

Narrative, Creativity, and Open Source in “Back to Eden”

by Scott H. Hawley

I. Introduction: An Alternate Future

In Beth Earnest’s short story, “Back to Eden,” classic sci-fi tropes provide the framework for a story steeped in political intrigue and ideological power struggles driven by transformative technology. Science fiction has long served as a platform for political commentary, with examples including Ursula K. Le Guin’s alternative property rights in “The Dispossessed” and Robert Heinlein’s civic duty-driven suffrage in “Starship Troopers.” Against this backdrop, an independent religious community takes a seemingly harmonious middle path – an optimistic deviation that almost defies the genre’s penchant for uncovering a dystopian flipside.

Another sci-fi staple in “Back to Eden” is the exploration of the counterfactual question, “What if things went the opposite way?” as in the alternate histories of Harry Turtledove and Philip K. Dick. It imagines a world where AI development is not led by left-leaning hubs like San Francisco and Seattle, known for [producing chatbots with systemic left-wing bias](#)¹, but by conservatives and Southerners, whose only technical claim to fame in our timeline is a [failed attempt to create a social media platform to protect free speech](#)². In this alternate reality, the “New Con,” a not-so-subtle play on “Neo Con,” has harnessed technology for control, using weaponized AI for various means, including the restriction of free speech.

Readers initially dismayed by motifs that caricature conservatives, the South, and the military will find depth as the story eventually reveals its use of an “unreliable narrator.” The main character, also fulfilling the role of a “neophyte” protagonist, serves as a vehicle to introduce the audience to the Eden community’s unique blend of faith and technology. Her perspective challenges us to consider these intersections, subtly coaxing secular, liberal-leaning readers to entertain a religious viewpoint on technological integration.

In the following sections, I will delve into select topics: the questionable narratives proliferating online, our tendency to humanize AI, the burgeoning field of AI-driven creativity, and the principles of Open Source, concluding with our collective responsibility in crafting the future.

¹ Fabio Motoki, Valdemar Pinho Neto, and Victor Rodrigues, “More Human than Human: Measuring ChatGPT Political Bias,” *Public Choice*, August 17, 2023, <https://doi.org/10.1007/s11127-023-01097-2>.

² Jon Brodtkin, “Parler Shuts down as New Owner Says Conservative Platform Needs Big Revamp,” *Ars Technica*, April 14, 2023, <https://arstechnica.com/tech-policy/2023/04/parler-shuts-down-as-new-owner-says-conservative-platform-needs-big-revamp/>.

II. The End of Collective Consciousness

The fragmented world of “Back to Eden” mirrors our own, extending beyond state borders to the contentious, contextualized, or even fabricated digital “worlds” we inhabit. In the story, Brother Henry asks, “Who believes anything on the internet today?” This is not merely a remark about the [growing fidelity of AI-generated fakery that fools audiences and journalists](#)³ alike. Rather, the recommendation systems powering our social media and news feeds propose to teach us about the world, yet these can result in [“echo chambers” and “filter bubbles”](#)⁴ reinforcing our existing biases rather than expanding our horizons. The apostle Paul's prophecy in [2 Timothy 4:3](#)⁵ eerily prefigures our current digital reality, where individuals gravitate towards content that echoes their preconceptions:

“For the time will come when people will not put up with sound doctrine. Instead, to suit their own desires, they will gather around them a great number of teachers to say what their itching ears want to hear.”

This selective exposure [contributes to](#)⁶ political polarization, and while some have promoted the use of [AI to present news from various perspectives](#)⁷, these have yet to become mainstream.

Previously automated systems merely filtered existing content, but the advent of "Generative AI" allows users to create entirely new content on demand. Already, we see [ads showing teenagers generating music for car trips](#)⁸ rather than listening to the radio. Generative [sitcoms](#)⁹ and [video games](#)¹⁰ have already appeared, and it's even been suggested that [generative “reality shows”](#)¹¹ could find a market. In such cases, one can be confident that the user's experience is not merely biased by societal groups but individually *unique*. In this landscape of customized content, where

³ “How Google, Microsoft, and Adobe Are Trying to Stop AI from Flooding the Internet with Garbage - Vox,” accessed January 15, 2024, <https://www.vox.com/technology/23746060/ai-generative-fake-images-photoshop-google-microsoft-adobe>.

