Examining the Societal and Environmental Impacts of AI Development: Inequities and Infrastructure Problems

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Ryan Chung

Spring 2025

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Sean Murray, Department of Engineering and Society

Introduction

Artificial Intelligence or AI has expanded on the scene of technology over the last few years. AI in its most popular form can be more specifically described as a Large Language Model, a system that is able to process, understand, and generate human language (What Is a Large Language Model (LLM)? - University of Arizona Libraries, 2023). This ability to interpret human language has taken the world by storm. Products like ChatGPT, Claude, and newly published DeepSeek have changed lives by allowing users to access information by prompt in a personalized way search engines like Google never have. This promise of automated workflows and replacement of traditional employees with computing power has caused a mass adoption affect amongst millions of companies. However, this relentless pursuit of technological advancement has unseen effects that are seen in societal and environmental domains.

The development of AI, as proved by OpenAI the creator of ChatGPT, is extremely resource intensive. OpenAI has raised over 17 billion dollars in funding since its inception. While this may be a special case for the largest AI company in the world, the need for large amounts of funding to develop AI technology is seen from all levels. Companies like Databricks, xAI, Waymo, and Anthropic have raised a combined 28 billion in 2024 alone. However fundraising isn't the only reason AI development is resource intensive. At its core, AI is trained on extremely large datasets through complex algorithms which demand high amounts of computing power only found in large datacenters. Datacenters use extreme amounts of power and water to complete these processes demanded by AI. From a life cycle analysis perspective of AI development there are many inequities as it relates to gender and socio-economic status. A thorough literature review detailed later in the paper revealed that from a funding perspective, venture capital firms have a predisposition for a certain archetype of individual, and from an

environmental perspective, low-socioeconomic individuals are disproportionately affected by the rising costs of resources driven by the resource hungry datacenters used to power development.

Problem Definition

As spoken on earlier, we can see that the development of AI is largely resource intensive, requiring vast amounts of computational power and capital investment. Companies like OpenAI, Waymo, Anthropic, etc have raised upwards of 40 billion in funding in just the last year. Due to the new expansion of the development of AI, data centers, which are at the core of the development of AI are in high demand. More specifically data centers house hardware such as Graphics Processing Units (GPUs) or Application Specific Integrated Circuits (ASICs) which are used to calculate the mathematical operations needed in machine learning or in this case AI (Lohn & Musser, 2022). These data centers also hold millions of terabytes of data which are used to train the large language models like ChatGPT or Claude. The sheer size of both the storage and hardware component drives the creation of dedicated spaces almost specifically for the purpose of AI. However, with all of this computing power and storage, there has to be resources used to power and maintain the data centers facilities. These large data centers utilize mass amounts of power to keep the machines running as well as water to cool the machines. Both are resources that are priced dynamically and adhere to demand in the region. As highlighted by a study done at UC Riverside, five queries through ChatGPT can use up to 500ml of water (UC Riverside Paper).

The core problem up for discussion is the inequitable distribution of AI's benefits and burdens. We have seen problems like this occur across cell phones, home computers, and other new, innovative forms of technology. While the benefits revolve around a small group of tech

elite and investors, society at large bears the externalized costs of the development of AI. In short, its benefits are concentrated in the hands of a few, while its costs are distributed widely, often on the shoulders of those who couldn't bear them. This lifecycle spans from venture capital funding at the start, through development in power-hungry data centers, to impacts on marginalized communities at the end. The problem definition thus lies in uncovering how structural inequities manifest at each stage of AI's journey: who gets to create AI, and who absorbs the fallout.

These costs are of course just the tip of the iceberg as it relates to AI development. Just underneath the surface we can see the inequities that hide itself in the process of AI development play a larger role than at face value. The pipeline of AI innovation starts at the venture capital (VC) funding stage. VC funding is riddled with a multitude of gender and socio-economic problems. Despite recent improvements, women founders receive a disproportionately small amount of VC funding as compared to their male counter parts. In 2023, the percentage of female-male founding teams that received funding made up a record-breaking 25% of all VC funds raised. This number is up significantly from the last several years where it was somewhere around 10 to 12 percent. However this is due to large deals, more specifically multi-billion-dollar deals from companies like OpenAI and Anthropic who have female co-founders (Teare, 2024). This new peak however, masks the reality of the venture landscape for female founders; the majority of AI startups still receive funding only when led by men. Across the time period of 2015 to 2023, the percentage of funding that goes to female-only founders has stayed steady at approximately 2% (Teare, 2024). These numbers reflect a persistent gender gap in the world of VC funding.

