September 16, 2025

The Honorable Darrell Issa, Chairman
Subcommittee on Courts,
Intellectual Property, Artificial Intelligence,
and the Internet
Committee on the Judiciary
2138 Rayburn House Office Building
Washington, D.C. 20515

The Honorable Hank Johnson, Ranking Member
Subcommittee on Courts,
Intellectual Property, Artificial Intelligence,
and the Internet
Committee on the Judiciary
2142 Rayburn House Office Building
Washington, D.C. 20515

RE: Hearing entitled "AI at a Crossroads: A Nationwide Strategy or Californication?"

Dear Chairman Issa, Ranking Member Johnson, and Honorable Members of the Subcommittee on Courts, Intellectual Property, Artificial Intelligence, and the Internet:

Engine is a non-profit technology policy, research, and advocacy organization that bridges the gap between policymakers and startups. Engine works with government and a community of thousands of high-technology, growth-oriented startups across the nation to support a policy environment conducive to technology entrepreneurship. As artificial intelligence is increasingly used, developed, and deployed by startups, Engine has a strong interest in ensuring a regulatory environment for AI conducive to startup success at home and in markets around the globe.

Startups are driving innovation in AI benefitting every corner of the economy—from agriculture to manufacturing, healthcare to education, finance to retail, and more.[1] But policymakers' approach to regulating AI will determine who is able to participate in the AI ecosystem and the speed at which innovations are disseminated to benefit the public. The U.S. has generally followed a model of permissionless innovation,[2] enabling entrepreneurs to build beneficial new products unencumbered by strict, expensive regulatory regimes antithetical to invention, experimentation, and iteration. This approach has made the U.S. tech sector the envy of the world.

But poorly conceived AI regulatory frameworks that are overbroad and over-reliant on ex-ante approaches threaten to undermine U.S. startup competitiveness and innovative capacity. The U.S. must avoid importing such costly frameworks like the European Union's AI Act. According to EU-funded estimates, the EU AI Act will create well over $200,000 in initial compliance costs and

[1] *See, e.g.*, *#StartupsEverywhere*, Engine, https://www.engine.is/startupseverywhere.
[2] *See, e.g.*, Adam Thierer, *Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom*, Mercatus (Mar. 2016), https://www.mercatus.org/research/books/permissionless-innovation-continuing-case-comprehensive-technological-freedom.

nearly $80,000 in annual ongoing costs.[3] A think tank analysis released shortly after pegged those costs even higher, at nearly half-of-a-million dollars in initial costs.[4] For comparison, a seed-stage startup is working with around $50,000 per month in resources—meaning such costly frameworks can literally take months off of their life.[5]

Individual U.S. states are unfortunately adopting elements of the costly EU approach and are pursuing their own unique AI rules, setting the stage for a patchwork of varying regulations that will steer where startups scale, cause them to degrade the quality of their products, and undermine their competitiveness. Startups have experience with burdensome state patchworks of regulations on issues from HR to data privacy that illustrate the threat to competitiveness that differing rules about the same topic pose for small companies. On data privacy, startups invest hundreds of thousands in privacy compliance, but face $15,000-$60,000 for each time states add or amend their laws.[6] Replicating this patchwork for AI is certain to undermine U.S. leadership.

We appreciate the Subcommittee's attention to this important issue for U.S. AI startups and the recognition that rules for AI being set at a national level is essential for startup competitiveness. We look forward to being a resource for the Subcommittee as you continue to explore AI issues and push to enhance America's innovative capacity. To that end, we have attached a recent report published by Engine below, *Mapping the future: charting the AI ecosystem & a policy blueprint for startup success*, which examines how U.S. startups are building with AI and how policymakers can support their success.

Sincerely,

Engine Advocacy

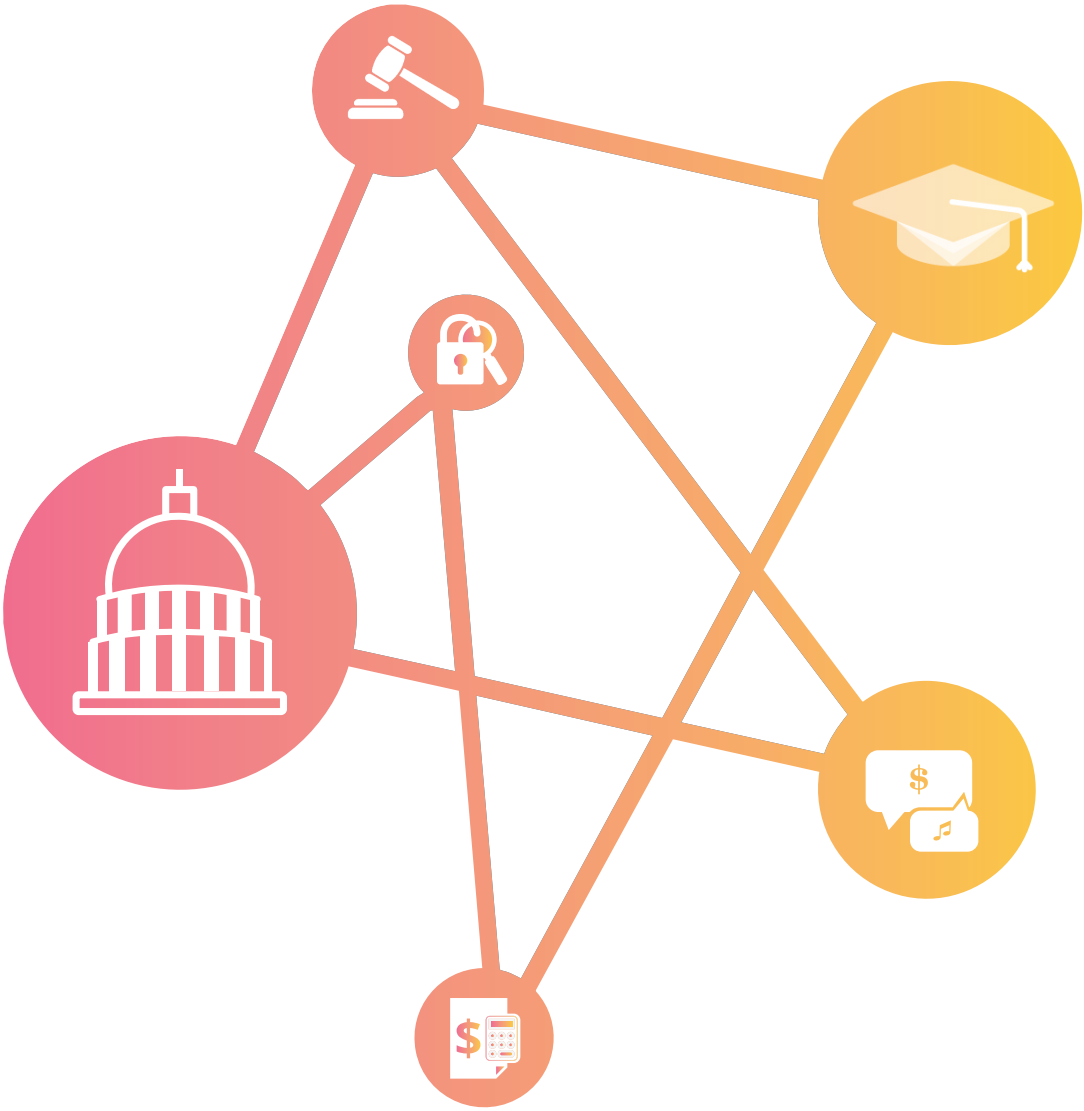
700 Pennsylvania Ave SE
Washington, D.C. 20003
policy@engine.is

[3] Andrea Renda, et al., *Study to support an impact assessment of regulatory requirements for Artificial Intelligence in Europe*, European Commission (Apr. 2021), https://op.europa.eu/en/publication-detail/-/publication/55538b70-a638-11eb-9585-01aa75ed71a1.
[4] Benjamin Muller, *How Much Will the Artificial Intelligence Act Cost Europe?*, Center for Data Innovation (July 2021), https://www2.datainnovation.org/2021-aia-costs.pdf (estimates converted from Euros to U.S. Dollars).
[5] *the State of the Startup Ecosystem*, Engine (Apr. 2021), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/60819983b7f8be1a2a99972d/1619106194054/The+State+of+the+Startup+Ecosystem.pdf.
[6] *Privacy Patchwork Problem: Costs, Burdens, and Barriers Encountered by Startups*, Engine (Mar. 2023), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/6414a45f5001941e519492ff/1679074400513/Privacy+Patchwork+Problem+Report.pdf.