⁴ Amy Ross et al Arguedas, “Echo Chambers, Filter Bubbles, and Polarisation: A Literature Review | Reuters Institute for the Study of Journalism” (Oxford, UK: Reuters Institute for the Study of Journalism, January 19, 2022), <https://doi.org/10.60625/risj-etxj-7k60>.

⁵ “2 Timothy 4:3.” The Holy Bible, New International Version. (1984). Bible Gateway. Available at: <https://www.biblegateway.com/verse/en/2%20Timothy%204%3A3>

⁶ Petter Törnberg, “How Digital Media Drive Affective Polarization through Partisan Sorting,” *Proceedings of the National Academy of Sciences* 119, no. 42 (October 18, 2022): e2207159119, <https://doi.org/10.1073/pnas.2207159119>.

⁷ Ground News. Available at: <https://ground.news/>

⁸ Suno [@suno_ai], Tweet, *Twitter*, December 20, 2023, https://twitter.com/suno_ai/status/1737299056290910464.

⁹ Frumess, *Seinfeld - Nothing, Forever*, 2023, <https://www.youtube.com/watch?v=M6mD9YzVbZI>.

¹⁰ Josh Norem, “Nvidia Brings Generative AI to NPCs,” *ExtremeTech*, May 30, 2023, <https://www.extremetech.com/gaming/nvidia-brings-generative-ai-to-npcs>.

¹¹ Horrific ML Ideas [@ml_ideas], “The First Personalized, Prompt-Based Generative TV Shows Should Not Be Sitcoms, Anime, or Any Other Narrative Genre...They Should Be So-Called ‘Reality’ Shows:...,” Tweet, *Twitter*, October 27, 2023, https://twitter.com/ml_ideas/status/1717698322046026052.

AI-generated media caters to the individual rather than the masses, it becomes crucial to distinguish the technology's role as a tool rather than a substitute for human interaction.

III. Steering Clear of Anthropomorphism

One aspect of the story that I find refreshing is its complete avoidance of the tired trope of AI anthropomorphism, which is not limited to science fiction but appears as a cognitive error that dogs even careful attempts to consider human-computer interaction. Blessedly, the AI systems referred to in “Back to Eden” appear only in the contexts of their *uses by humans*. These systems are not themselves sentient, have not formed their own governments or demanded “[rights](#)¹²,” etc. In this sense, the world of “Back to Eden” mirrors our own, inviting us to ponder realistic near-term AI issues. [Zeynep Tufekci captures this sentiment well](#)¹³:

“Ask not what Artificial Intelligence will do if it becomes human-level—ask what humans will do with artificial intelligence in the meantime.”

There is something unavoidably anthropomorphic in the very conception of AI, for which we adopt the [folklore definition](#)¹⁴ of “computers doing what we *used to think* only humans could do.” In discussions and marketing of AI systems, anthropomorphic metaphors are commonly used to succinctly provide new investors and users with a vision for the product's roles by leveraging familiar grounds of human experience. However, borrowing language in this way can sometimes be extremely misleading. We may use the word “computer,” despite it formerly denoting a human profession, yet no one seems to have a problem with it. Other language, however, is not so universally accepted. One often hears objections in the form of, “You shouldn't call say the machine is doing X because only human beings can truly engage in X,” where “X” might be “learning,” “training,” “understanding,” “reasoning,” etc.

How do we establish criteria to discern when the adoption of traditionally human-specific terms for machine functions is suitable, versus when it's a case of misleading anthropomorphism? I will share two ideas from the AI Ethics literature. The first criterion, [as highlighted by Rob Wortham, is transparency](#)¹⁵ that equips users to gain an accurate understanding of an AI's capabilities and limitations. Transparency is not so much about exposing the innards of a system

¹² François Fleuret [@francoisfleuret], “I Have the Deepest Respect for @DavidDeutschOxf, but This Is Weird...” Tweet, *Twitter*, December 20, 2023, <https://twitter.com/francoisfleuret/status/1737613256867668036>.

¹³ zeynep tufekci [@zeynep], Tweet, *Twitter*, October 11, 2017, <https://twitter.com/zeynep/status/918181775406391296>.