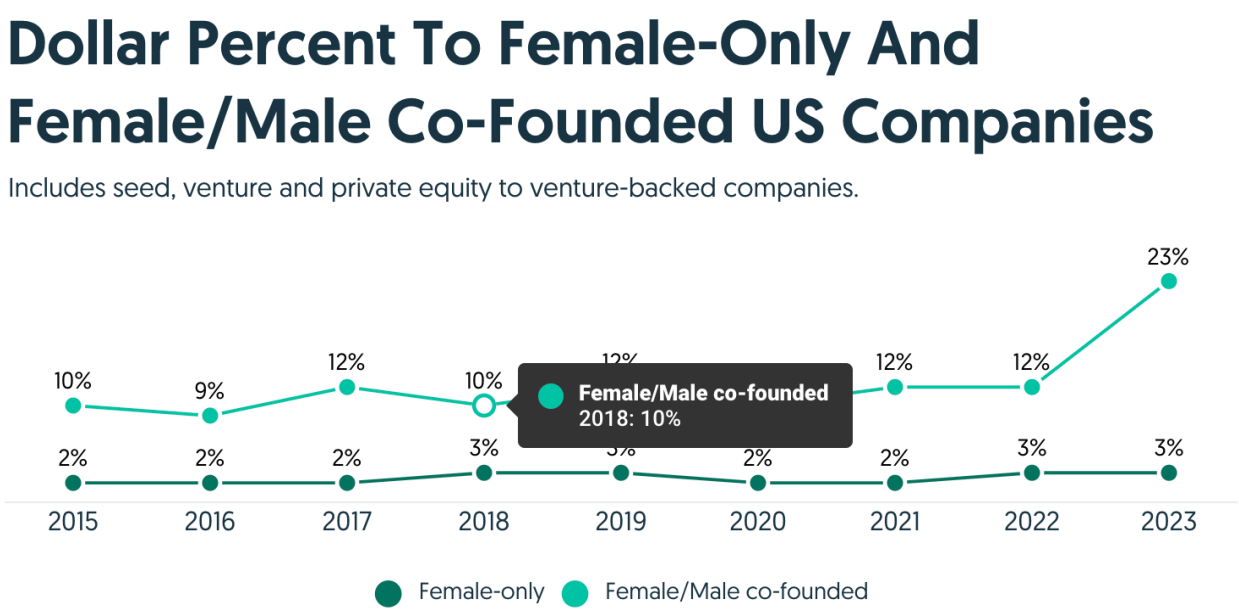

Figure 3: Percentage of Venture-Backed Funding going to Female-Only Founders (Teare, 2024, p.1)

Knowing these underlying issues with AI and its lifecycle, we can hone in our field of research and the analysis we need to draw conclusions. The glaring question is, why then do we continue to push AI despite its harms? Who ultimately benefits? In order to answer this, we need to unpack the full life cycle of its development from venture capital priorities, development practices, and the deployment of AI systems. This allows us to see inequities in social justice, sustainability, and problems in infrastructure as they directly relate to the relentless pursuit of AI development.

Research Approach

In order to examine the lifecycle of AI and its development from funding to social impact, this study will employ a multi-method research approach grounded in both data analysis and case-based investigation. The goal here is to connect the abstract funding patterns from the epicenter of AI company ideation and funding with concrete community impacts in a specific region. In this case the specific region is Northern Virginia, which has been hailed as one of the

largest datacenter areas in the world. In order to do so, this study will approach the biases in silicon valley with a quantitative approach to analyze specific funding trends since the late 2010s. Also, this study will use qualitative case studies and literature review to draw conclusions about AI's social lifecycle. By combining evaluations of both the venture capital stage and the implications of AI's social lifecycle, this research aims to capture the full spectrum of AI's lifecycle, from Silicon Valley board rooms to 'data center alley' neighborhoods.