Engine

Mapping the future:

charting the AI ecosystem &
a policy blueprint for startup success



May 2025

ABOUT ENGINE

Engine was created in 2011 by a collection of startup CEOs, early-stage venture investors, and technology policy experts who believe that innovation and entrepreneurship are driven by small startups, competing in open, competitive markets where they can challenge dominant incumbents. We believe that entrepreneurship and innovation have stood at the core of what helps build great societies and economies, and such entrepreneurship and invention has historically been driven by small startups. Working with our ever-growing network of entrepreneurs, startups, venture capitalists, technologists, and technology policy experts across the United States, Engine ensures that the voice of the startup community is heard by policymakers at all levels of government. When startups speak, policymakers listen.

CONTENTS

INTRODUCTION

Artificial intelligence is a foundational technology driving innovation in every corner of the economy. Rapid advances and breakthroughs in AI have drawn policymakers' attention—as evidenced by the flurry of hearings, task forces, executive orders issued, and bills introduced. These breakthroughs have enabled a flourishing ecosystem of startups building AI, building with AI, and using AI to better everyday tasks. As policymakers continue to dive in to varying facets of AI with proposals, guidance, frameworks, and incentives, it is imperative that they have a solid grasp of the breadth and interconnectedness of the AI ecosystem.

> *One of the things that concerns me is when policy and AI are discussed, policymakers only focus on how to regulate big corporations and they unintentionally hurt a ton of businesses and startups that use AI...*[1]
>
> **- Constanza Gomez, Co-Founder, Sortile, New York, New York**
>
> *Sortile is a technology platform designed to identify textile types in order to facilitate recycling.*

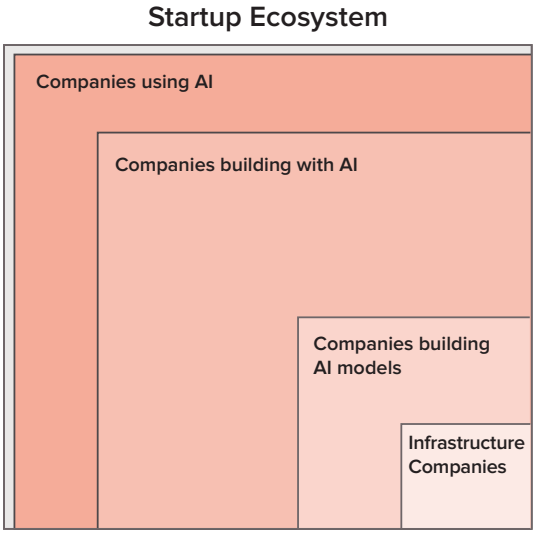
By volume, startups make up the vast majority of the AI ecosystem. But like many technology policy issues, the perspectives of startups can be drowned out by attention on the largest players in an industry.

Startups are building with AI in a few distinct ways. Some startups are building their own machine learning models to perform specific tasks. Most startups are leveraging foundation models—either by licensing from market leaders or accessing open source—to fine-tune and build unique products. Often, startups leverage multiple foundation models, finding some perform better at certain tasks and others better at different tasks. This increasingly describes how startups are building—90 percent of the startups that joined our network in the past year used AI in their main product. And finally, most startups are leveraging others' AI tools as part of their product or to augment basic business functions.

AI is nearly ubiquitous in the startup ecosystem.

Startups are implicated at every level of the AI ecosystem. For example, a startup leveraging the latest foundation model to create a product giving feedback on startups' pitch decks sits in the middle and touches companies up and down the stack. Policies impacting one part of the stack—like rules on deploying AI, or another, like requirements for foundation model development—will reverberate across the ecosystem.

Policymaking in the AI space requires not one bill or one framework, but smart policies across all issues impacting AI developers, deployers, and users. To chart a course forward, policymakers must begin with a strong foundation.



> *Policymakers need a deeper understanding of this space and the differences in the participants, because there is a general lack of knowledge about AI's transformative potential, leading to an overemphasis on worst-case disaster scenarios. They need to know the costs and impacts on smaller firms like ours.*[2]
>
> **- Vance Reavie, CEO and Founder, Junction AI, Bentonville, Arkansas**
>
> *Junction AI is an AI platform that automates marketing, merchandising workflows, and content generation processes for brands and retailers.*

GLOSSARY

There are few simple definitions in the AI space, and the difficulty of defining and scoping AI topics becomes increasingly so as policymakers put pen to paper. Rather than inform legislative text, this glossary is designed to help stakeholders understand and navigate key terms and topics in AI.

AI: Technology that imitates or enhances human capabilities to perform tasks like communicating, reasoning, problem-solving, understanding language, recognizing images, and more.

> The John S. McCain National Defense Authorization Act for Fiscal Year 2019 includes a five-part definition of AI that is relied upon by most federal agencies and other stakeholders.[3]
>
> > (1) Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.
> >
> > (2) An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.
> >
> > (3) An artificial system designed to think or act like a human, including cognitive architectures and neural networks.
> >
> > (4) A set of techniques, including machine learning, that is designed to approximate a cognitive task.
> >
> > (5) An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision making, and acting.

Copyright: An intellectual property right designed to benefit the public by incentivizing the creation and dissemination of works by granting creators exclusive rights to their works. Copyright protects the expression of an idea, it does not protect the idea itself or facts.

Compute: The hardware resources—processors (CPUs, GPUs, or TPUs), memory, and storage—that make AI models work, allowing them to train on data, process information, and generate predictions.

Digital replica: Videos, images, or audio recordings that have been digitally created or modified using AI to realistically depict an individual using their Name, Image, and Likeness, and other indicators of identity.

Fair use: A doctrine where copyrighted content may be used without the rightsholder's consent for purposes like criticism, comment, news reporting, teaching, scholarship, or research. Determining fair use involves a four-part test: the purpose and character of the use; the nature of the work; the amount and substantiality used; and the effect on the market for the work.

Fine-tuning: The process of adapting a pre-trained model to perform specific tasks. Rather than training a model from scratch, fine-tuning leverages the general knowledge already acquired by the model during its initial training, specializing it for a developer's specific needs.

FLOPs / FLOPS: FLOP stands for FLoating-point OPeration. A floating-point operation is one arithmetic operation — such as addition, subtraction, multiplication, and division — on a floating-point number. The amount of floating-point operations, or FLOPs, is a measure of how much computational work a model requires to process data and make decisions. Meanwhile, FLOPS — FLoating-point Operations Per Second — measures a computer system's computational performance, quantifying the number of calculations it can perform per second.

Foundation model: Large-scale AI model trained on vast amounts of diverse data, making them adaptable to a variety of tasks. These models serve as the backbone for many AI applications, providing a base level of knowledge that other developers, like startups, can fine-tune for specific uses.

Frontier model: A foundation model considered to be at the leading edge of current capabilities.

Large Language Model (LLM): A type of foundation model focused on natural language processing.

Machine learning: A branch of artificial intelligence where algorithms and models are developed to make predictions based upon data.

Model weights: Numerical parameters learned during training that determine the importance of features in a dataset. Weights shape model behavior making them an essential part of model accuracy and utility.

Open source: Open source has historically described software that "anyone can view, modify, and distribute." In AI, this generally means making the model, or the algorithms, code, weights, and data used for AI development available to the public.

Pre-training: Initial training phase in model development where the model is exposed to a large range of data sources to learn general patterns and information relationships.

Training: Following pre-training, training involves exposing the model to specific types of data to improve its performance or perform a specific task. Training and fine-tuning are sometimes used interchangeably.

