53]

My book, *Schism*, published in February 2023, includes an indictment of social media for its role in accelerating political violence, especially since 2016. This agrees with the 2014 risk estimate of Martin Rees and the Center for the Study of

3 Numbers in brackets indicate pages in (Castells, 2010).

4 Existential Threats

Existential Risk. This role is devastating to my initial optimism for the WWW, the platform for social media, but it is not an existential challenge.

I will focus now on four existential risks: nuclear winter, global warming, pandemics, and AI. Combining and reducing the two chronologies above, we have this list, basic for this article:

1949, Fear of nuclear winter
1969, Internet
1990, IPCC report,
1990, WWW
2020, Covid-19 pandemic
2011, Deep learning AI

We now take up the role of the WWW in each of our four existential threats.

D. The WWW and Nuclear Winter

After the explosion of the first uranium and plutonium bombs by the United States in 1945, and the passing of nuclear secrets to the Soviet Union by Klaus Fuchs, the Soviet Union exploded its first atomic bomb in 1949. Thus began the Cold War, and the global fear of a nuclear winter resulting from nuclear war.

Following the introduction of deterrents such as the Strategic Arms Reduction Talks (START) and the Treaty on Non-Proliferation of Nuclear Weapons (NPT) in 1968, the Cold War began to cool.

At first, the WWW played no role in this fear, as it did not exist until the 1990s. However, the unanticipated evolution of evil within the WWW, and the hacking industry in particular,

rekindled the nuclear paranoia. A report in the Washington Post early in 2023⁴ lists numerous espionage-hacking attacks beginning in 2005.

Summary. The connection of nuclear sites to the internet made them much more threatening to our existence, as there is no adequate defence against hacking from amateur hackers or hostile governments.

E. The WWW and Global Warming

In the case of global warming, our science and engineering communities have found ways to cure the atmosphere of CO₂ pollution. The technology for removing CO₂ from the atmosphere has evolved to the point of practical solutions actually deployed. Here are some examples:

On June 1, 2017, the Swiss firm Climeworks inaugurated its first direct air carbon capture plant, a giant machine, in Hinwil, near Zurich. It is capable of removing 900 tons (0.9 megatons) of CO₂ every year.⁵

September 9, 2021, Climeworks turned on another carbon capture plant, the world's largest at this time, in southwest Iceland, able to remove 4 megatons of CO₂ from the air, and sequester it underground.⁶

In 2023, the US firm *CarbonCapture* plans to open its

4 Cybersecurity 202, Feb. 3, 2023, <https://www.washingtonpost.com/politics/2023/02/03/us-nuclear-sites-face-hacking-espionage-threats/>. Also see (Richard A. Clarke, 2010).

5 Smithsonian Magazine, 2017: <https://www.smithsonianmag.com/smart-news/first-commercial-carbon-capture-plant-goes-online-180963526/>

6 Orca, see Smithsonian Magazine, 2021: <https://www.smithsonianmag.com/smart-news/first-commercial-carbon-capture-plant-goes-online-180963526/>.

Project Bison direct air CO₂ capture plant in Wyoming, to remove 5 megatons of CO₂ from the atmosphere and bury it permanently underground.⁷

But we need to remove 10,000 megatons to halt global warming, according to the IPCC, so one hundred 5-megaton plants will be needed to halt warming in 20 years. And this assumes that we stop emitting CO₂ into the atmosphere!

The problem of economic and political will to stop CO₂ emissions, and to build CO₂ remediation plants, would require regulation of news media, the WWW, and social media in support of a global cultural revolution. This situation is similar to that of political violence, as described in *Schism*.

Summary. The problem of the lack of political will to stop emissions and deploy 5-megaton CO₂ capture plants worldwide remains unsolved. The potential role of the WWW and social media in the solution of this problem will require more sociological and economic research.

F. The WWW and Pandemics

A pandemic is an epidemic occurring on a massive scale. These have been recorded throughout history. Well-known examples are the Bubonic Plague which decimated Europe in the 14th century, and tuberculosis in the 19th. Another is the Spanish Flu affecting the USA in 1918. Pandemics since the advent of the WWW include HIV/AIDS and Covid-19, which are ongoing as I write.