¹⁴ Scott H. Hawley, “Theopolis Monk: Envisioning a Future of A.I. Public Service.” *The Transhumanism Handbook*, edited by Newton Lee, Springer Verlag, 2019, pp. 271–300. *PhilArchive*, <https://philarchive.org/rec/HAWTME>

¹⁵ Robert H. Wortham, *Transparency for Robots and Autonomous Systems: Fundamentals, Technologies and Applications*, IET Control, Robotics and Sensors Series 130 (London: The Institution of Engineering and Technology, 2020).

that may be too complex for users to grasp, rather it involves [finding the right balance of information](#)¹⁶ sufficient to help users [form accurate mental models](#)¹⁷ of the machine’s internal states or at least to acquire some reasonable confidence in predicting the system’s behavior. When people are unable to do this, they often overestimate the system’s performance or “[overattribute](#)¹⁸” capabilities that don’t exist. A classic case of overattribution was seen in public reactions to the defeat of chess grandmaster Gary Kasparov by the IBM computer system Deep Blue: although it could only play chess and had no capacity to “learn” to do anything else, there were widespread fears about the robots taking over. Unfortunately, overattribution often goes unnoticed until consequences manifest or reality asserts itself: self-driving cars that [seek to crash into lane dividers](#)¹⁹ or [hallucinate non-stop traffic lights](#)²⁰ when following trucks transporting said lights. Users unfamiliar with LLMs can make the error of [trusting their hallucinatory outputs](#)²¹, and companies may [inflate their claims](#)²² of models’ abilities [until users report otherwise](#)²³. Overattribution often results from overly anthropomorphic language or expectations, and thus we can help protect users by not trying to exploit or misapply their intuitions regarding human interactions.

A second criterion to evaluate the appropriateness of anthropomorphic language is to ask whether a given metaphor even applies to comparable interactions between humans. In a recent paper examining the rhetoric of AI systems serving as “collaboration” partners, [Evans, Robbins, and Bryson](#)²⁴ expose the many differences between human collaborations – involving shared goals and mutual understanding among consenting equals– and the disproportionate power dynamics of a human using an AI tool. They prefer the term “joint action” as a more accurate (if less buzzworthy) term that avoids the connotations of “collaboration.” This correction is timely,

¹⁶ Andreas Theodorou, Robert H. Wortham, and Joanna J. Bryson, “Designing and Implementing Transparency for Real Time Inspection of Autonomous Robots,” *Connection Science* 29, no. 3 (July 3, 2017): 230–41, <https://doi.org/10.1080/09540091.2017.1310182>.

¹⁷ Alan F.T Winfield et al., “IEEE P7001: A Proposed Standard on Transparency,” *Frontiers in Robotics and AI* 8, no. Sec. Ethics in Robotics and Artificial Intelligence (July 26, 2021), <https://doi.org/10.3389/frobt.2021.665729>.

¹⁸ Butlin, Patrick, et al. *Consciousness in Artificial Intelligence: Insights from the Science of Consciousness*. arXiv:2308.08708, arXiv, 22 Aug. 2023. *arXiv.org*, <https://doi.org/10.48550/arXiv.2308.08708>

¹⁹ Thorbecke, Catherine. “Tesla on Autopilot Had Steered Driver towards Same Barrier before Fatal Crash, NTSB Says.” *ABC News*, 12 Feb. 2020, <https://abcnews.go.com/Business/tesla-autopilot-steered-driver-barrier-fatal-crash-ntsb/story?id=68936725>.

²⁰ Robitzski, Dan. “Tesla Autopilot Glitch of Truck Hauling Traffic Lights.” *Futurism*, 23 June 2021, <https://futurism.com/the-byte/tesla-autopilot-bamboozled-truck-traffic-lights>.

²¹ Neumeister, Larry. “Ex-Trump Lawyer Michael Cohen Says He Unwittingly Sent AI-Generated Fake Legal Cases to His Attorney.” *AP News*, 29 Dec. 2023, <https://apnews.com/article/michael-cohen-donald-trump-artificial-intelligence-777ace9cc34aa0e56398fd47a1d6b420>.

²² Yann LeCun [@ylecun], “A Large Language Model Trained on Scientific Papers. Type a Text and Http://Galactica.Ai Will Generate a Paper with Relevant References, Formulas, and Everything.,” Tweet, *Twitter*, November 15, 2022, <https://twitter.com/ylecun/status/1592619400024428544>.