AI ESSENTIALS

Artificial Intelligence can be technical, but it is essential that stakeholders have a grasp of the key topics in AI and how they intersect with startups and public policy. These topic-based explainers aim to bridge the knowledge gap, complete with useful analogies and takeaways for both startups and policymakers.

HOW DO NEURAL NETWORKS WORK?

A **neural network** is a computer model that learns to identify complex relationships between input and output data. Structurally, it consists of interconnected layers: an input layer, several hidden layers, and an output layer. Information flows from the input to the output layer through these hidden layers.

Layers are composed of **nodes**, basic information processors. Each node takes a piece of information (input), applies a mathematical function to it, and then passes the result (output) to the other nodes in the next layer. Connections between nodes across layers carry **weights**, numerical values that determine the importance of the transmitted information. The network "learns" by adjusting these weights based on the errors it makes.

You can think of a neural network as a water filtration system. Each layer represents a filtration stage: the input layer is where raw, murky water enters, and the output layer is where clean water exits. Each node is an individual filter unit, where the amount of filtering it performs can be calibrated based on water quality, similar to how weights can be adjusted to amplify the most significant pieces of information. The system fine-tunes each filter to produce the cleanest water possible, much like how a neural network adjusts weights to produce the most accurate outputs.

Training most neural networks involves two main steps. First, the network is given a dataset called **training data**, which consists of examples with known outcomes (labels). Starting with random weights, the network takes examples, makes predictions, compares them to the correct outcomes, and then adjusts its weights to reduce errors. As training progresses, the network's predictions get better and better. This process is repeated for many cycles (epochs) until these adjustments no longer result in significant improvements to the network's performance, at which point the model has **converged**.

After training, the network is tested on a separate pool of data known as **testing data.** This testing phase evaluates the network's prediction accuracy, showing its ability to generalize and handle real-world data outside of training examples.

Data is the lifeblood of neural networks because it fuels their learning process.[4] These networks identify patterns by adjusting weights to highlight relevant information and minimize attention to less important details. Through iterations, the network learns to recognize which features in the data are most important for making accurate predictions. The more comprehensive and diverse the dataset a network trains on, the better it becomes at identifying pertinent information and understanding patterns, which enhances its performance in real-world applications.[5]

For startups in the AI space, access to abundant, high-quality data is crucial. Larger tech companies typically have vast data resources from existing services or partnerships that startups often lack. This lack of data can significantly impede their ability to build models that are accurate and reliable. Therefore, open data initiatives, public-private partnerships, balanced intellectual property frameworks, and uniform regulatory environments play a critical role in helping startups obtain the data necessary to build robust neural networks, foster competition, and drive innovation in AI.[6]

WHAT ARE MODEL WEIGHTS?

As AI systems take in vast amounts of **data**, they have to determine which characteristics of the data are important. For example, if an AI model is being trained to differentiate between dogs and cats, the model will place more importance on relevant distinguishing features (like the shape of the ears or length of nose) and less importance on less relevant features (like the number of legs or color of fur).

These distinctions are reflected in **model weights**, which are numerical parameters that determine the importance of features in a dataset. Highly complex AI systems can have billions of weights — like GPT-3, which has over 175 billion model weights.[7] You can think of weights as volume knobs that control how much influence each input (like an image detail or a text entry) has on the final decision by the AI. During training, these weights are continually adjusted as the model learns from the data, refining its accuracy by emphasizing and de-emphasizing certain inputs.

Model weights play a crucial role in determining the outputs of AI systems and access to model weights can enable an individual to make beneficial changes to the model (addressing a biased result or creating a new product) or to make malevolent changes (allowing the model to create harmful or illegal content). For certain models, like one designed for fraud detection, securing model weights is tantamount because access to weights could enable criminal circumvention.

Conversely for general-use models, access to model weights allows for transparency and enables innovation. These models, often called **open source** or **open-weight models**, are pre-trained models with publicly available weights. Since the functionality of AI relies primarily on the configuration of weights, pre-trained models enable startups to sidestep the significant data and **compute** resources required for training AI from scratch. Furthermore, open-weight models promote transparency and allow for more thorough scrutiny by researchers and policymakers,[8] helping address many core concerns related to AI outputs like biases or hallucination.

Policymakers have highlighted these tradeoffs, for example, the Biden Administration's AI executive order directed the study of such models.[9] The resulting report highlighted the benefits — for innovation, research, and transparency — and risks of abuse.[10] The report concluded restrictions were not appropriate, but open source and access to model weights promises to be subject of AI policy debates into the future.

WHAT IS COMPUTE AND HOW IS IT MEASURED?

If you wanted to haul large amounts of water from a well, you would (at one point in history) need a resource, like a horse, to pull the load. You would measure how much weight the horse can pull per minute in terms of "horsepower." In the world of AI, the resource required to make models work is **compute**, and it's measured in **FLOPS**.

Compute refers to the hardware resources that make AI models work, allowing them to train on data, process information, and generate predictions. Without sufficient compute, even the most sophisticated models struggle to perform efficiently. The implications for startups and policymakers are twofold. First, compute is scarce and expensive, meaning startups are constrained by their access to compute. And second, given compute's role in the effectiveness of models, policymakers are including compute-based thresholds in regulatory frameworks they are pursuing.

Compute involves the processors (CPUs, GPUs, or TPUs), memory, and storage needed to perform the numerical calculations for AI models. These resources are especially critical during **training**, where models adjust internal parameters (called **model weights**) based on patterns found in massive **datasets**. Having more compute means models can more quickly and effectively learn from data, leading to more accurate predictions, improved decision-making capabilities, and the ability to handle more complex tasks. Inadequate compute, on the other hand, limits a model's complexity, slows down training, and can hinder innovation.

We can measure both the computational work required by AI models, as well as the theoretical capacity of compute resources using units with confusingly similar names — **FLOPs** and **FLOPS**.

FLOP stands for FLoating-point OPeration. (A floating point number is a standardized format in computing to precisely and uniformly encode large and small values. A floating-point operation is one arithmetic operation — such as addition, subtraction, multiplication, and division — on a floating-point number.[11]) The amount of floating point operations, or FLOPs, is a measure of how much computational work a model requires to process data and make decisions.

The higher the FLOPs, the more complex the model and the more compute it demands. Older models may require trillions (or 10^{12}) FLOPs, but today's leading AI models demand compute on a massive scale due to the enormous datasets they process and the intricate neural networks they utilize. For example, training a model like GPT-4 can require septillions (or 10^{24}) FLOPs of compute.[12]

Meanwhile, FLOPS — FLoating-point Operations Per Second — measures a computer system's computational performance, quantifying the number of calculations it can perform per second. It is a measure for how powerful a given piece of hardware is or its theoretical capacity — the higher the FLOPS, the more powerful the hardware.

Both measures are used in AI regulatory efforts. For example, the Biden Administration Executive Order on Artificial Intelligence last fall included disclosure requirements for models trained with more than 10^{26} FLOPs of compute and for compute clusters with theoretical capacity of more than 10^{20} FLOPS.[13] Europe's AI Act and a (since vetoed) controversial California bill also use compute-based thresholds.[14] And given the financial costs of large quantities of compute, some policymakers have also included cost-of-compute-based thresholds in regulatory efforts.

It's unclear whether these thresholds — which are arguably arbitrary — will hold up over time. Technological improvements in both models and in compute will lead to more capable models with lower compute requirements and lead compute costs to fall.

At present, compute remains a main cost center for startups in AI (and for startups with their own compute resources, associated costs like energy and cooling). Most startups lack the resources to invest in their own infrastructure and must rely on cloud services to access the necessary compute power. That means startups often compete and approach AI development on a different plane, either developing niche models to perform specific tasks (as opposed to a large language model), or **fine-tuning** others' pre-trained models. Larger companies by comparison can innovate more freely with in-house compute or even afford to invest in custom hardware, such as Tensor Processing Units (TPUs).

Policymakers are now exploring ways to level the playing field by ensuring that startups can access the compute they need to innovate. Some proposed solutions include the National AI Research Resource, providing access to high-performance compute resources and datasets for academic researchers and smaller companies — helping to democratize access to the tools necessary for AI development for a more equitable and innovative AI ecosystem.[15]

WHAT ROLE DOES DATA PLAY AND WHERE DOES IT COME FROM?