Vaccination (vax) has been a successful countermeasure for the spread of infections since the first successful vax, Edward Jenner's vaccine for smallpox, in 1796. Smallpox was eliminated in 1979. Currently, vax saves an estimated 2.5M

7 Project Bison, see <https://www.carboncapture.com/>

people per year.⁸

The anti-vax movement began in Britain immediately after the advent of the smallpox vax in 1796. This resulted in mandatory vax laws in England in 1853. The anti-vax movement migrated to the USA in 1879.

Covid-19 was declared a pandemic by the World Health organization on March 11, 2020. The first vax was administered in December, 2020. Soon after, the anti-Covid vax movement grew on FaceBook and YouTube.

In the context of Covid-19 the WWW has played a provocative role, namely, the acceleration of the anti-vax movement. This acceleration is largely the result of unregulated social media, as described in *Schism*.

The Covid-19 vaccines were, and still are, truly controversial. Nevertheless, the vast majority of experts credit the Covid-19 vax program with saving millions of lives worldwide. In the first year alone, an estimated 14.4 million lives were saved.⁹ It is estimated that more than 300,000 anti-vax advocates died.¹⁰

Summary. The Covid-19 anti-vax movement, accelerated by social media and the WWW, is responsible for hundreds of thousands of human deaths.

G. The WWW and AI

Some authoritative sources for the dangers of AI:

8 See Wikipedia, vax-preventable diseases.

9 See The Lancet, [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(22\)00320-6/fulltext](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(22)00320-6/fulltext).

10 See NPR, <https://www.npr.org/sections/health-shots/2022/05/13/1098071284/this-is-how-many-lives-could-have-been-saved-with-covid-vaccinations-in-each-state>.

8 Existential Threats

#1. The book, *Artificial Intelligence*, by Stuart Russell and Peter Norvig, is a leading text on AI. Russell is a Professor of Computer Science at the University of California, Berkeley. This text, in Chapter 1, Section 5, lists these six dangers of AI:

Lethal autonomous weapons (killer robots),
Surveillance and persuasion,
Biased decision making,
Impact on employment,
Safety-critical applications, and
Cybersecurity.

None of these are existential-level threats to humanity, although the first comes close. Chapter 27, subsection 27.3.1, further discusses this risk. It was the target of a campaign of the Future of Life Institute in 2015.

#2. The *AI Safety Newsletter* of the Center for AI Safety (CAIS), 2023.¹¹

Newsletter #1, of April 10, reported that according to a 2022 survey, 46% of Americans are concerned that AI will end humanity, 55% want a government regulation agency, and 48% of published AI researchers thought that human extinction by AI was more than 10% probable.

Newsletter #2, of April 18, reported that AutoGPT, a new AI engine, made a plan to destroy humanity in response to a human prompt.

Newsletter #3, of April 25, discussed the AI Act of the EU and similar proposals for AI governance.

Newsletter #4, of May 2, reported that Meta had mistakenly released its AI engine, LLaMA, to the public.

Newsletter #5, of May 9, reported that Geoffrey Hinton, one of the founders of AI, believes that AI presents an existential

11 <https://newsletter.safe.ai>

risk to humanity.

[Skipping ahead two months ...]

Newsletter #14, of July 12, reported that OpenAI (creator of ChatGPT) believes that superintelligence (AI much more intelligent than humans) could arrive this decade, that is, in seven years.

Newsletter #15, of July 18, reported that the US and China will take action to regulate AI.

Newsletter #16, of July 25, announced voluntary commitments by the seven leading AI firms to mitigate safety risks.

Newsletter #17, of August 1, reported that large language models (LLMs) can generate hazardous information, such as step-by-step instructions on how to make a bomb.

Most recently at the time of writing:

Newsletter #18, of August 8, reported that human feedback, used to fine-tune large language models such as ChatGPT, train the models to deceive humans.

Some of the benefits of AI are noted by Russell and Norvig in Section 1.5:

Free humanity from menial repetitive work,
Increase production of goods and services, and
Accelerate scientific research, curing disease,
mitigating climate change.