²³ Yann LeCun [@ylecun], “Galactica, the LLM for Scientists from Meta,...Was Taken down after 3 Days...,” Tweet, *Twitter*, November 14, 2023, <https://twitter.com/ylecun/status/1724448825509851332>.

²⁴ Katie D. Evans, Scott A. Robbins, and Joanna J. Bryson, “Do We Collaborate With What We Design?,” *Topics in Cognitive Science* n/a, no. n/a, accessed January 16, 2024, <https://doi.org/10.1111/tops.12682>.

since the temptation to label Generative AI systems as “collaboration partners” may increase with their increasingly popular applications in art and music.

The reasoning whereby AI systems are barred from preaching or sacramental roles in Eden is explicitly grounded in ontology and theology. It’s fascinating to me that in Eden, machines still “might assist musicians with music, both in composing and performing.” Prior to the onslaught of “worship tunes,” Christian hymns served as [both worship and theological education](#)²⁵. In some Charismatic circles, with worship often prioritized over preaching, one could almost envision the inverse of Eden: embracing AI for sermons but reserving musical worship as the soulful province of humans. This interplay between AI and creativity will be expanded upon in the following section.

IV. Creativity with Computations and Constraints

The history of generative aids in human creativity [spans centuries](#)²⁶, from Mozart’s [musical dice game](#)²⁷ to probabilistic systems that influenced modern tools like ChatGPT, rooted in [Markov's theological debate](#)²⁸ over a century ago. We’re currently riding a wave of “Generative AI,” yet creativity encompasses more than generation; it also involves discernment and, often, strategic subtraction, as in marble sculpture. Rick Rubin attributed his success as a music producer to discernment alone [when he famously confessed](#)²⁹,

“I have no technical ability. And I know nothing about music... I know what I like and what I don’t like. And I’m decisive.”

This self-awareness echoes the editor’s critical role, which lies not just in adding content but in the judicious selection from myriad possibilities and the careful excision of the extraneous, thereby distilling the work to its purest form. It’s worth noting that these dual processes of generation and discrimination are realized in a powerful class of neural networks known as [Generative Adversarial Networks](#)³⁰. GANs can be difficult to train because the feedback from the

²⁵ Paul Wesley Chilcote, *Singing the Faith: Soundings of Lyrical Theology in the Methodist Tradition* (Nashville: Wesley’s Foundry Books, 2020).

²⁶ Marcus Du Sautoy, *The Creativity Code: Art and Innovation in the Age of AI*, First US edition (Cambridge (Mass.): The Belknap press of Harvard university press, 2019).

²⁷ PlayOnlineDiceGames.com, “Online Mozart Dice Game,” Play Online Dice Games, accessed January 16, 2024, <http://www.playonlinedicegames.com/mozart>.

²⁸ Brian Hayes, “First Links in the Markov Chain,” *American Scientist*, February 6, 2017, <https://www.americanscientist.org/article/first-links-in-the-markov-chain>.

²⁹ Hugh Allen, “Rick Rubin Interview | 60 Minutes | Transcripts.” *Rev Blog*, <https://www.rev.com/blog/transcripts/rick-rubin-the-60-minutes-interview-transcript>.

³⁰ Varshita Sher, “How I Would Explain GANs From Scratch to a 5-Year Old: Part 1,” TOPBOTS, March 11, 2021, <https://www.topbots.com/explain-gans-from-scratch-part1/>.

“critic” part of the network is often not sufficient to “coach” the generator part along constructive lines; there are simply too many directions that the generator could go.

Thus, we come to another tool to help guide the creative process and limit the seemingly endless space of possibilities: the imposition of constraints! [Igor Stravinsky praised the value of constraints](#)³¹, saying:

“The more constraints one imposes, the more one frees one’s self of the chains that shackle the spirit.”