Data is a fundamental resource that powers all AI systems. There are different types of data that are utilized for different functions in AI development, and varying sources of that data. Both of these can be highly context-dependent on stages of development, resources that the developer has on hand, and the task(s) that the model is being trained to perform. Legislators, regulators, and rightsholders each can impact the availability and types of data used for AI, and shape which companies can participate in AI innovation in the first place.
AI systems require not just large quantities of **data**, but data that's properly structured, labeled, and divided for different phases of development. The specific requirements depend on the model's purpose and learning method.

In **supervised learning**, we give the AI system examples with clear "answers" or **labels** — imagine teaching a child by showing them pictures of animals and telling them "this is a cat" or "this is a dog." For a medical AI, this means X-rays must be marked by experts exactly where fractures appear to teach the system what a broken bone looks like and where to find it.

In contrast, **unsupervised learning** uses **unlabeled** examples, allowing the AI to find patterns on its own. This is like asking someone to sort a drawer full of socks without telling them how; they might stumble across groups by color or size as the most effective, thereby 'learning' the groups. In AI applications, this can help discover natural groupings without predefined categories.

Meanwhile, **semi-supervised learning** combines both approaches, using a small amount of labeled data together with larger amounts of unlabeled data. This is particularly valuable when labeling is expensive or time-consuming, like in medical imaging where specialist doctors would need to annotate thousands of images.

Different stages of AI development require datasets that serve distinct purposes. **Training data** teaches initial patterns, while **testing data** — kept entirely separate — verifies whether the AI can handle new situations. For medical imaging, this means not just training on thousands of fracture X-rays, but testing on a separate set of cases to ensure the AI can spot breaks it hasn't seen before.

For data, quality often matters more than quantity. Good training data needs to be accurate, well-labeled, and representative of real-world scenarios. A customer service AI tool trained mostly on routine queries will struggle with complex problems, even if trained on millions of examples. Similarly, an autonomous vehicle AI model learns more from a few thousand carefully annotated hazard scenarios than millions of normal driving situations.

Such specialized data can originate from various sources. Some datasets are publicly available, such as government weather data or academic research. However, many valuable datasets are often collected from real-world interactions, high-quality sources, or must be curated specifically for the intended AI application. Large technology companies generally hold a competitive advantage given the significant differences in the resources available to curate those datasets, the ability to negotiate licenses or fight lawsuits with rightsholders, and the access to large volumes of highly relevant data collected through user interactions.

AI companies source data from a range of sources, including from themselves or their users. Information directly collected from users, or **first-party data**, can be especially useful for improving models because it captures real-world interactions and is continuously updating. For example, an AI product designed to detect elderly people falling benefits from such data to differentiate falling versus sitting in a chair. Large companies with lots of users have access to more interactions that help improve their models. In contrast, startups typically don't have large user bases or multiple services generating continuous data streams. Data privacy rules and regulators also factor in here, with the Federal Trade Commission last year warning AI companies against changing their terms to leverage user data for AI training.[16]

Many companies large and small **ingest** data from the open web or from open data sources. **Open data sources** are freely accessible datasets that anyone can use, often maintained by public institutions, governments, or nonprofits. These sources provide essential resources for startups that lack extensive proprietary data. The non-profit Common Crawl,[17] for instance, maintains a public web archive of over 9.5 petabytes of data, accessible to anyone–from small startups to big players such as Stable Diffusion, who use filtered versions of this data through another nonprofit organization, LAION.[18]

Original, expressive content — from everyday folks and the most well-known organizations — is plentiful online and therefore throughout many datasets scraped from the public corpus. Large rightsholder organizations and well-known celebrities have alleged ingesting this data for training purposes amounts to copyright infringement and filed lawsuits against the largest AI companies and startups alike.[19] Some large companies have negotiated agreements

to license data for AI from these entities in response.[20] Those deals can run into many millions of dollars annually, beyond the budgets of startups, meaning startups might not be able to participate in AI innovation. Ingesting data — facts about the world — to learn is not an infringing practice (see below section on Copyright for further discussion of this topic), and policymakers will need to support this understanding if the AI ecosystem is to remain competitive. (Outputs, on the other hand, can sometimes be infringing, and if stakeholders are worried about AI diluting human creativity, considering whether outputs are infringing is a key part of the solution.)

Other types of data used for AI include **custom datasets** compiled to tailor the data to meet specific industry needs. For example, an autonomous vehicle company might collect video footage of different driving scenarios, then label objects like pedestrians, cars, and traffic signs in each frame. This purpose-built dataset ensures that the data aligns precisely with the company's goals, but it requires significant time, resources, and infrastructure to organize and label at scale.

To aid where this real-world data is scarce, some companies may use **synthetic data** — artificially generated data that mirrors real-world patterns. Rather than collecting thousands of real-world images of rare driving scenarios, a company could simulate various traffic conditions, weather patterns, and road layouts and record that data. This approach can help expand the dataset quickly and provide diverse examples that might be challenging to capture in real life. However, recent research warns models may become overly tuned to synthetic patterns that do not align with real-world data, resulting in "model collapse" when training exclusively on synthetic data.[21]

Policymakers — on issues from data privacy to intellectual property rights — have wide remit to impact the competitiveness of the AI ecosystem (or lack thereof), depending upon the actions they take when it comes to data and AI. They should seek a balanced and competitive landscape that ensures small startups with few resources can continue to innovate, grow, and compete.

HOW DOES COPYRIGHT INTERSECT WITH AI?

Startups are increasingly innovating in AI, but unresolved questions about copyright hang like a sword of Damocles over the entire AI ecosystem. Numerous ongoing AI lawsuits turn on whether including copyrighted content, such as written works, images, or music, in datasets to train generative AI models constitutes infringement.[22] How these cases are resolved will determine the pace of AI innovation and whether startups can afford to participate in the AI ecosystem at all.

Much of the current wave of innovation in AI is based upon closed or **open-source foundation models** that startups often **fine-tune** to perform a specific task. These models are trained on a large corpus of training data inputs so that they can accurately learn about the world and document the relationships between words, pixels, tones, and more. Many large models learn from what is publicly available on the Internet. And since most content created is copyrightable — including anything expressive, such as articles, songs, and even meaningless tweets[23] about brushing your teeth — this training data may include copyrighted content.

Copyright law is designed to benefit the public by incentivizing the creation and dissemination of works through granting creators exclusive rights to their works. While copyright protects the expression of an idea, it does not protect the idea itself or facts. Additionally, the fair use doctrine allows copyrighted content to be used without the rightsholders' consent for purposes like criticism, comment, news reporting, teaching, scholarship, or research when the application is weighed in favor of fair use. These limitations and exceptions are key to understanding how debates around copyright and AI should be resolved.

For startups, it is crucial that using copyrighted content as inputs does not constitute infringement and falls under fair use.[24] Generative AI models learn from inputs similar to how humans learn from articles, books, or art to produce a new creation. AI models use inputs to understand and interpret concepts. This learning process does

not result in the AI model directly copying the content it is trained on, but instead documenting relationships and patterns as vectors. Once training is finished, a model can produce outputs, like new writing or images, informed by those relationships.

You can think of an AI model as a composer who takes tempos, rhythms, and chord progressions from various known classical pieces to create a new composition. Similar to how the composer pulls and combines existing musical elements to generate new, original music, AI models learn from and interpret existing creative content to then produce a new output. (AI models produce unique creations, but users can still prompt outputs that are materially similar to copyrighted material, often despite safeguards designed to prevent it.)

If inputs are deemed to be infringing, courts will need to determine whether fair use justifies using copyrighted content in training data. The fair use doctrine allows for certain unauthorized uses of copyrighted content, and when evaluating a fair use defense claim, courts weigh four factors: the purpose and character of the use, the nature of the copyrighted work, the amount and substantiality of the portion use, and the effect on the market.