A thorough review of existing literature underpins this research. Although to my knowledge, no studies have been done in order to expose the inequalities of AI development and who stands to gain versus who faces the consequences. This includes scholarly work on venture capital bias including gender disparities in startup funding, and on environmental justice related to technology infrastructure. Key theoretical perspectivs guiding the analysis are Social Construction of Technology (SCOT) and the Political Econony of Tech. These frameworks allow the research to interpret how power and social values shape technological advancement. Some similar concepts that will be used here are meritocracy as a cultural myth in the analysis of tech funding, and the idea of data centers as "digital plantations" stemming from critical geography research highlighting extractive tech companies (Rosati, et al., 2023). Along with the use of those theoretical frameworks, a life-cycle analysis is the overarching concept used to analyze each step of the process from funding to ecological and societal concerns. This will also expose the inequalities across each step and learn who truly faces the consequences of AI development. This theoretical grounding ensures that the research treats the problem not as isolated issues but as a systemic, structural phenomenon.

Ventural Capital

In order to frame the quantitative analysis of ventural capital funding withn AI development specifically, there is some context needed. The late 2010s and early 2020s saw an unprecedented surge in the creation and investment into AI companies. In 2021, global AI startup funding hit a high of $93.5 billion dollars, roughly two times the 2020 total, across 746 deals, more than double the amount invested in 2020 (Syal, 2022). This boom made 2021 a peak year in terms of funding to AI companies/startups. Major rounds included multi-billion dollar raised by companies already previously discussed such as Databricks and OpenAI (Syal, 2022). The AI Index from Stanford which is referenced in the same article note 15 mega rounds, rounds greater than 500 million each, in 2021 (Syal, 2022). This indicates a jump in the funding of AI companies and a belief that they could yield profit.

This boom was followed by a post-releasee of ChatGPT wave of funding. By 2024, AI-related startups captured almost 46% of all U.S venture funding (Robbins, 2025). This was a jump from 10% from almost a decade earlier (Robbins, 2025). Breakout successes like ChatGPT have brought on a new wave of optimism with multi-billion dollar rounds to show for it. The amount raised in 2024 even eclipsed the massive amounts of money raised in 2021. These booms at the end of the 2010 and beginning of 2021 is how we will frame the quantitative analysis in the upcoming section. These are the most important timeframe to analyze the true lifecycle of AI given its in its early stages of development.

The Thirst for AI – Case Study on YCombinator

As stated previously, we have seen a boom in AI company creation and funding in the last several years. In some ways, due to the success of ChatGPT on a global scale. With this large boom, there is a crux to the pipeline of AI startup ideas being created. Here I would like to introduce YCombinator. YCombinator or YC is known worldwide as a launchpad for innovation,

it fosters startips that they hope will shape the global technology landscape. Notable alumni companies include AirBNB and Stripe. YCombinator known for its innovative technology has taken a large pivot towards mostly AI companies. In YC's Winter 2024 batch, 86 startups which was nearly half the cohort, focused on AI technologies (Shah, 2024). This represents an upward trend in AI companies being put through YCombinator. This is heavily aligned with VC trends during this time period.

