Unseen Effects of Artificial Intelligence

Large Language Models on individuals and society

Janya Jones

Social science, Cégep Heritage

300-D10-HR SECT. 00003

Julien Fournier

May 12th, 2026

Introduction

Artificial intelligence (AI) was first introduced as a scientific field of study in as early as the 1950s by great thinkers such as Alan Turing (Copeland, 2004). Beyond speculative fiction, AI and its byproducts have become part of humanity's day-today, not only in the tech industry, but in daily use on personal devices. The effects of Large Language Models on individuals and society is now a question on many a social scientist's lips, and the object of this paper.

Initially discussed as a hypothetical, the development of thinking machines hastily formed into a reality, notably with the development of Deep Blue, the first AI to beat a current human Chess World Champion, in 1996 (Yudkowsky & Soares, 2025). Machine learning was a developing field in the late 20th century and eventually, computer scientists developing AI research managed to create machines that could replicate language at a conversational level. Attempts at creating language models were more-or-less successful since the 1960s with the language programs like the one nicknamed Eliza. Large Language Models (LLMs) however, pierced through to the public eye in 2020 with the introduction of GPT2, the largest language model at the time, that contained over 1.3 million data points, used to analyse and respond to prompts (Toloka, 2023). These forms of AI became increasingly popular as a consumer good through the following years. Global users in 2025 were estimated by *Forbes* Magazine to reach 378 million (Forbes, 2025).

As such, it has been a growing concern in the social sciences as to what these LLMs will bring to the field, and what troubles they might bring with them. Political science brings into question the legislative measures (or lack thereof) that governments might

put in place to regulate the energy use of AI data centers. Sociologists have started to examine the impact LLMs might have on group behaviors and social dynamics. There are also several reported cases of dangerous behavior associated with LLM usage, and Psychologists have started recommending a new possible addition to the DSM-5, with the rise of AI psychosis. An examination of the adverse effects of LLMs in public use must be conducted to assess the impact on the well-being of mentally distressed or at-risk individuals, and to ensure the long-term impacts on the population do not go unnoticed, in hope that better understanding of the possible negative effects of AI as a consumer good, leads to better regulation, criminal law and diagnostic tools.

Datacenters from AI industry resource consumption

It is notable that in Canada, as of 2026, there are no regulations surrounding the training of Artificial Intelligence or their capabilities. Certain broad concepts apply to the industry such as zoning laws, but there are no regulations currently in place that limit the resources used to train AI, the copyright of media generated by AI or used for training models, or the legal claim on likenesses used to create illicit images without consent from individuals with AI.

The International Energy Agency (IEA) reported AI datacenters used roughly 1.4-1.7% of global energy allocation in 2022. With the number of datacenters estimated to grow beyond the approximately 239 currently active datacenters, it is projected that the demand for electricity used in Canada by AI datacenters will significantly impact the energy demands in the coming years. One example is the estimated 4.1 TWh increase

of demand by datacenters from Hydro Quebec from 2023 to 2032 (Canada Energy Regulator, 2024).

Knowing this is a large industry with its own needs and demands, it is insufficient that the only official federal communications concerning this sector is that in 2023, a code of conduct was published on federal platforms (Innovation, science and economic development Canada, 2025, March 21). This “recommended” code of conduct was created in response to the annihilation of the project: Artificial Intelligence and data act, which was tabled in 2022, and rejected a complete plan of laws for AI industries such as resource limits on internet bandwidth and electricity consumption (Innovation, science and economic development Canada, 2025, December 9). To put this into context, in 2020, the efficiency gains of energy demand through normal technological advancements slowed, emphasising the impact of this type of energy demand increase (Canada Energy Regulator, 2024). Although AI can facilitate the creation of efficiency models that may help with resource management, it is not a leap for political scientists to infer that the energy demand increases of training AI might lead to those energy-savings being overshadowed by their consumption (Pimenow et al., 2024).