In the context of AI training, these factors support an applicable fair use defense. The purpose and character of AI's use of copyrighted content is transformative because it does not replicate the original content but generates new outputs with added character or a further purpose. Regarding the nature of the work, AI processes creative inputs in a way that transforms it into new content. While AI models may ingest entire copyrighted works, they do not simply replicate this content in its entirety in the output. Finally, AI's use of copyrighted content is unlikely to harm the market for the original work because AI developers don't directly profit from the distribution of the generated content. Instead, they monetize the tools that enable users to generate outputs.

While these fair use factors are generally favorable for AI developers, the application is not always straightforward, and different cases can lead to varying outcomes. Fair use is a fact-specific, case-by-case determination made by the courts. For AI companies, particularly startups, this legal uncertainty poses a significant challenge. The risk of facing a copyright infringement suit can result in costly litigation and could stifle innovation.

Rightsholders — like publishers, authors, record labels, and others — have construed AI training as infringement, saying it constitutes an unauthorized use of their works. Many have sued AI companies of all sizes, ranging from startups to market leaders. If the rightsholders prevail, it will upend the AI innovation ecosystem.

Under such an environment, AI developers would need to seek licenses before they can begin development. In response to the current legal uncertainty, larger AI companies are opting to make licensing agreements for copyrighted content to avoid potential lawsuits or settle ongoing litigation.[25] This is unworkable for startups and threatens smaller companies' participation in the AI ecosystem. Licenses would be prohibitively expensive for startups and negotiations unbalanced given the comparative size difference between small startups and large rightsholder organizations. This would make it both difficult for startups to fine-tune existing models with anything besides their own data (which is often in short supply as new entities), and make it near impossible for a new startup to challenge existing players in frontier model development.

> *For smaller companies like us, what happens with the copyright litigation for the largest entities is going to trickle down and impact what we are doing. We are concerned about how it will affect us, but at the end of the day, we cannot spend large amounts of money on it because we don't have the same access to capital that these larger AI companies do.*[26]
>
> **- Chandler Malone, Co-founder and CEO, Path, New York, New York**
>
> *Path provides AI test-prep tools for standardized exams and professional certifications.*

Moreover, requirements to license training data would remake the ecosystem most startups build upon. Licenses will increase costs for frontier model development, which will be passed on to startups that rely on access to them. Open source development will be harmed as open source developers are unlikely to release their models for free and with documentation that reveals how they were trained and what they 'know.'

To promote innovation and maintain a competitive AI ecosystem, startups need legal clarity to move forward with AI development without the fear of facing costly copyright infringement lawsuits related to training data.

WHAT IS A FOUNDATION MODEL?

Foundation models are large-scale AI models trained on vast amounts of diverse **data**, making them adaptable to a variety of tasks. These models serve as the backbone for many AI applications, providing a base level of knowledge that other developers, like startups, can refine for specific uses. You can think of foundation models as a handyman, individuals that are generally useful for a wide range of home improvement tasks, but with a bit more specialized training can become very good at specific tasks, like plumbing or electrical.

Rather than developing an AI system from the ground up, companies can leverage foundation models to accelerate innovation, reducing both costs and complexity. These models can be closed or **open-source** — for example, OpenAI's GPT-4 (closed) and Meta's LLaMA (open) are both foundation models.[27] Through techniques like **fine-tuning**,[28] foundation models can be adapted for specific applications, like reviewing pitch decks,[29] improving public services,[30] or enhancing the capacity of small businesses.[31]

Foundation models streamline AI development by providing a starting point, saving startups from the immense resource burden of training a model from scratch. Training a foundation model requires vast **computational** resources, often amounting to millions of dollars in cloud costs and specialized hardware. Most startups do not have access to those sorts of resources. Even if they did aim to build their own foundation model to then fine-tune, it would likely be out of date by the time they finished training, based upon the rapid increases in performance in models from leading AI labs.[32] By using a foundation model (or multiple), developers can focus on optimizing performance for their unique needs, significantly lowering barriers for startups to bring AI-based products to market.[33]

Policymakers' approach to AI regulation — and to foundation models in particular — will impact far beyond just the foundation model developers themselves, given the ecosystem of startups and others building upon them.[34] For example, AI rules that incorrectly assign liability to developers of foundation models rather than malign actors (who misuse models to break the law or commit antisocial behaviors like creating misinformation) will undermine the availability of those models. Developers will not want to be liable for actions of others that they do not have control over and will restrict access to their models, potentially harming startups building upon them. This disincentive will be particularly acute for open-source models because open source developers lack formal relationships with and awareness of those who use and build with the technology they make widely available.[35]

WHAT ARE TRANSFORMERS?

Humans don't make very good multitaskers, but the same isn't true for AI. Recent innovations in AI have enabled AI models to quickly and efficiently analyze massive amounts of data in parallel, meaning they process different pieces of inputs at the same time, rather than step-by-step like older models. These AI models are called **transformers**.

Transformers are far faster than traditional models, enabling AI systems to understand the relationships between different pieces of data — whether that's words in a sentence, pixels in an image, or even chunks of code. Originally designed for tasks like language translation and text generation, transformers have since expanded their reach to other fields like computer vision and even code generation.

Most leading AI models today are transformers — it's the "T" in ChatGPT, for example — and startups are leveraging transformers to deploy AI solutions that scale, work faster, and use fewer resources — key advantages when budgets and timeframes are tight.

Transformers' efficiency lies in their design. They use a **self-attention** mechanism to focus on different pieces of information within the data to understand relationships among them; **positional encoding** to keep track of the order of data; and an **encoder-decoder** structure, where one part of the model processes input data (the encoder) and the other generates outputs (the decoder).

If we were translating "Humpty Dumpty sat on a wall" into Spanish, those elements of the transformer model would each play a role. As data was being encoded, **self-attention** would capture the relationship between "sat" and "wall." Even though that entire phrase would be processed at the same time — that is, in parallel — **positional encoding** would keep track of the order of the words. Finally, the **decoder** would generate the output: "Humpty Dumpty se sentó en una muralla."

Transformers are adaptable to a wide range of tasks, and a similar process would follow if we were processing an image of Humpty Dumpty, except with pixels instead of words. The ability to use one model for tasks across multiple modes — like text, images, or other media — can help reduce time and development costs compared to older models created for each task.

WHAT IS OPEN SOURCE?

Open-source software is in just about every tech product in existence, from your phone, to your car, to your refrigerator. It has led to orders-of-magnitude reductions in costs to start a company, helped improve security, and fostered innovation.[36] In many policy conversations about AI, the term "open-source" often comes up, with similar implications for startups, but there aren't yet the same conventions about what exactly open-source AI entails.

Open source has historically described software that "anyone can view, modify, and distribute."[37] In AI, this generally means making the algorithms, code, weights, and data used for AI development available to the public. This approach facilitates collaboration and innovation among researchers, developers, and startups.

But openness in AI generally isn't binary — fully closed or fully open. Instead, some open-source models embody a philosophy of full transparency by releasing all aspects of the model, while other developers retain some resources as proprietary information, only releasing a combination of the pre-trained **model weights**, code, or datasets. Further, unlike open-source software, which has a well-developed set of licensing norms,[38] open-source AI is less standardized. Some open-source AI resources come with restrictive licenses that prevent their use for, e.g., commercial use or distribution (which make them comparatively less useful for startups).

Open-source models are crucial for startups because they reduce the need to develop AI models from scratch. Training AI models requires vast amounts of data, storage, and computational capabilities, which startups typically do not have. Having access to pretrained models and their weights enables startups to build on and adapt models to their needs through a process called **fine-tuning**. This accessibility accelerates innovation by drastically reducing costs and reducing barriers to entry for startups.[39] Additionally, it enables startups to build better products by enabling them to focus their limited resources on their true innovation rather than foundational technology.

Policymakers are currently evaluating the benefits of promoting open weights to foster innovation, and it's critical that they strike a careful balance in their regulatory approach to avoid imposing excessive burdens on startups or undercutting a key path for them to innovate in AI. The Federal Trade Commission even recently weighed in to underscore the benefits for startup competitiveness of open-weight models.[40] But openness can get caught in the crosshairs of overzealous regulation. For example, proposals like California's recently-vetoed SB 1047 are aimed at regulating AI model development but would make model developers responsible for future (mis)uses of the model.[41]

Under that arrangement, no developer would open source their AI models to others, because they could be held accountable for actions they had no control over.[42]

Concerns about potential misuse merit attention, such as malicious actors exploiting open models to develop harmful technologies, but these risks are inherent to AI—and all technology—broadly and not specific to open-source systems.[43] In fact, there are ways that open source can actually enhance AI safety by allowing a bigger and more diverse group of developers to use the technology and identify problems. In contrast, "closed" models place the burden of identifying vulnerabilities and biases solely on the original developers, potentially increasing the risks of adversarial attacks or unethical practices.