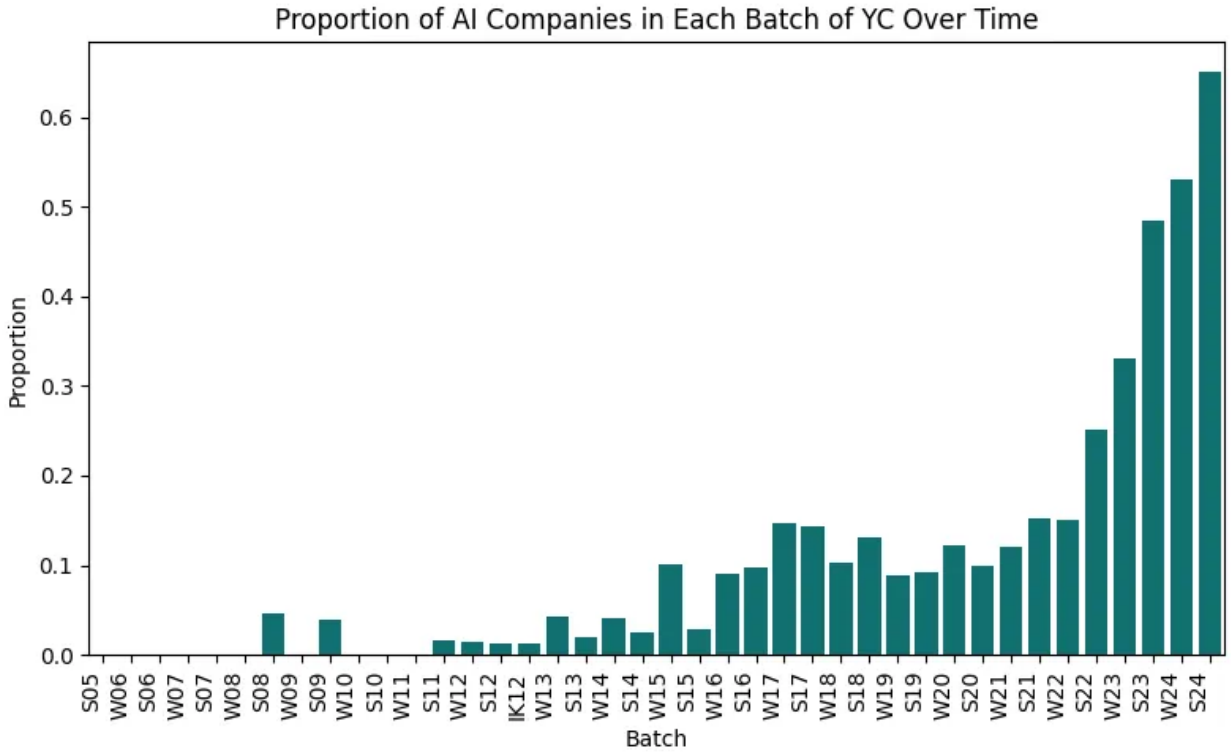

Figure 2: Proportion of AI Companies in Each Batch of YC Over Time (Shah, 2024, p.1)

An important notion that also needs to known about YCombinator is its founder profile. They consistently show a disparity in who actually participates in YCombinator. Reports have shown that 45% of founders are graduates from a specific subset of elite institutions such as Harvard, UC-Berkeley, and Stanfard (Chung, 2024). Also, self-reported from YC itself is the shocking statistics on female participation in YC. Gary Tan, the founder of YC, has come out and said that in their recent batches only 11% of the founders identified as women (Tan, 2024).

In summary, YC's recent cohort demographics reveal a domination of male, highly-educated, well-networked individuals. This stems from the central theory of meritocracy, assuming that every applicant is on a level playing field, which ignores any barriers applicants may face to get where they are.

Northern Virginia – The Epicenter of AI Hardware

Northern Virginia is the world's largest data center hub. They host roughly 25% of all United State data center capacity and roughly 200 data centers are located in Loudoun County alone. Not only are there already roughly 200, 117 more are in the pipeline (Turner, 2025). Global leaders in the sector like AWS, Meta, and Google have a foothold on this area. County officials note "there has not be a single day in the past 14 years when a data center was not under some type of construction or expansion in Loudoun County." Northern Virginia's favorable policies, pre-existing fiber infrastructure, and ample power attract these companies, while AI innovation and productization drive up demand for server capacity (Carey, 2025). As AI training expands, data center construction has accelerated.

This growth, of course, comes with significant environmental demands. 25% of all of Virginia's electricity now powers data centers, a number that is supposed to double by 2040. They also use massive amounts of water, as much as three hospitals (Carey, 2025). Overall, data centers emit more carbon than the entire airline industry (Monserrate, 2022). Although these costs are partly externalized, tech firms benefit from their immense computing capabilities.

