WEEK 4 CASE STUDIES

Study 1: "Godfather of AI" Geoffrey Hinton Warns of the 'Existential Threat' of AI"

Study 2: "Generative AI Has an Intellectual Property Problem"

Hannah Graham [1, 2,]†

1 Northwestern University

2 GitHub: https://github.com/hannah-r-graham

† hannahrg24@gmail.com

Table of Contents

Case 1: "Godfather of AI" Geoffrey Hinton Warns of the 'Existential Threat' of AI"

The Data:

The primary source and only source for this case study is an interview with Geoffrey Hinton a year ago from writing this paper and at a time that was only a few months after Hinton left Google. The interview did not emphasize much data, but more on the philosophical issues surrounding AI, its capabilities, and how to best develop guard rails while developing models and using AI assets in the real world. Hinton referenced a few scenarios but most notably intuition: cats are more feminine, and dogs are more masculine according to human and AI intuition. Hinton stated this is what makes him think that AI is learning quicker and better than humans can, and what makes them such a threat to society. This was by no means a full-deep analysis, but Hinton is as expert as it gets in the AI world, so the examples he puts forth are something we should heed. As such an expert, I feel its safe to assume that Hinton has done much further analysis than what is mentioned in this brief interview.

Relevance to Data Governance:

Hinton described AI and its ability to surpass human intelligence as an existential threat: AI will become more intelligent than us and gain control – which includes taking away jobs. This crosses into legal, ethical, and management issues. Ethical because what are the downsides of humans not working or having a job due to technology. Hinton believes that this will be a Mutually Assured Destruction issue that companies will seek to mitigate, resulting in people collaborating, especially large companies, to avoid replacing humans completely. Do I personally believe that this is true, and companies will be able to think this far ahead? No. I think that is overly optimistic, but I hope to be proved wrong.

This is a legal issue because if someone does replace another person's job with AI, should there be repercussions. Last but not least, this is also a management issue because Hinton believes that the key to navigating these dynamic times is managing the development of the models with tests and understanding of how to keep the model under control.

Relevant Legislation:

Hinton suggests legislation around AI should be similar to many governments laws pertaining to fake currency. Producing fake currency and distributing fake currency is illegal in many countries, including the United States. He thinks the same should apply to AI content wherein any fake images, fake videos, and fake voices are required to have a disclosure or label that AI was used or else 10 years in prison for the creator or distributor (for example). Otherwise, we will be overwhelmed with fake content, and we won't know what is real or what is not. What is truth and what is lie. No such federal laws, at least in the USA, are actively doing that. Despite no legal requirement, companies have begun changing internal policies to require transparency of when AI is used, either to create artifacts such as images/videos/voices or on peoples data, speaking from personal experience.

Hinton also suggests that it's not only a philosophical issue but that people who are developing these models need to be doing tests and experiments. Hinton believes we can't stop the development of AI, it is too valuable and impactful, but we can learn how to manage the AI while actively developing the models.

Data Governance Solution:

AI has many, many benefits both socially, ethically, and financially. Therefore, the development of AI is not an if or should. It's happening whether or not we want or agree with it. Hinton believes the same. He is not pessimistic nor optimistic about the future of AI but says it's an unknown and "uncertain time". To help combat this, he suggests transparency laws that require AI generated artifacts to be transparent about their use of AI and a repercussion of 10 years in prison if the creator/distributor does not adhere. Hinton also suggests that developers and researchers actively test and experiment with models to maintain understanding of how they work and can be controlled in various scenarios. I agree with both suggestions. I think these suggestions are common sense, and I have even seen companies such as Microsoft or places like Medium.com implement transparency actions like "made with AI" under AI content as part of a culture shift.

Case 2: "Generative AI Has an Intellectual Property Problem"

The Data:

This case study used various sources and examples including Andersen v Stability AI et al., Getty v Stable Diffusion, as well as more historical precedence not related to AI like: Google defending web scraping for use in search, and a case against Andy Warhol foundation for using a licensed image of Prince. Google won by claiming its "transformative use allowed for the scraping of text from books to create its search engine" (Appel et al, 2023). The Warhol, Anderson and Getty cases are still to be determined. I believe for analysis, this is a great start at what it means to have a transformative work versus copyright infringement. In art it's a fine balance and up to interpretation at the moment.

Relevance to Data Governance:

copyright infringement on art is focused primarily on the ethical and legal bucket. Ethics, deem that someone's art is their own and should be respected as such. As an artist, go find your own voice. Don't steal from others…. but that's exactly what AI is. Its output is based on training/historical data. Its artwork is derived from others artwork by design. Legally this is an issue because in art, how do you deem something duplicative of another's work if it's not an exact match? If AI changes only a few pixels of an image, is it still a copy or is it considered a brand-new work? How would the law even regulate/enforce that. It's a very slippery slope with not necessarily a clear answer.

Relevant Legislation:

The main "law" discussed in this article is of the Fair Use Doctrine "…which allows copyrighted work to be used without the owner's permission "for purposes such as criticism (including satire), comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, and for a transformative use of the copyrighted material in a manner for which it was not intended" (Appel et al, 2023).The big question, as mentioned above, is what deems transformative? When is art changed enough to not deemed a copy?

For mitigation, the authors suggest 1) Developers ensure they have received permission from owners to leverage IP, by licensing or sharing profits. 2) Companies should require owners

to opt-in for their works being used by AI, not required to opt-out. 3) Content creators should check regularly to see if their works have been used in AI and are out in the world in a way they did not intend. Finally 4) businesses should require any AI platforms have proper licensing from data they used to train on.

Data Governance Solution:

I think the authors suggest reasonable solutions for the most part, including asking the law to be clearer on what constitutes copyright infringement with art. However, the suggestion that owners should hold large companies accountable for stealing their works is ridiculous, and blames the victim while undermining years of precedence in copyright law. It's not a reasonable solution. However, I do agree that companies should offer opt-in for art works and create a symbiotic relationship with artists so the benefits aren't so one-sided. One solution that the author did not mention is transparency – let's say an artist does give permission for AI to use their artwork, then I believe that artwork should give credit to that original artist.

References

Amanpour and Company. 2023. "Godfather of AI" Geoffrey Hinton Warns of the 'Existential Threat' of AI." May 9, 2023. YouTube Video, 18'8".
https://www.youtube.com/watch?v=Y6Sgp7y178k

Appel, Gil, Juliana Neelbauer, and David A. Schweidel. 2023. "Generative AI Has an Intellectual Property Problem." Harvard Business Review, April 7, 2023.
https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem

Appendix A