Open-source AI — or AI systems developed using open-source resources — should not be subjected to special or additional rules simply because they're open-source. The inherent risks of AI systems are similar, whether they are "open" or "closed," and regulations should be specifically tailored to their use. Supporting open-source AI not only means supporting startups but also contributes to the broader advancement of AI technologies.

> *Without open-source models and tools when we started, we would have been years behind on deployment. [...] The real innovation comes from how these models are fine-tuned and applied to specific industries. [...] These tools are crucial for startups, enabling us to adapt, innovate, and bring solutions to market faster.* [44]
>
> **- Lauren McCullough, CEO and Co-Founder, Tromml, Durham, North Carolina**
> *Tromml is an AI-driven platform designed to deliver tailored insights for the auto parts industry.*

WHAT IS FINE-TUNING?

Fine-tuning is the process of adapting a pre-trained model to perform specific tasks. Rather than training a model from scratch, fine-tuning leverages the general knowledge already acquired by the model during its initial training, specializing it for a developer's specific needs. For example, OpenAI's GPT-4 is a generalized large language model capable of performing a variety of tasks, ranging from writing essays to planning parties. Fine-tuning involves taking a model like GPT-4 and adapting it for a particular use, such as creating e-commerce product descriptions.

The process of fine-tuning is similar to training an AI model; however, instead of starting with random weights, fine-tuning uses the weights of a model that has already learned general patterns from a large and diverse dataset, often available as open-source or open-weight models. The process begins by gathering a dataset specific to a new task. To continue the metaphor from the earlier section on foundation models, this would be like sharing plumbing handbooks and YouTube videos to prepare your handyman to repair a bathroom drain.

As a startup-specific example, fine-tuning a model to create e-commerce product descriptions would involve collecting numerous examples of product descriptions, sales statistics, and order histories. This dataset is used to pass examples through the model, which adjusts its internal weights to learn patterns unique to this task. As the model learns from the new data, it specializes its knowledge, becoming especially proficient in the task at hand.

Startups fine-tuning existing models to circumvent the high costs and complexities of building an AI model from scratch. This approach allows them to sidestep the initial training phase, which consumes vast amounts of data and compute resources, focusing instead on innovating with existing models. Fine-tuning not only cuts costs and saves resources but also can reduce risks related to bias and user data privacy, as it leverages models trained on diverse, high-quality data. Overall, fine-tuning democratizes AI development, enabling startups to build innovative AI solutions that compete effectively with larger incumbents and address unique service gaps.

Fine-tuning, which relies on access to pre-trained models and model weights, is deeply intertwined with **open-source** AI. To foster innovation powered by fine-tuning, policymakers should also support open-source initiatives. As an additional benefit, open-source and open-weight models can actually make AI safer by attracting testing and input from researchers and developers worldwide.

POLICY LANDSCAPE:
HOW ARE POLICYMAKERS APPROACHING AI?

Policymakers in state legislatures, Congress, and around the world have put forward a number of varying proposals touching nearly every aspect of the AI ecosystem. The number is dizzying, too—with U.S. states putting forward over 900 bills touching AI in some way through the first four months of this year alone.[45] There are several key themes to these efforts—like safety, content, and national security—that are worth breaking down in the context of their likely impact on startups.

AI SAFETY

Many AI rules in place or being contemplated by policymakers focus on potential harms that could arise from AI, whether near-term (like discrimination), or far-off (like killer robots). Nearly all of the approaches to these issues involve pre-market requirements and restrictions, like independent audits, risk assessments, certification, or licensing. Such frameworks can impose steep costs and pose clear problems for startups.

Many AI frameworks categorically impose obligations for developers, deployers, or end users. For example, the European Union's AI Act,[46] Colorado's AI law,[47] many proposed state laws, and rulemaking in California are each primarily designed to regulate "high-risk" AI used for "consequential decisions."[48] These categories include AI impacting education, employment, finance, lending, housing, healthcare, and more.

Assigning obligations by category can make sense in the abstract, but often becomes overinclusive or even duplicative of existing law. For example, an AI tool that helps schedule a job interview and an AI tool making a hiring decision are subject to the same obligations. And an adverse outcome—like racial discrimination—is already subject to scrutiny under existing law and would be violative of civil rights laws.[49] Moreover, the benchmark of comparison matters. Humans have biases and informed use of technology can actually help to mitigate them and improve outcomes.

A smaller, yet very impactful set of proposals is focused on "existential" risks potentially arising from the most advanced "frontier" AI models. An example of such a framework that gained much attention in 2024 was California's SB 1047.[50] That bill would have regulated model development by requiring a safety determination at the outset and then held developers liable for meeting it—disincenting model development and the availability of open-source AI models. Even though the bill was aimed at only the most resource-heavy models, it would have rippled through the startup ecosystem since those are the foundation models startups are leveraging to innovate in AI.

> *...a lot of amazing solutions that society is coming to rely on that wouldn't be possible without AI. It's important for policymakers to be careful when thinking about how to handle technologies that are still evolving.*[50]
>
> **- Laura Truncellito, Founder, Enployable, Tysons, Virginia**
> *Enployable is AI-powered platform designed to unlock hidden talent and address labor shortages in the construction, energy, transportation, and tech sectors through two-way matching between employers' missions and cultures and job candidates' values, beliefs, and soft skills.*

CONTENT

Policymakers are concerned with both inputs—content used to train AI models—and outputs—content generated by AI models. Both will impact how AI models are developed, who can use them, and for what end uses. On inputs, there is a fierce debate over the inclusion of copyrighted content in AI training data. Large rightsholder organizations, authors, artists, and others have sued alleging that the inclusion constitutes infringement. Policymakers have generally waited to act on this front, not wanting to get ahead of the courts. However, they are advancing requirements for developers to disclose training data—including the enumeration of copyrighted or potentially copyrightable materials.[52] Outside of the clear trade secret and competitiveness implications for developers, such requirements would add significant costs to model training while exposing startups to potentially ruinous litigation.

On outputs, policymakers have put forward a number of proposals that would create potential liability for AI-generated content and for content-hosting platforms—both of which could include startups. One proposal in Congress and several states would create a new intellectual property right for individuals' Name, Image, and Likeness (NIL), set up a notice-and-takedown regime, and enable lawsuits against individual creators, AI companies, and the content hosting platform for creation and dissemination of unauthorized "digital replicas."[53] Related proposals create watermarking requirements for generated content, including some that require those watermarks to be immutable (which may be technically infeasible).[54] Most of the proposals are responding to AI-generated "deepfakes" or "digital replicas,"[55] but they lack sufficient safeguards for protected uses—like parody—and recognition of innovative or benign uses—like sending a personalized video message to a customer thanking them for making a purchase.[56]

One area of clear harm that policymakers want to address is the distribution of non-consensual intimate imagery (NCII), including AI-generated NCII. One federal proposal that has momentum is the Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks (TAKE IT DOWN) Act, which would criminalize the publication of NCII, including imagery generated by AI, and require that Internet platforms remove it within 48 hours of being notified of the content.[57] This targeted bill has less potential for unintended consequences for startups, since the vast majority of Internet platforms already take steps to remove pornographic content and do not wish to host this type of content.