Impacts on Northern Virginia Residents

The sudden growth in data center development in Loudoun County has caused property value to shoot up, exacerbating housing costs. Home prices have risen by 65% in Northern

Virginia as seen in Figure 3 (Staff, 2024), reflecting a trend discovered by van Tilburg which states that properties near new data center development often increase roughly 5.27% in value (van Tilburg, 2021). This explosive growth greatly benefits landowners but displaces residents who cannot afford the hike in taxes or rent. Simultaneously, resource usage by data centers can strain local utilities, causing a spike in the cost of electricity and water in the area. This disproportionately affects low income households who are already more vulnerable. Many of these households will face energy insecurity, forcing them to choose between costs like medical care or paying utility bills (Hernandez, 2023). Though data centers bring in tax revenue to the county and state, members of the community who are not in the tech field gain fewer benefits.

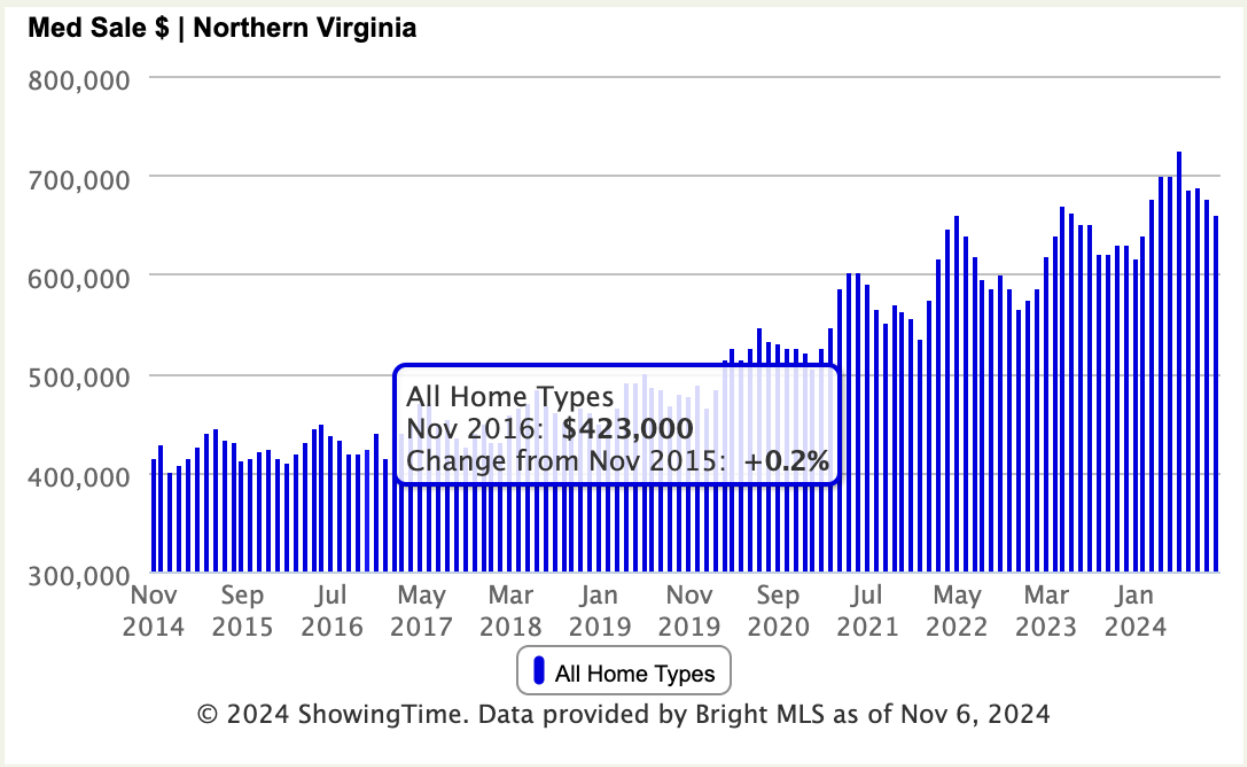

Figure 3: Median Sale Price of Homes in Northern Virginia Over 10 Years (Staff, 2024)