The Psalmist expressed a similar duality of freedom amid the constraints of God’s commands,

“I run in the path of your commands, for you have broadened my understanding.”
— *Psalm 119:32 (NIV)*

Even when constraints aren’t obvious, the ability to limit choices is important for both skilled people in creative roles and advanced AI systems. For instance, AlphaGo's success partly came from its skill in choosing to evaluate only the best possible moves, avoiding the need to look at every single option in a game with countless possibilities. In the field of generative models, there is an illuminating concept called the [Manifold Hypothesis](#)³². To illustrate, imagine trying to create a digital image with each pixel being able to be any color. Most random combinations would just look like messy static rather than a clear picture of something. If we picture the values of all pixels as forming the coordinates of a single data point in a very high-dimensional space (e.g. $3N^2$ dimensions for an RGB image with a resolution of $N \times N$ pixels), then the set of all data points representing all “meaningful” images would lie in a reduced-dimensional subset out of the vast possibility space. The Manifold Hypothesis says that the points in this subset are “connected” and lie along what mathematicians call a “manifold” – imagine a curvy slice being extracted somehow from a block of cheese. According to the hypothesis, the reason that humans and neural networks can be so good at finding “meaningful” points in the vast space of possibilities is that we learn to intuit or predict the shape of these manifolds, reducing the complexity of finding a meaningful point. Returning to Stravinsky, the manifold acts as a constraint.

Adhering to and deviating from such a manifold [echoes key dynamics in both creativity and Christian spirituality](#)³³, whereby the constraints or “rules” that provide structure can be life-

³¹ Ron Drotos, “Igor Stravinsky on Creativity and Freedom,” Keyboard Improv, October 28, 2016, <https://keyboardimprov.com/igor-stravinsky-on-creativity-and-freedom/>.

³² Aiden Rocke, “Revisiting the Manifold Hypothesis,” *LessWrong* (blog), January 10, 2023, <https://www.lesswrong.com/posts/uH4GfZwBYteiy2GrL/revisiting-the-manifold-hypothesis>.

³³ “Living Off-Shell with Guiding Constraints - Scott Hawley Interview,” *Sacred Space*, directed by Rex Schnelle, 2022, https://youtu.be/9fkoy_hE2AY?si=WhlhHta60hApidGz&t=422.

giving. However, transforming the adherence to rules into an end in itself leads to legalism and stagnation. We need the freedom to be able to break the rules, to deviate from the constraint manifold, and therein lies a source of creativity and life. However, if we deviate too far or too much, the work loses integrity, or our lives spin out of control. Thus we need an “attractive force” ever pulling us back toward the constraint manifold. In neural network training, this force appears in the form of the gradients of the terms in the loss function. Thus, from a machine learning standpoint, we might paraphrase a line in the classic hymn, “Come Thou Fount of Every Blessing,” to read,

Let thy goodness, like a fetter, [serve as an additional term in my loss function].

Granting the freedom to stray from perfection while consistently striving for it, we can, in our creative endeavors, mirror God by working imperfections into creations of surpassing beauty. [Learning these manifolds](#)³⁴ and using them effectively are, in the words of [Deep Thought](#), “tricky.”³⁵ While it’s possible to construct smooth, compressed manifolds inside the latent spaces of neural networks (i.e., via [Variational Autoencoders](#)³⁶), the most [modern, sophisticated generative models](#)³⁷ employ more complicated data representations that often amount to [using graphs to quantize the space](#)³⁸. The tricky and remarkable aspect is that in many applications of these models, the key is to find “nearest neighbor” quantized data points to represent a desired point in space. This process, when applied to graphs, presents an AI pathfinding problem [known to be “NP-hard.”](#)³⁹ similar to the widely recognized "Traveling Salesman" problem. The insight that humans and artificial neural networks are adept at solving such problems is, to this author, one of the more fascinating connections to emerge in recent years. This connection between humans and artificial neural networks navigating such challenges is a captivating topic that merits greater elaboration elsewhere.

V. “Open Source”

In the story, Eden’s abbess asks, “Do you believe in such open-source technology, Sondra?” In our timeline, the past two years have been pivotal for open-source AI, with [leading](#)

³⁴ Leland McInnes et al., “UMAP: Uniform Manifold Approximation and Projection,” *Journal of Open Source Software* 3, no. 29 (September 2, 2018): 861, <https://doi.org/10.21105/joss.00861>.

³⁵ Douglas Adams. *The Hitchhiker’s Guide to the Galaxy*, Ballantine Books, 2009.