NATIONAL SECURITY

AI is a foundational and powerful technology that can be used for many purposes, including both military and commercial uses, meaning leadership in AI innovation is geopolitically important. Accordingly, many policymakers are focused on U.S. leadership in AI and have leveraged or explored leveraging export controls and other restrictions on key technologies. For example, earlier in 2025, the U.S. put forward a "Framework for AI Diffusion," that regulates the diffusion of chips needed for AI development and model weights for certain closed models.[58] That framework sees export controls imposed by "tiers," with most countries subject to restrictions—including neighbors, top trading partners, and North Atlantic Treaty Organization (NATO) Allies. Policymakers have also explored restrictions on open-source models to prevent their use by individuals in other countries.[59]

The U.S. should lead in AI development, but poorly calibrated export controls are likely to backfire in the long run.[60] Restricted countries will not idly accept their exclusion from access to the most advanced technologies. Non-U.S. supply chains will be developed, creating a market for technologies hailing from adversarial nations. Eventually, innovators in the U.S. and around the world will be building with those tech products. Instead, maintaining global sales of leading U.S. AI technologies is essential to securing U.S. AI leadership, promoting cycles of investment in the AI ecosystem, and attracting the top talent to build in America.

RESOURCES

Innovation in AI is a resource-heavy endeavor with needs for expensive data, compute, and technical expertise. Adoption of AI can similarly be a barrier for government and conventional small businesses alike. Policymakers both in the states and at the federal level have conceived efforts to address these issues. For example, both federal agencies and state governments have set up programs to provide compute and AI-ready data sets to students, researchers, and startups.[61] Federal and state governments have made steps toward government adoption, and Congress has advanced bills to provide guidance for AI adoption to conventional small businesses.[62] Federal agencies have also developed voluntary guidance and best practices for AI development.[63] Taken together, these efforts can improve the competitiveness of startups—directly, by growing the talent pool, and by helping to create a market for their AI products.

THE PRO-STARTUP POLICY AGENDA TO POWER AI INNOVATION

Policymakers need to support pro-startup policy if they want a world-leading AI ecosystem made up of U.S. companies building AI, building with AI, and using AI to better everyday tasks. AI is a foundational technology that is implicated in a range of policy issues impacting the competitiveness of startups and the speed of innovation. Policy to power AI innovation, then, is not accomplished with one bill or one framework, but instead requires smart policies across all issues impacting AI developers, deployers, and users.

Startups are driving innovation in AI benefitting every corner of the economy—from agriculture to manufacturing, healthcare to education, finance to retail, and more. But policymakers' approach to regulating AI will determine who is able to participate in the AI ecosystem and the speed at which innovations are disseminated to benefit the public. The U.S. has generally followed a model of permissionless innovation,[64] enabling entrepreneurs to build beneficial new products unencumbered by strict, expensive regulatory regimes antithetical to invention, experimentation, and iteration. This approach has made the U.S. tech sector the envy of the world. But poorly conceived AI regulatory frameworks, that are overbroad and over-reliant on *ex-ante* approaches threaten to undermine U.S. startup competitiveness and innovative capacity.

> *I come from a regulatory background, and I know regulation in the AI space is coming. But regulation needs to be done with startups and entrepreneurship in mind, rather than big corporations. If not, there will be a pause in AI growth and development following the regulation implementation, as small businesses will have to jump through more hurdles to get their ideas off the ground. There must be space for R&D and innovation within the regulation to prevent the stalling of AI growth, as well as resources to help startups with the increased expenses associated with compliance.*[65]
>
> **- Darryl Keeton, Founder and President, Sensagrate, Scottsdale, Arizona**
> *Sensagrate is a software platform that leverages cameras, LiDAR, radar, and other sensors to monitor traffic and improve safety.*

To address potential harms without burdening innovation, any regulation that policymakers pursue should be outcome-focused, begin from a position of existing law, and target bad actors. Many potential harms associated with AI are already illegal,[66] and enforcing or enhancing those protections is the most straightforward way to address those concerns without new, duplicative, and burdensome regulators and frameworks.

AI also doesn't stop at state or even national borders. Policymakers must act to avoid a state patchwork of varying or conflicting AI rules and take steps to avoid a global patchwork of incongruous, competing approaches. States and localities can (or even should!) take steps to encourage AI investment, research, and development, but must avoid enacting their own unique AI rules. A patchwork of varying rules will burden startups, slow down innovation, and undermine U.S. AI leadership.

Winning the AI race can be accomplished by pursuing a policy agenda that unleashes innovation, encourages investment, creates clear rules, and opens markets, guided by these realities:

STARTUPS NEED CLEAR RULES TO BE COMPETITIVE

Policymakers need to be sure their approach to regulating AI is tailored to discrete harms and is not duplicative of existing law to avoid harming startup competitiveness. For example, employment discrimination is already unlawful—whether that discrimination involves the use of AI or not—meaning a new law specific to hiring and AI is likely to be duplicative.[67] Thanks to existing law, firms offering those services already have market and legal

incentives to ensure their products function properly—the new rules will only add cost and burden innovation without delivering justifiable benefit. Moreover, if each state has its own AI rules, the overlapping obligations will create additional costs and further undermine startup competitiveness.[68] Maintaining one set of rules—achieved through federal preemption, if necessary—is imperative to creating the clarity needed to enable startup success.

AI: Startups want to build socially-beneficial AI tools and routinely look to standard-setting organizations and industry best practices for guidance. To bolster responsible AI innovation and enhance U.S. global influence, policymakers should support business-led development of voluntary standards. Leveraging safe harbors can help further incent adoption of best practices. Policymakers can also utilize regulatory sandboxes, which foster innovation while enabling startups and regulators alike to learn from each other and find a balanced approach. Sandbox programs must be set up so that it is beneficial for startups to participate, and they require sufficient time, staffing, and resources to properly function.

Capital: Policymakers should avoid forcing startups to expend their few resources following new, overbroad rules—especially where existing law already covers the motivating concern. Overbroad, imprecise definitions in regulatory frameworks can scope-in far too many activities, creating compliance obligations that strain startups' limited budgets without cognizable benefit. For example, definitions of AI can scope-in common technologies like calculators or spreadsheets. Categorically defining risk creates the same obligations for AI models that diagnose diseases as those that help schedule appointments.

Capital: Poorly conceived regulatory frameworks foist disproportionate compliance burdens upon startups that drains their limited capital, discourages innovation, and undermines economic dynamism. The obligations created by rules themselves, like audit requirements, can be prohibitively expensive. Since development costs and barriers to market steer where startups innovate, these provisions could be net-negative if they discourage socially beneficial innovations.

Trade: To support U.S. leadership in AI, it is essential to avoid a patchwork of disjointed AI regulations at home and to leverage smart digital trade policies to avoid a global patchwork. Engaging with trading partners and in multilateral fora can help smooth global regulatory approaches, and supporting longstanding U.S. digital trade priorities—like enabling cross-border data flows, opposing data localization, protecting source code, and prohibiting tariffs on digital trade—will help break down barriers to foreign markets, especially for U.S. startups.[69]

> *We need more clear, accessible guidelines and resources that would help us innovate responsibly without being overwhelmed by the compliance costs or confusion.*[70]
> **- Remy Meraz, Co-Founder and CEO, Zella Life, Los Angeles, California**
> *Zella Life offers an AI-powered executive coaching platform*

STARTUPS BUILD OFF OF FOUNDATION MODELS

Startups are building and leveraging AI in a few distinct ways, and to avoid stifling innovators, policy needs to account for all of them. Many startups are leveraging foundation models, often large language models (LLMs)—both open and closed source—and fine-tuning them to create unique services.[71] This means AI policy needs to support open-source development and avoid policies that will increase costs or disincent both open and closed-source developers to make their models available for others to build upon.[72]

Liability: Foundation models are suitable for a wide range of tasks, but end users often determine how the model will be used. AI rules that incorrectly assign liability to developers of foundation models rather than malign actors will restrict the availability of those models, because developers

will not want to be liable for actions of others that they do not have control over. This disincentive will be particularly acute for open-source models, because open source developers lack formal relationships with and awareness of those who use and build with the technology they make widely available.[73]

> *In the AI space, everything's moving so fast. … We don't have the funding to compete on the foundational model level and we don't necessarily want to. What we can compete on is understanding the value proposition of our customers the best. We use open-source models for the broad LLM needs we have. Then, we build a couple of custom models on our end to add to the foundation model.*[74]
>
> **- Paul Ehlinger, Co-Founder, Flamel.ai, Covington, Kentucky**
>
> *Flamel.ai is an AI platform that helps multi-location and franchise brands stay true to their brand guide while delivering content at scale.*

STARTUPS NEED DATA TO BUILD AI

AI innovation is data-driven, and startups need data to build, train, and fine-tune AI models. Data acquisition can be expensive, and startups have few resources to acquire vast data sets needed to build accurate and useful models.[75] Policy must enable startups to access the data they need to build, test, and improve AI models.[76]

IP: Policy should recognize that training models with data sets that include copyrighted content is permissible under law. This is imperative to support innovation, deter costly litigation, avoid gatekeeping by large entities, and prevent prohibitively expensive licensing requirements.[77]

Privacy: Startups need one consistent nationwide framework that creates clarity, streamlines costs, and fosters data-driven innovation. At present, a patchwork of data privacy laws risks uncertainty around data use, creates duplicate requirements, and weighs on already-strapped startup budgets.[78] A federal privacy law should ensure that startups can collect and process the data necessary to create new and beneficial products.