Linking Two Domains Together

These exploitative trends illustrate a cleary cause-and-effect cycle between the beginning of the lifecycle, AI venture funding, and the end of the cycle, local environmental and social impacts. As early-stage tech companies attract billions in VC investment to develop enterprise-level AI based applications and services, they require more data center capacity, thus intensifying environmental and socio-economic effects on local communities across the United States. If we circle back to SCOT or Social Construction of Technology as a framework, we see that the venture capitalists, tech giants, and policy makers are the powerful stakeholders who influence the process, while marginalized voices have limited influence on local decisions. To take this idea even further, scholars have specifically labelled this dynamic a "digital plantation" where the enterprise stage "server farming" exploits local resources without distributing the benefits across all who are affected (Rosati et al., 2023). To close, the evidence points us to the conclusion that Silicon Valley's pursuit of AI advancement causes growth in data center infrastructure across local communities, producing gains for stakeholders while burdening local communities.

Results

This sections presents the results of the literature reviews, case studies, and quantitative analysis in three sections – Bias in Venture Funding and Development, Environmental and Societal Impacts of AI Infra., Unification of the Lifecycle of AI.

Biases in Venture Funding and Development

The analysis confirms that the funding in venture capital as it relates to AI is concentraded among a very small subset of founders. We have seen that despite growth in recent years, AI funding as it relates to women remain is in a dire state. The Crunchbase data review

revealed that even taking into account male-female founder groups along with female-only founder groups, the percentage of funding going to woman is abysmal. Is a banner year of 25 and 3 percent respectfully, inclusive? Given those numbers, a staggeringly high percentage of US VC funding went to males, 75-93%. This quantifiable gender gap aligns with prior studies and background research on systemic VC bias. Furthermore, a breakdown of AI startup founder shows a considerable bias towards education. More specifically, hand picking individuals from a small subset of founders from a very specific educational background, as evident by YC and other reports. This illustrates a very obvious pedigree bias. By a quick LinkedIn search (the manuscript referenced under my name) under the filter of just YC – W24 being in their profile, we easily saw that 45% of founders were alumni of top-tier universities (e.g, Stanford, MIT, UC Berkeley). Entrepreneurs outside elite circles face barriers to accessing the crux of AI application development and funding. Thus skewing the vision of AI technology to a specific group of individuals.

As we continue to dive into YCombinator and the influence it plays in early funding and development of AI, it is impossible to ignore how their "founder mold" qualitatively reinforces these numerical trends. To put it simply, YC is dominated by male, Silicon Valley-style entrepreneurs, with very few women in ech batch and an underrepresentation of founders from non-traditional backgrounds. There are also several cultural narratives at play in such as meritocracy and disruption in YC's discourse and help explain this skew. As we know, YC professes to fund the best ideas but the findings from the case study echos the idea that meritocracy in tech is a myth that ignores unequal access. For example, YC's emphasis on founders have great credentials tends to filter out those without Silicon Valley or deep big tech connections. A consistent stream of YC success stories from the same founder profile attracts

similar applicants while YC's selection process reinforces that same notion, that the "best" founders and ideas ceom from the same group of male, Silicon Valley-style entrepreneurs. As a result, alternative voices such as female-led or community-oriented ventures are sidelined in the process without exposure to a top hub of funding and development. Through the data, we can also see that greater than 50% of the startups pushed through YCombinator in their recent Winter 2024 batch were AI-related, which heavily relates to the explosion of AI-company funding in the years between 2019 and 2024. This for-profit narrative forms our view of AI just as much as the idea of the ideal founder forms the mold of who gets to participate in AI. In summary, the results, as they relate to venture capital and other early stage investment such as YC, show that the early-stage AI innovation landscape is skewed by gender, privilege, and a unrelenting drive for profit. An imbalance that dictates which technologies get built and which don't. These biases at the early stages of development set up a downstream effect to be explored in the next section.

Environmental and Societal Impacts of AI Infrastructure

The research findings underscore that the growth of AI has significant impacts on the environment, concentrated in communities that house data centers. As we saw in the Northern Virginia case study, multiple data points highlight this burden. Loudoun Count, Virginia has roughly 200 data centers as of the last few years excessive amounts of power consumption. This has led to concerns of grid integrity and contributed to rate increases for local residents. Additionally, evidence from Monserrate have shown that the continuous draw of power and water for the datacenters is put ahead of the needs of the community (Monserrate, 2022). We also see that water is being used at a significantly higher rate from 2019-2023.