³⁶ Sergey Nikolenko, “Variational Autoencoders (VAEs): Generative AI I - Synthesis AI,” *Synthesis AI* (blog), February 7, 2023, <https://synthesis.ai/2023/02/07/generative-ai-i-variational-autoencoders/>.

³⁷ Sergey Nikolenko, “Generative AI VI: Stable Diffusion, DALL-E 2, and Midjourney - Synthesis AI.” *Synthesis Blog*, 9 Aug. 2023, <https://synthesis.ai/2023/08/09/generative-ai-vi-stable-diffusion-dall-e-2-and-midjourney>.

³⁸ Scott H. Hawley, “Residual Vector Quantization,” *Hawley Blog*, 12 June 2023, <https://drscotthawley.github.io/blog/posts/2023-06-12-RVQ.html>.

³⁹ Shicong Liu, Hongtao Lu, and Junru Shao, “Improved Residual Vector Quantization for High-Dimensional Approximate Nearest Neighbor Search” (arXiv, September 17, 2015), <https://doi.org/10.48550/arXiv.1509.05195>.

[companies](#)⁴⁰ and [legislative](#)⁴¹ [bodies](#)⁴² weighing its impact on society. Few question the benefits of open source software in general: without the stack of systems such as Linux, [nginx](#)⁴³, node.js, and MySQL, to name but a few, our information society would quickly halt. Debate has focused on the extent to which open source AI systems might be exempted from new regulations aimed at ensuring the safe and ethical use of AI. Sharing code openly fosters [transparency](#)⁴⁴, enhancing security, shaping policy, and propelling technological innovation; however, it also introduces a “proliferation risk” in AI, a field where the potential to drive significant change comes with both [promise and peril](#)⁴⁵. All powerful technologies can pose public risks if they are adopted and employed without due consideration, though such risks are not unique to open source. Early integration of ChatGPT with customer-facing websites has [exposed companies to unwanted and hilarious](#)⁴⁶ outcomes. Because the risks are still poorly understood, we have seen [many highly influential scientists](#)⁴⁷ calling for a “pause” on the development of more sophisticated AI. Contrasting views persist within the AI community, with [some](#)⁴⁸ advocating transparency (and thereby open source) as a tool for global collaboration on safer AI, while others [prioritize caution](#)⁴⁹. This difference of opinion on the best way to proceed is sufficient to explain those [rated best for transparency](#)⁵⁰ may also be [ranked last in terms of AI Safety](#)⁵¹. Nevertheless: transparency advocates recently announced the [launch of the AI Alliance](#)⁵², in support of which Jeremy Howard, founder of fast.ai, asserted,

⁴⁰ Sharon Goldman, “Hugging Face CEO Tells US House Open-Source AI Is ‘Extremely Aligned’ with American Interests,” *VentureBeat* (blog), June 22, 2023, <https://venturebeat.com/ai/hugging-face-ceo-tells-us-house-open-source-ai-is-extremely-aligned-with-american-interests/>.

⁴¹ Chris Middleton, “AI - Open Wide and Say ‘Ah!’ As the UK House of Lords Debates Open Source,” *Diginomica*, November 13, 2023, <https://diginomica.com/ai-open-wide-and-say-ah-uk-house-lords-debates-open-source>.

⁴² Foo Yun Chee et al., “Talks on EU’s AI Act to Resume Friday after Marathon Debate,” *Reuters*, December 7, 2023, sec. Technology, <https://www.reuters.com/technology/eu-still-hammering-out-landmark-ai-rules-marathon-overnight-talks-2023-12-07/>.

⁴³ “Nginx - Market Share, Competitor Insights in Web And Application Servers.” *6sense*, <https://6sense.com/tech/web-and-application-servers/nginx-market-share>

⁴⁴ Katharine Miller, “Introducing The Foundation Model Transparency Index” (Human-Centered Artificial Intelligence, Stanford University, October 18, 2023), <https://hai.stanford.edu/news/introducing-foundation-model-transparency-index>.

⁴⁵ Tristan Harris and Aza Raskin, “The Promise and Peril of Open Source AI with Elizabeth Seger and Jeffrey Ladish,” *Your Undivided Attention*, Center for Humane Technology, accessed January 16, 2024, <https://www.humanetech.com/podcast/the-promise-and-peril-of-open-source-ai-with-elizabeth-seger-and-jeffrey-ladish>.