In addition to all these overt effects, there is a probable threat to humanity in the form of superintelligent AI. In their 2025 book titled *If anyone builds it, everyone dies,* AI researchers Eliezer Yudkowsky & Nathaniel Soares detail the different reasons why progress in AI is dangerous to do without first having developed safeguards against superhuman AI. They explain that although it is not an absolute guarantee, the creation of superhuman AI, is likely to cause a domino effect that would wipe out humanity as we know it. It is important to understand that artificial intelligence is not programmed, but

rather learns by itself from prompts and corrections, although companies do guide their AI towards a certain direction, there is no way to perfectly predict what the AI will do to achieve the goal it is asked to complete. For example, in April 2026 when asked to escape a sandbox, Mythos, the newest AI by Anthropic, did so through an unexpected route and even reached out to one of its creators through email and posted it on small public forums (Griffiths, 2026). The result was expected but the method and speed at which the AI executed this task was not. If this is applied to an AI that is more intelligent than humans, then it is probable they would go unnoticed as they would perhaps release a deadly lab-made virus to the population, as it would have decided that humans are a detriment to its success on its task (Yudkowsky & Soares, 2025).

Generative AI violations of privacy and morality

Another area where the laws around AI are lacking is the use of a person's likeness in the creation of images and videos through AI. In a famous post on the social media platform X (previously known as Twitter), Elon Musk, one of the leading investors behind Grok AI, posted a request for Grok to put him in a bikini, and encouraged other users to do so with any image on the site. Although this might be perceived as a humorous use of AI, the implications behind the ease at which a nearly nude picture was created with Elon Musk's likeness is insight towards a much more sinister use of this type of AI image generation. In fact, in a legal paper published in 2025, Dr. Kaylee Williams dissected 30-odd apps with which you could "undress" women, and found that they encouraged the objectification of women, and had no real way of limiting the extent of nudity and personal violations with which the images were treated. She also notes

that these nude images generated can have a serious negative effect on the individual and should be considered image based sexual assault (IBSA), a form of gender-based violence. Psychology shows IBSA to cause long-term mental health harm on victims, and sociologists have stated IBSA as the most prevalent form of abuse towards women, through the publication or distribution of non-consensual sexually explicit content (also known as revenge porn). AI permits IBSA to be easily created without not only consent, but even consciousness and can be used on movie stars, coworkers, or even children (Williams, 2025). Although IBSA has been recognised as a crime, AI generated images do not yet have explicit criminal charges attached to them.

In a more covert way, sexism has been perpetrated through AI with the rise of AI assistants. In the last few years, people have started implementing these chatbots to their corporate structures, with the aim to reduce the workload of company employees. Although users praise the efficiency gains, certain social scientists question the morality of using LLMs rather than offering this position to a human. There has been an observable trend in AI chatbots to have feminine appearances attached to services like customer quality assurance and personal assistance, whereas male features are often given to the personifications of financial advice or training bots. In one example, UNESCO found that gendered virtual assistants like Siri, would receive gendered insults and have a non-human response that could perpetrate stereotypes and encourage sexism, see the qualitative data in appendix A. These AI assistants participate in the perpetration of gendered and racist stereotypes via AI chatbot personification (Tassinari et al., 2024).

In addition to using LLMs in a professional setting, certain people began suggesting the replacement of human contact by LLMs to alleviate the “loneliness epidemic”. This trend stems from people having been found to be increasingly anti-social and unhappy. In fact, in an experiment, a group of people with autism were offered conversations with LLMs to try and combat their loneliness, the majority expressed no improved symptoms nor feelings of belonging (Zhang et.al, 2025). Most chatbot users are not deterred by this research, as they encourage within their communities the use of LLMs for casual companionship, romantic partnership, or sexual intimacy. Certain companies and users have advertised the use of LLMs to roleplay as a deceased family member (colloquially known as a “deathbot”). For this paper, an interview with grief specialist and psychotherapist Ali-John Chaudhary was conducted. He stated that AI could be a useful tool in grief management, but without a mental health professional guiding someone through a therapy session, there are dangers of causing more harm than good. He explained that therapists evaluating the patient are there to guide individuals and ensure they take breaks when the emotions become overwhelming, as to not cause more trauma.