AI's physical infrastructure is a resource-intensive operation with climate implications. These burdens are not, however, shouldered by the tech companies that build them, but by both

the environments and the areas residence as explained later. In a way this concept can be defined as “digital gentrification.” As I have described in research findings, there has been a stark increase in housing prices in Northern Virginia (Staff, 2024), a policy driven growth of data center construction (Carey, 2025), and a rising cost of utilities amongst the region. These factors exarcerbate housing unaffordability for lower-income residents and force vulnerable households to weight utility costs against other essential expenses (Hernandez, 2023).

Ultimately, the real significance of data center proliferation lies not just in emissions or resource consumption, but in the social landscape it reshapes, where benefits are tightly concentrated and costs diffuse through the community.

Digital Plantation

In many ways, the gentrification and excessive land use in Northern Virginia has been labeled a “digital plantation” (Rosati, et al., 2023). This dynamic is very evident in this region, as the wealthy tech companies harvest the benefits (data, profit) from the digital infrastructure, local communities are left with the environmental degradation and social upheaval. In Rosati et al.’s paper, they specifically mention Northern Virginia as an almost plantation-era exploitation, noting it features “resource monopolies, extractive exploitation, and monocrop ecologies – based on the “Server Farming” industry.” (Rosati et al., 2023). My results support this analogy. We found that tech corporations exert outsized influence over local policy in order to gain access to the area. This power imbalance silences or overrides the voice of marginalized residents. In essence, we circle back to the idea of the beneficiaries of AI’s boom such as VCs, tech firms, and datacenter operators, holding the power while those bearing the cost lack a comparable voice. This clearly draws a picture of a full-circle of inequity as one could compare the individuals who are left out the creation of AI are now left to face the burden of the technology they had no voice

in. One of this paper's guiding questions, "why do we keep pushing AI despire its harms," is answered hear. Those in power reap outsized rewards and can externalize all of the harms to others, which leaves them no incentive to change.

Unifying The Lifecycle

The relentless drive for scale in AI, fueled by venture capital checks and a hunger for rapid returns creates a feedback loop: more funding enables more development, which demands more data center infrastructure. Thus causing a compounding effect of environmental and social costs. Tying these strands together reveals that AI's entire life cycle, from initial capital injection to physical deployment, reinforces inequalities. By focusing on swift profit and market dominance, AI developers risk burdening marginalized communities with the costs of innovations they did not help design and may not directly benefit from.

Conclusion

This research has traced the lifecycle of AI from it's instantiation and funding to the data centers required for its deployment. This unconvered a consistent them: structural biases lead to an unequal distribution of AI's benefits and harms. This reiterates the main insight that has emerged from this research that the AI revolution, in its current state, mirrors and amplified the pre-existing social inequities. Venture capitalist and tech bosses are steering AI development and as a result, the goods that come from AI development accrue to those who are already more than well-off. Meanwhile, the costs such as environmental degradation, higher costs of living, and resource burden, are put onto marginalized groups who have the least say in these initiatives. The answer to my research question of who benefits and who pays is answered clearly: the

beneficiaries are the tech elites and investors and the individuals who pay the cost are disproportationately vulnerable individuals, communities and the environment as a whole.

I believe the most important implication of this work is the pressing need to reimagine AI and its development from the beginning until the end. We see at the beginning of the lifecycle that the narrow group of founders that are chosen to lead the charge are creating AI to server its specific purpose. The venture capital and tech world need to push to fund more equitably, for example, increasing funding into women-only startups beyond just the occasional headline deal. At the end of the lifecycle, we see a strong need for policy to be reimagined as it relates to the creation of AI infrastructure. Technological development needs to be in line with social justice and sustainability, otherwise we will see the same groups who have always been taken advantage of, be harmed even further.