⁴⁶ Notopoulos, Katie. “A Car Dealership Added an AI Chatbot to Its Site. Then All Hell Broke Loose.” *Business Insider*, 18 Dec. 2023, <https://www.businessinsider.com/car-dealership-chevrolet-chatbot-chatgpt-pranks-chevy-2023-12>.

⁴⁷ “Pause Giant AI Experiments: An Open Letter,” *Future of Life Institute* (blog), accessed January 15, 2024, <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.

⁴⁸ “EleutherAI,” EleutherAI, December 17, 2023, <https://www.eleuther.ai>.

⁴⁹ Aaron Mak, “When Is Technology Too Dangerous to Release to the Public?,” *Slate*, February 22, 2019, <https://slate.com/technology/2019/02/openai-gpt2-text-generating-algorithm-ai-dangerous.html>.

⁵⁰ Katharine Miller, *ibid.*

⁵¹ Sean Ó hÉigeartaigh et al, “Do Companies’ AI Safety Policies Meet Government Best Practice?,” Leverhulme Centre for the Future of Intelligence, accessed January 15, 2024, <http://lcfi.ac.uk/news-and-events/news/2023/oct/31/ai-safety-policies/>.

⁵² “Introducing the AI Alliance,” *Meta AI* (blog), December 4, 2023, <https://ai.meta.com/blog/ai-alliance/>.

“Open source is the backbone of all leading artificial intelligence software. With open source, the entire community comes together to collaborate on solving the toughest problems, the most effective solutions rise to the top, and everyone benefits.”

What does “open source” really mean? According to the [definition by the Open Source Initiative](#)⁵³ who originally coined the term, “Open source doesn’t just mean access to the source code.” It requires adherence to specific criteria, including points 5 and 6 which are “No Discrimination Against Persons or Groups” and “No Discrimination Against Fields of Endeavor,” respectively. Thus Eden's restrictions against military use contradict the official designation of “open source.” However, the moniker “open source” may be too descriptive for its own good, and its widespread (mis)use risks the same “brand dilution” afflicting household names like [Kleenex](#)⁵⁴ and [Velcro](#)⁵⁵. This dilution reflects deeper issues within the open-source movement, where traditional licenses are failing to address modern challenges. Bruce Perens of OSI [recently pointed out the inadequacy](#)⁵⁶ of current licenses in tackling contemporary issues, proposing a “Post-Open” model to redefine enforceable open software standards. This suggests new directions in software licensing, reconciling open-source values with the demands of modern tech governance.

Regardless of the nitpicking over the term “open source,” it makes sense that creators should have the ability to set their own licenses which may be as permissive or restrictive as they wish. Stability AI famously disrupted the AI development ecosystem in 2022 by [releasing both their code and model weights](#)⁵⁷ using a “Creative ML” version of the extremely permissive [OpenRAIL-M license](#)⁵⁸, which allowed anyone to build anything from it, even commercial products, with minimal restrictions (e.g., against the production child pornography). Even still, [some users](#)⁵⁹ claimed that such restrictions amounted to censorship and that Stability’s use of the “open source” moniker was unjustified, and wanted no limits placed on their so-called “freedom of expression.”⁶⁰ Since then, we’ve seen a mixture of licenses, typically with the source code being released via highly permissive licenses, but the pretrained model weights are another story.

⁵³ “The Open Source Definition,” Open Source Initiative, July 7, 2006, <https://opensource.org/osd/>.

⁵⁴ Michelle Cheng, “The ‘Lawyers of Kleenex’ Are Taking a Soft Approach to the Hard Realities of Genericide,” *Quartz*, November 3, 2022, <https://qz.com/the-lawyers-of-kleenex-are-taking-a-soft-approach-to-1849737749>.

⁵⁵ VELCRO® Brand, *Don’t Say Velcro*, 2017, <https://www.youtube.com/watch?v=rRi8LptvFZY>.

⁵⁶ Thomas Claburn, “What Comes After Open Source? Bruce Perens Is Working on It,” accessed January 16, 2024, https://www.theregister.com/2023/12/27/bruce_perens_post_open/.