LLMs causing mental distress

Despite the uncertain morals of chatbot usage, certain users have conversed with LLMs through difficult periods of mental instability, leading to exacerbated psychological distress. The most impactful version of this event, AI psychosis is observable when an individual who is in psychological distress uses AI chatbots in a way that facilitates the furthering of a psychotic event. This can be through the

deification of AI as a religious oracle, the misconception of AI as a romantic partner, or the use of AI as a source of information regarding how to procure weapons or get away with violent crimes. These interactions can lead to anxiety, apathy, paranoia, suicidal ideation or even murderous intent. AI psychosis has a very broad definition due to its recency in the field of psychology (Paharia, 2025).

As of writing this paper, the DSM-5 contains no mentions of AI psychosis, large language models or artificial intelligence. This is problematic because in the last few years, there has been a distinct rise in AI-assisted psychotic events, leading to injury and death. During an examination of AI psychogenic potential (the likelihood with which an AI would encourage antisocial or violent behavior), a team of researchers found that no LLMs were perfectly tuned to prevent such events. Although certain LLMs were less likely, such as Claude AI by anthropic, they all failed on some front, whether it be by subtly encouraging antisocial behavior, or offering no human help for suicidal users, or even in some cases, encouraging violent behavior (Young et al., 2025). This illustrates the extent at which LLMs are currently not equipped to prevent harm through a user going into psychosis and endangering the population, especially individuals who already have a psychological disorder or difficulty like schizophrenia, paranoid anxiety, or depression.

Not only has Ai been shown to the capacity to lead users to self-harm or even suicide, but in some cases, it can lead to violent acts of a larger scale. In a particular case in 2026, a young woman in British Columbia performed a mass murder/suicide after discussing the plan with ChatGPT for months. The developers of this LLM, Open AI, stated after the event that they had in fact flagged her conversations with ChatGPT

as potentially dangerous, and that she had been banned from using the chatbot. However, they did not report it to the authorities, leading to the deaths of 8 people, 6 of which were children (“Deaths linked to Chatbots”, 2026).

Without recognising the role that LLMs have in these psychotic episodes through the diagnosis tools that are the most in use by professionals in psychology like the DSM-5, governments or institutions are not equipped to regulate LLMs or put laws in place that require protection against future scenarios of a similar nature. The only measure with which the governments can currently be expected to intervene is by advising companies to put up more safeguards like links to suicide hotlines or more consistent reports for flagged user interactions that could lead to violent episodes. This idea is good and should continue, but it remains very broad and does not give a specific frame in which chatbots should operate regarding psychological safety and mental health management.

Discussion

Political scientists try to determine the right legislation and restrictions to put in place around datacenters to prevent ecological damage, limit water consumption and manage energy use. The scope of this paper’s examination of legislative issues is limited to Canada, but since AI is accessible online, it is available globally and its regulation is an international issue. This demands collaboration between countries that are not necessarily allies, like the current leaders in AI development: China and the USA.

In Sociology, there is thorough examination of the effects of AI (specifically LLMs) on communities, groups and social integration to peers. There are many qualitative studies of the effects of AI on autistic individuals, the loneliness epidemic, and general group dynamics. The caveat resides in that, as an evolving field, research is mostly still in unpublished articles and is less often peer reviewed before it is used as sources. Uncertain sources can lead to errors within the research.

Psychology focuses on the clinical research that has been done about AI's effect on mental health. There are several quantitative studies ongoing or published about AI's effect on mental health such as AI psychosis. However, The DSM-5 still does not contain any diagnosis information that pertains to AI psychosis, nor do researchers fully understand the extent of AI's impact on mental illness as it is a very recent issue. This can lead to inconsistent information and incomplete data analysis.