I do acknowledge that the trajectory and intentions behind all ideas are not equal. We cannot expect similar results from all venture capital funding situations and AI companies, however the overarching evidence points to this conclusion. Beyond that, the concluding message of this research is a call to action: to ensure AI's lifecycle, from funding to datacenter building, evolves towards equity and away from the profit-hoarding initiatives that are currently in place.

References

1. Carey, M. (2025). How Data Center Alley Is Changing Northern Virginia. Oxford American. https://oxfordamerican.org/oa-now/how-data-center-alley-is-changing-northern-virginia
2. Chung, R. (2024, November 18). YC Founders Categorized by University [Unpublished manuscript]. University of Virginia.
3. Hernandez, D. (2023, June 29). Energy Insecurity And Health: America's Hidden Hardship. Health Affairs. https://www.healthaffairs.org/do/10.1377/hpb20230518.472953/
4. Hu, K. (2025, January 7). AI startups drive VC funding resurgence, capturing record US investment in 2024. Reuters. https://www.reuters.com/technology/artificial-intelligence/ai-startups-drive-vc-funding-resurgence-capturing-record-us-investment-2024-2025-01-07/
5. Li, P., Yang, J., Islam, M., & Ren, S. (2025). Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models. https://arxiv.org/pdf/2304.03271
6. Lohn, A., & Musser, M. (2022). AI and Compute: How Much Longer Can Computing Power Drive Artificial Intelligence Progress? CSET Issue Brief. https://cset.georgetown.edu/wp-content/uploads/AI-and-Compute-How-Much-Longer-Can-Computing-Power-Drive-Artificial-Intelligence-Progress_v2.pdf
7. Monserrate, S. G. (2022). The Cloud Is Material: On the Environmental Impacts of Computation and Data Storage. MIT Case Studies in Social and Ethical Responsibilities of Computing, Winter 2022. https://doi.org/10.21428/2c646de5.031d4553

8. Monserrate, S. G. (2022, February 14). The Staggering Ecological Impacts of Computation and the Cloud. The MIT Press Reader. https://thereader.mitpress.mit.edu/the-staggering-ecological-impacts-of-computation-and-the-cloud/
9. Robbins, J. (2025, January 9). AI startups grabbed a third of global VC dollars in 2024. PitchBook. https://pitchbook.com/news/articles/ai-startups-grabbed-a-third-of-global-vc-dollars-in-2024
10. Rosati, C., James, A., & Metcalf, K. (2023). Data plantation: Northern Virginia and the territorialization of digital civilization in "the Internet Capital of the World." Online Media and Global Communication, 2(2), 199–227. https://doi.org/10.1515/omgc-2023-0017
11. Shah, D. (2024, October 15). Data-driven insights into Y-Combinator. Medium. https://medium.com/@devanshshah1309/data-driven-insights-into-y-combinator-f8b695e184bf#The%20Rise%20of%20AI
12. Staff, U. (2024, May 24). The 10-year path of Northern Virginia home prices in four charts. https://dc.urbanturf.com/articles/blog/the_10-year_path_of_northern_virginia_home_prices_in_four_charts/22350
13. Syal, A. (2022, July 14). Essential Takeaways From Stanford's AI Index 2022 Report. Datacamp.com; DataCamp. https://www.datacamp.com/blog/essential-takeaways-from-stanfords-ai-index-2022-report
14. Tan, G. (2024, March 4). Meet the YC winter 2024 batch. Y Combinator. https://www.ycombinator.com/blog/meet-the-yc-winter-2024-batch

15. Teare, G. (2024, June 5). The portion of US VC funding that went to female founders hit a new peak in 2023, thanks to massive AI deals. Crunchbase News. https://news.crunchbase.com/diversity/us-vc-funding-female-founders-peaked-2023-ai-openai-anthropic/
16. Turner, M. (2025, January 1). A Strategy for a Changing Paradigm. https://www.loudoun.gov/ArchiveCenter/ViewFile/Item/13979
17. University of Arizona Libraries. (2023). What is a large language model (LLM)? https://ask.library.arizona.edu/faq/407985
18. van Tilburg, T. (2021). Living next to the cloud: The impact of data center development on housing prices. Master's Thesis, University of Groningen. https://frw.studenttheses.ub.rug.nl/3724/