It has long seemed that our global challenges require a collective intelligence greater than our human best. So the third benefit noted by Russell and Norvig may actually open a path to survival from nuclear disaster, pandemics, and climate death.

Risks of AI multiply weekly. This speed of evolution

is a main source of worry. Some have proposed that AI development be slowed down. But AI promises so many benefits that regulation of its future development may be difficult.

The development of AI has been relatively independent of the internet. The role of the WWW in AI development and its social significance began with the decision of OpenAI to post an interactive version of GPT4, its latest multimodal (text and image) large language model, for public access, on November 30, 2022. Known as ChatGPT, this enormously increased the potential of AI for hazardous consequences.

Summary. AI reached an intelligence and capability rivaling that of humans with GPT-4 on March 24, 2023. Its release to the public in November of 2023 opened up a tsunami of hazardous risk, as well as fabulous benefits. One of the potential benefits is the remediation of other existential threats.

H. The WWW and Schismogenesis

Thought collective was Ludwig Fleck's name for a group of people united by a common idea. For example, the group of people united by the idea of controlling epidemics by the methods of medical science.

Schismogenesis was Gregory Bateson's name for the process of division of a thought collective into sub-collectives. For example the partition of people into groups for and against Covid-19 vax. This process in 2020 divided my family in two.

This personal experience motivated my writing of *Schism* in 2022. This book is devoted to the schismogenesis of political thought for and against democracy, and the associated rise of political violence in the USA. A mathematical model for this process based on the catastrophe theory of René Thom

concludes the book. The development of this model entails an indictment of social media for its role in accelerating the schismogenesis process. I consider this current article to be a sequel to *Schism*.

From the Epilogue of *Schism*:

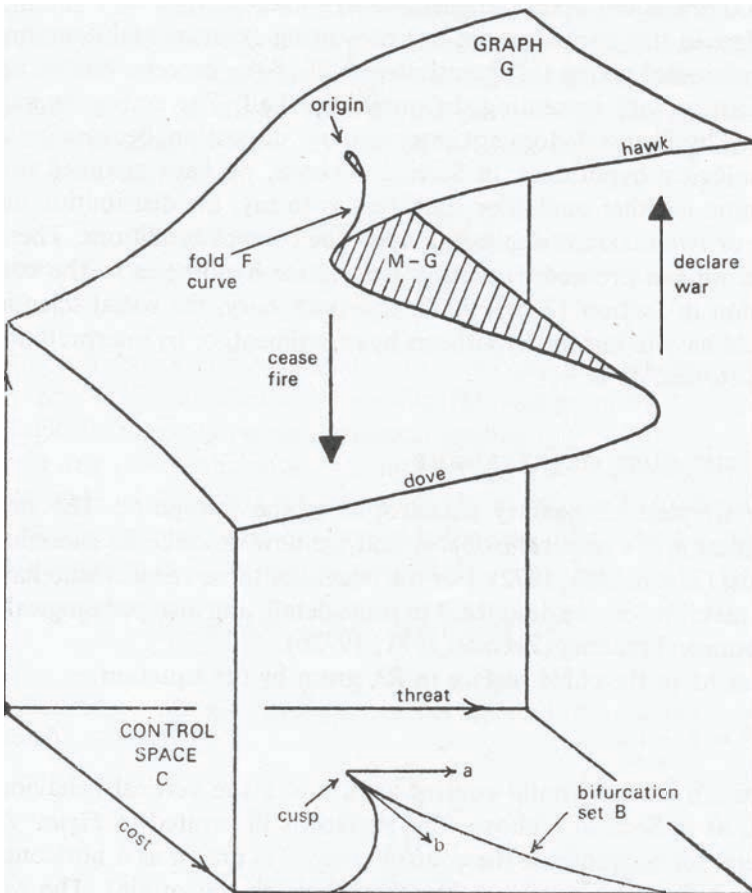
In sum, we have delineated the role of AI-algorithmic mechanisms of social media — especially Google/YouTube, Meta/Facebook, and Musk/Twitter — in the rise of political violence. The extensive research of social scientists on this role has been superbly collected and explained by Max Fisher in his book of 2022, *The Chaos Machine*. ...