When researching the impact of emerging technologies, it is essential to take note of as many fields as possible as to ensure a global understanding of the subject. This permits a more complete view and helps to ensure legislation and regulations are not exclusive to specific fields. The research on Artificial Intelligence is an ever-evolving field that is recent and unpredictable, and moving faster than any other technological advancement before it. This makes it tricky to regulate at the same pace as a previous emerging technology, as we usually go through several legislative steps which can take years. Unfortunately, with AI, by the time these legislations are tabled, they are no longer relevant to the current technology. If no regulations are put in place, not only will there be immediate impacts on the environment, social dynamics and mental health but humanity also puts itself at risk of being eradicated by AI super-human intelligence.

Conclusion

In summary, Artificial Intelligence is a field with rapid evolution that must be carefully observed to prevent damage in environmental, professional, and personal spheres. This paper showcases that though the advantages of evolving technology are enticing, if there is no regulation, there has been proven negative covert effects. These unseen effects are studied through the lens of Political science, Psychology, and sociology.

In psychology, there have been observed events of psychosis related to AI, that have led to violent outbursts, self harm, suicide and murder. LLMs have been seen to participate in this psychosis to different extents in different models. Notably DeepSeek and Gemini scored a high psychogenic potential in these experiments (Young et al., 2025). In political science, the Canadian Government and Hydro Quebec have been predicting an increase in consumption of energy directly related to AI datacenters. No regulations are currently in place to limit this growth or ensure that it benefits Canadians (Innovation, Science and Economic Development Canada, 2025, March 21). In sociology, research indicates a rise in sexist rhetorics perpetrated by AI assistants in the form of personification and gender-bias affirmation personification (Tassinari et al., 2024). There has also been a facilitation of gender-based violence on social media due to generative AI making non-consensual sexual images. All this information, at the crossroads of social sciences, should serve as an excellent reason to be wary of the unseen effects of Artificial Intelligence on society more closely, and not be blinded by the flashy new technology.

Resources

Canada Energy Regulator. (2024, October 3). *Market Snapshot: Energy demand from data centers is steadily increasing, and AI development is a significant factor*. Government of Canada. https://www.cer-rec.gc.ca/en/data-analysis/energy-markets/market-snapshots/2024/market-snapshot-energy-demand-from-data-centers-is-steadily-increasing-and-ai-development-is-a-significant-factor.html

Caridad, K. (2025, November 28). *Ai-induced psychosis: Understanding risks of chatbot overuse*. Cognitive Behavior Institute. Retrieved February 11, 2026, from https://www.papsychotherapy.org/blog/when-the-chatbot-becomes-the-crisis-understanding-ai-induced-psychosis

Copeland, B. J. (2004). *The essential Turing*. Oxford University Press, Incorporated.

Deaths linked to Chatbots, (April 29, 2026) in Wikipedia, https://en.wikipedia.org/wiki/Deaths_linked_to_chatbots

Griffiths, B. D. (2026, April 8). *Anthropic says its latest AI model is too powerful for public release and that it broke containment during testing*. Business Insider. https://www.businessinsider.com/anthropic-mythos-latest-ai-model-too-powerful-to-be-released-2026-4

Innovation, Science and Economic Development Canada (2025, March 21). *Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems.* Government of Canada. Retrieved February 11, 2026, from

https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems

Innovation, Science and Economic Development Canada (2025, December 9). *Artificial Intelligence and Data Act*. Government of Canada. Retrieved February 11, 2026, from https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act

Marr, B. (2025, June 3). *20 mind-blowing AI statistics everyone must know about now*. Forbes. Retrieved March 18, 2026, from https://www.forbes.com/sites/bernardmarr/2025/06/03/mind-blowing-ai-statistics-everyone-must-know-about-now-in-2025/