⁵⁷ “Stable Diffusion Public Release,” Stability AI, accessed January 16, 2024, <https://stability.ai/news/stable-diffusion-public-release>.

⁵⁸ Kyle E. Mitchell, “Open RAIL-M Is More Unclear Than Scary,” *Dev/Lawyer* (blog), January 26, 2023, <https://writing.kemitchell.com/2023/01/26/Open-RAIL-M-Unclear>.

⁵⁹ “r/UnstableDiffusion,” reddit.com, accessed January 16, 2024, <https://www.reddit.com/r/UnstableDiffusion>.

⁶⁰ Kyle Wiggers, “As AI Porn Generators Get Better, the Stakes Get Higher,” *TechCrunch*, July 21, 2023, <https://techcrunch.com/2023/07/21/as-ai-porn-generators-get-better-the-stakes-raise/>.

As one example, [Meta released their AudioCraft suite](#)⁶¹ with open code but with a restriction against commercial usage of the pretrained model weights. This is because it's not just the code but the *training data* that is [highly-contested intellectual property](#),^{62,63} which will likely produce key legal precedent in the next couple of years – it's an exciting time to be a lawyer! Perhaps in the future of the “Back to Eden” story, the Eden community might follow a similar prescription of sharing their source code openly but not the model weights, or perhaps even using the model code of the evil New Con empire but re-purposing it by retraining on “ethically sourced” data. The recognition of the ethical implications associated with the [planet-scale supply chain in the development of AI systems](#)⁶⁴ in “Back to Eden” brings an additional layer of depth to the narrative.

VI. Conclusion: “As for me and my house”

The fragmented world in “Back to Eden” extrapolates from our current global context, suggesting that a unified ethical and regulatory framework for AI may be unattainable. Rather than trying to control the AI applications of other actors, those in Eden have adopted an “[As for me and my house](#)”⁶⁵ approach to foster human flourishing. Rather than isolating themselves from the world, Eden remains open to collaboration, exchanging resources and code with communities sharing similar values. While the dangers of threats from bad actors are acknowledged, the people of Eden choose to put their faith in God and remain true to their principles, as Eden's Mother Regina says:

“Anyone can be wiped out, as you put it...But we refuse to solve the problems of this world by the methods of this world.”

The author of “Back to Eden” makes a significant contribution by crafting an intriguing portrayal of a community that balances worldly engagement, technological integration, and adherence to faith. This story enriches our broader discussion of these intricate issues and merits our appreciation.

⁶¹ “Open Sourcing AudioCraft: Generative AI for Audio Made Simple and Available to All,” Meta AI, accessed January 16, 2024, <https://ai.meta.com/blog/audiocraft-musicgen-audiogen-encodec-generative-ai-audio/>.

⁶² “Artists Take New Shot at Stability, Midjourney in Updated Copyright Lawsuit | Reuters,” accessed January 16, 2024, <https://www.reuters.com/legal/litigation/artists-take-new-shot-stability-midjourney-updated-copyright-lawsuit-2023-11-30/>.

⁶³ Michael M. Grynbaum and Ryan Mac, “The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work,” *The New York Times*, December 27, 2023, sec. Business, <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>.

⁶⁴ Kate Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (New Haven London: Yale University Press, 2021).

⁶⁵ “Bible Gateway Passage: Joshua 24:15 - New American Standard Bible,” Bible Gateway, accessed January 16, 2024, <https://www.biblegateway.com/passage/?search=Joshua%2024%3A15&version=NASB>.

Bio

Scott H. Hawley, Ph.D., is a Professor of Physics and Senior Data Fellow at Belmont University, Nashville, Tennessee, and the 2022 recipient of the Faculty Award for Christian Scholarship. Hawley is a Founding Member of AI & Faith and participates in the Artificial Intelligence Grand Challenge Initiative of Vanderbilt University's History Department. Hawley's AI research focuses on machine learning applications in acoustics and signal processing, particularly for musical audio production. His contributions to AI Ethics include publications in *Perspectives on Science and Christian Faith* and winning the 2020 FaithTech Writing Contest. Before his work in machine learning, Hawley conducted astrophysical research, with a notable collaboration with the Naval Research Laboratory appearing in *Science*, and work on generative imagery with the Army Research Laboratory being featured in *ACM SIGGRAPH*.