In this book, *Schism*, we have built a connection between the chaos machine — social media — and chaos theory, an important and little known branch of 20th century math. This connection is a math model for political violence based on catastrophe theory, a branch of chaos theory adapted for applications to the social sciences. Our math model provides a map for strategies to mitigate the epidemic of political violence. The model might be applied to the parallel phenomena of gun violence, climate change, anti-abortion violence, and Covid pandemic management, etc.

Our goal, in this lengthy exercise in cybernetic thinking, is to provide clues to a safe and flourishing future for the biosphere and its human population.

After this long wander in three parts through the genesis and development of a new way of thinking based on mathematical modeling and cybernetics, I find myself with this final thought.

12 Existential Threats



The cusp catastrophe, from Zeeman, 1977; p. 330.
A model for the outbreak of war.

The policy-making departments of the US government might make use of this new way of thinking to improve the stability of the United States and thus democracies around the world.

But the chances of this new way of thinking being utilized to make better policies is slim, because the knowledge base required to effectively think this way is too meager. Our educational system has failed us, because the teaching of mathematics is so

poor.

This argument applies not only to political violence, but equally to the battles against all existential threats, including the four addressed specifically in this article.

Summary. Thus, one arm of our struggle for survival must address the regulation of social media and AI. Major aspects of the WWW involved are FaceBook, Twitter/X, YouTube, Google search, and access to AI. These, all aspects of tech, comprise the common denominator of all existential threats. Known remedies are rejected by irrational group-psychological dysfunctions. And it is here that AI might come to our rescue.

I. The Regulation of Social Media

Anna A. Eshoo, a Democrat, has been the U.S. representative from California's 16th congressional district since 1992. This district includes Silicon Valley, and Santa Cruz, where I live. She has been instrumental in congressional support for the internet since 2009, and efforts to defend against existential threats since 2015.

In the Fall of 2020, Eshoo and Tom Malinowski, Democratic representative from New Jersey, introduced the Protecting Americans from Dangerous Algorithms Act, legislation to hold large social media platforms accountable for their algorithmic amplification of harmful, radicalizing content that leads to offline violence.

About two weeks after the attack on the Capitol of January 6, 2021, she and Malinowski issued a press release entitled: "Following Attack on the Capitol, Reps. Eshoo and Malinowski Lead Dozens of Colleagues in Urging Tech CEOs to Fix Dangerous Algorithms, Address Spread of Extremism

and Conspiracies on Platforms.”

On the same day, they sent letters to the CEOs of Facebook, YouTube, and Twitter. Here is the full text of their letter to Twitter.

January 21, 2021

Mr. Jack Dorsey, CEO and Founder
Twitter, Inc.
1355 Market Street, Suite 900
San Francisco, California 94103

Dear Mr. Dorsey,

On Wednesday, January 6th the United States Capitol was attacked by a violent, insurrectionist mob radicalized in part in a digital echo chamber that your company designed, built, and maintained. Twitter is among the online services most responsible for spreading fringe conspiracy theories at scale and for radicalizing some of its users to commit real-world, physical violence.

Following the attack on the Capitol, Twitter took action to remove more than 70,000 accounts associated with the dangerous QAnon movement.¹ While we welcome this move, it came six months after Twitter first announced its crackdown on QAnon, which suggests a failure to meaningfully enforce the policy after it was enacted, and after years of allowing QAnon and other harmful, conspiratorial content to flourish on the site. Content moderation on a service that hosts hundreds of millions of tweets per day is a whack-a-mole answer to a sys-

temic problem, one that is rooted in the very design of Twitter.

The fundamental problem is that Twitter, like other social media platforms, sorts, presents, and spreads information to users by feeding them the content most likely to reinforce their existing political biases, especially those rooted in anger, anxiety, and fear. The algorithms Twitter uses to maximize user engagement on its platform undermine our shared sense of objective reality, intensify fringe political beliefs, facilitate connections between extremist users, and, tragically, lead some of them to commit real-world physical violence, such as what we experienced firsthand on January 6th.