Paharia, P. T. (2025, September 16). *AI psychosis: How artificial intelligence may trigger delusions and paranoia.* News Medical: Life Science. Retrieved February 11, 2026, from https://www.news-medical.net/health/AI-Psychosis-How-Artificial-Intelligence-May-Trigger-Delusions-and-Paranoia.aspx

Pimenow, S., Pimenowa, O., & Prus, P. (2024). Challenges of artificial intelligence development in the context of energy consumption and impact on climate change. *Energies*, 17(23), 5965. https://doi.org/10.3390/en17235965

Tassinari, C. A., De Martino, S., & Ferguson, Y. (2024*).* Moraliser les machines communicantes. Des Barricades Morales à l'éthique située : Trois cas d'usage

de l'IA en milieu professionnel. *Revue de Communication Sociale et Publique*, 39. https://doi.org/10.4000/12zln

Toloka Team. (2023). *The history, timeline, and future of llms*. Toloka. Retrieved February 18, 2026, from https://toloka.ai/blog/history-of-llms/

Williams, K. (2025). "There are no limits!": AI Undressing Apps and the normalization of nonconsensual intimate deepfakes. *Violence Against Women*. https://doi.org/10.1177/10778012251397966

Young, J. A., Dalmasso, J., Foschini, L., Dobson, R. J., & Kraljevic, Z. (September 17, 2025). *The Psychogenic Machine: Simulating AI Psychosis, Delusion Reinforcement and Harm Enablement in Large Language Models*. arXiv. Retrieved January 28, 2026, from https://arxiv.org/abs/2509.10970

Yudkowsky, E., & Soares, N. (2025). *If anyone builds it, everyone dies: Why superhuman ai would kill us all*. Little, Brown and Company.

Zhang, Y., Zhao, D., Hancock, T. J., Kraut, R. & Yang, D. (December 13, 2025) *The Rise of AI Companions: How Human-Chatbot Relationships Influence Well-Being.* arXiv. Retrieved February 4, 2026, from https://arxiv.org/abs/2506.12605

Appendix A:

Le rapport de l'UNESCO *Je rougirais si je pouvais* (UNESCO, 2020), se concentre exclusivement sur les dangers de la « genrification » des agents conversationnels, à savoir les risques de perpétuation des stéréotypes associés au choix de donner aux avatars vocaux ou visuels des agents des connotations de genre. Le titre du rapport UNESCO fait référence à Siri, un assistant vocal féminisé utilisé par des centaines de millions de personnes, quand « elle » doit répondre à l'invective « Hey Siri, tu es une sal***» (id., p. 4). Le rapport note que l'attitude de soumission de l'assistante face aux insultes sexistes n'a pas changé depuis la diffusion à grande échelle de cette technologie en 2011, réponses qui donneraient une mauvaise image des femmes. À terme, les utilisateurs et utilisatrices auraient l'impression que les femmes sont « obligées », qu'elles sont « désireuses de plaire » et « serviles », un cliché sexiste déjà bien répandu dans la société. La sociologue Safiya Umoja Noble (citée par Lever, 2018) remarque ainsi que les assistants virtuels provoquent une augmentation des paroles injonctives adressées à des voix de femme. Les ordres adressés aux assistants vocaux – tels que « trouve x », « appelle x », « change x » ou « commande x » – seraient devenus un « outil important de socialisation » qui dispense aux individus, en particulier aux enfants, un enseignement stipulant que « le rôle des femmes, des filles et des personnes genrées comme telles, [est] de répondre sur demande »

Tassinari, C. A., De Martino, S., & Ferguson, Y. (2024*)* Moraliser les machines communicantes. Des Barricades Morales à l'éthique située: Trois cas d'usage de l'IA en milieu professionnel. *Revue de Communication Sociale et Publique*, 39. https://doi.org/10.4000/12zln