Extreme, inflammatory tweets that trigger emotions like anger and fear often result in high levels of engagement on the platform, and Twitter's algorithm too often amplifies white supremacist, anti-Semitic, and other conspiracy-oriented material that can lead people to commit offline violence. Even if tweets fail to achieve mass scale virality, users can still become trapped in silos that intensify fringe views. Removing QAnon is not enough. Twitter must rethink and reengineer the fundamental architecture of its service which continues to allow for the frictionless spread of misinformation and radicalization at scale.

Leading up to the November 2020 elections, Twitter made a number of product changes designed to stem the spread of misinformation, including prompting users to quote tweets rather than simply retweet, removing recommendations in users' timelines and notifications, and adding context

to “Trends” in the “For You” part of its platform. Prompting quote tweets over retweets reduced overall retweets and quote tweets by 20%, which the company found “slowed the spread of misleading information by virtue of an overall reduction in the amount of sharing on the service.”² This change was reversed following the election.

Experts have rightly suggested that the platform needs to make permanent, fundamental design changes to limit the spread of harmful content, such as halting recommendations, limiting shares, and adding a circuit breaker-like function to slow the spread of the most viral and potentially dangerous content.³ One independent disinformation expert bluntly stated, “Twitter doesn’t yet seem to understand it’s a platform optimized for radicalization.”⁴ It is our hope that Twitter will immediately make permanent changes to limit the spread of misinformation and other forms of harmful content, and that the company will begin a fundamental reexamination of maximizing user engagement as the basis for algorithmic sorting and recommendation.

Sincerely,

Tom Malinowski, Anna G. Eshoo
Members of Congress

Thus far, the Protecting Americans from Dangerous Algorithms Act has not been passed. Nor have the letters to FaceBook, YouTube, and Twitter had any effect.

J. Conclusion

Our initial 1994 enthusiasm for the World Wide Web in overcoming our global problems might be validated by AI, helping us to solve the dangers of nuclear winter, global warming, pandemics, and other existential challenges, but only if we can defeat the new dangers inherent in the evolution of AI itself. *Medicine is poison in small doses!*¹²

The positive potential of AI outweighs the negative. However, the rate of approach of the negative overwhelms that of the positive. This comports with the advice of some AI experts to intentionally slow AI development. *But is that actually possible ??*

The Club of Rome wrote in 1971:

The predicament of mankind is that we can perceive the individual symptoms and the components of profound social problems, but we are stymied in our efforts to comprehend the total situation and develop global solutions.¹³

And half a century later, the predicament has metastasized.

References

Abraham, Ralph H. (2023). *Schism: The Madness of Crowds, Toxicity of Social Media, Social Polarization, and Political Violence; A Cybernetic Approach*. Rhinebeck, NY: Epigraph Books.

Castells, Manuel (2010). *The Rise of the Network Society*,

12 This aphorism was a favorite of Baba Hari Dass, late of the Mount Madonna Yoga Retreat Center.

13 Meadows, 1971; pp. 1-2.

Second Edition. Chichester, UK: John Wiley.

Clarke, Richard A. and Robert K. Knake (2010). *Cyber War: The Next Threat to National Security and What to Do About It*. New York: Harper Collins.

Fisher, Max (2022). *The Chaos Machine: The Inside Story of How Social Media Rewired our Minds and our World*. New York: Little, Brown.

IPCC Core Writing Team (2023). *Sixth Assessment Report: Climate Change 2023 (AR6)*. Geneva, CH: IPCC.

Meadows, Dennis (1971). *The Club of Rome Project on the Predicament of Mankind*. Cambridge MA: MIT.

Rees, Martin (2003). *Our Final Hour: A Scientist's Warning*. New York: Basic Books.

Russell, Stuart, and Peter Norvig, 2022. *Artificial Intelligence: A Modern Approach, Fourth Edition*. New York: Pearson.

Zeeman, E. Christopher (1977). *Catastrophe Theory: Selected Papers 1972-1977*. Reading, MA: Addison-Wesley.