Government: Government possesses troves of data useful for AI, especially for tailored, specialized models.[79] Policymakers should ensure that agencies make this data available in AI-ready formats, and where possible, without availability lags that can undermine its utility.

> *AI is based on statistics, and to have robust statistics, you need data. It's a numbers game, and for startups to be able to compete, they need to not be restricted from accessing this data.*[80]
>
> **- Dr. Carlos Gaitan, Co-Founder & CEO, Benchmark Labs, San Diego, California**
>
> *Benchmark Labs uses AI to provide farmers with actionable weather data relevant to their location to help them save water, energy, pesticides, and fertilizers.*

STARTUPS NEED RESOURCES TO INNOVATE

AI innovation can be technical and expensive, and policy should aim to facilitate investment in AI startups, strengthen skilled talent pools, and make available resources directly to startups.

Capital: Policy should enable and incent investment in AI startups. Measures to grow the pool of investors are essential to improve capital access for AI innovators—especially those located outside existing major hubs. Further, pursuit of adjacent policy goals—especially in the competition space—without fully considering downstream consequences will make it harder for startups to get

investment or leverage existing investments. Policymakers must avoid restricting investment in AI startups and should enable—not limit—successful startup exits.[81]

Tax: Tax policies should prompt investment in startups and in AI R&D. Preserving and expanding favorable tax treatment of qualified small business stock and introducing federal credits for angel investments will help to de-risk investment in AI startups, aid early-stage AI startups in attracting needed talent, and encourage positive cycles of reinvestment of returns. Restoring immediate expensing of R&D costs will increase startups' capacity by letting them put more resources toward AI innovation.

Talent: To grow the AI talent pool, policymakers should fund and support AI skilling and upskilling programs, enhance STEM education, and attract and retain talented immigrants. Skilling, upskilling, and STEM education efforts are essential long-term efforts to building a strong base of domestic talent prepared to work in AI companies and with AI tools. Reforming and expanding immigration pathways is essential to winning the talent race. Streamlining the O-1 visa program and creating a startup visa will ensure founders can start and grow AI companies in the U.S. Expanding the H-1B visa program will help to fill present shortages in skilled talent. Finally, many foreign students receive advanced STEM degrees from U.S. institutions but then return home where they end up competing with U.S. innovators. Instead, those graduates should have the option to remain in the U.S. with the stability—like permanent residency—needed to launch or work at a startup.

Capital: Government grants are important means of capital access that support American R&D, facilitate commercialization of new innovations, and de-risk investment in recipient startups. Flagship programs like Small Business Innovation Research grants should be made permanent, accessible to more startups, and improved to suit startup realities.[82] Further, key resources like compute and data can be very expensive and steer how startups innovate. Providing compute or data sets directly to startups can lower barriers to entry and bolster AI R&D.[83]

> *Accessing the right talent on our budget was one of our most challenging problems as a startup. [...] Being a startup on a shoestring budget, we went through a couple of student machine learning engineers, and then we got money for more professional ones. Still, some of them flopped as well. It took a lot of iterations to find a machine learning engineer.*[84]
>
> **- Dr. Cara Wells, Founder, EMGenisys, Driftwood, Texas**
>
> *EMGenisys is a machine learning-powered platform to analyze embryos and improve animal reproduction.*

STARTUPS CAN HELP TO IMPROVE GOVERNMENT

Policymakers can use levers of government to speed AI adoption and support U.S. AI leadership around the globe.

Government: Government is a large buyer of software and serves many functions that could be made more efficient through the use of AI. Policymakers should ensure acquisition processes work for startups so they can help to improve the provision of government services and through AI.

> *We are excited to be in a space that gets to serve government and public officials, especially local governments across the country. Oftentimes these are folks that work extremely hard, are incredibly underpaid, and choose careers that allow them to serve their communities with their skill set. The notion that we can provide tools for them to improve how they do their jobs, and serve their community more, is awesome. Not only is it a novel use of AI, but an incredibly important one.*[85]
>
> **- Chip Kennedy, CEO, and Lindsay Avagliano, COO, CivicReach.AI, Raleigh, North Carolina**
>
> *CivicReach is building local government-specific, on-demand, and employee-first voice AI agents to assist staff and call centers in answering the phone more often, more times of day, and in more languages.*

ENDNOTES

1 *#StartupsEverywhere: Constanza Gomez, Co-Founder, Sortile,* Engine (Sept. 13, 2024), https://www.engine.is/news/startupseverywhere-new-york-ny-sortile.
2 *#StartupsEverywhere: Vance Reavie, CEO and Founder, Junction AI,* Engine (Oct. 17, 2024), https://www.engine.is/news/startupseverywhere-bentonville-ark-junction-ai.
3 National Defense Authorization Act for Fiscal Year 2019, Pub. L. No. 115-232, § 238(g) (2018), https://www.congress.gov/115/bills/hr5515/BILLS-115hr5515enr.pdf.
4 *Infra, What role does data play and where does it come from?*
5 *See, e.g.,* Nick Ismail, *The success of artificial intelligence depends on data,* Information Age (Apr. 23, 2018), https://www.information-age.com/success-artificial-intelligence-data-10142/.
6 *See, e.g.,* Min Jun Jung and Nathan Lindfors, *Startups and AI policy: how to mitigate risks, seize opportunities and promote innovation,* Engine (Sept. 8, 2023), https://www.engine.is/news/category/startups-and-ai-policy-how-to-mitigate-risks-seize-opportunities-and-promote-innovation.
7 Tom Brown, et al., *Language Models are Few-Shot Learners,* (July 22, 2020), https://arxiv.org/abs/2005.14165.
8 *See, e.g., Coalition letter to Gina Raimondo, Sec'y of Com.,* (Mar. 25, 2024), https://www.rstreet.org/wp-content/uploads/2024/03/Civil-Society-Letter-on-Openness-for-NTIA-Process-March-25-2024.pdf.
9 *See, e.g.,* Nathan Lindfors, *What's in the AI executive order, and what does it mean for startups?,* Engine (Nov. 3, 2023), https://engineadvocacyfoundation.medium.com/whats-in-the-ai-executive-order-and-what-does-it-mean-for-startups-3ae23059d518.
10 *See, e.g., Dual-Use Foundation Models with Widely Available Model Weights,* Nat'l Telecomm. & Info. Admin (July 2024), https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.
11 *IEEE Standard for Floating-Point Arithmetic,* IEEE Std 754-2019 (July 22, 2019), https://standards.ieee.org/ieee/754/6210/.
12 *See, e.g., Tracking Large-scale AI models,* EpochAI, https://epochai.org/blog/tracking-large-scale-ai-models#appendix.
13 Exec. Order No. 14110, 88 Fed. Reg. 75197-8 (2023), https://www.govinfo.gov/content/pkg/FR-2023-11-01/pdf/2023-24283.pdf.
14 *See, e.g., Letter from Engine to Gavin Newsom, Governor of Cal.,* Engine (Sept. 3, 2024), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/66d9f473ab2cb563476cb8e0/1725559924043/Engine+SB+1047+veto+letter.pdf.
15 *See, e.g.,* CREATE AI Act of 2025, H.R. 2385, 119th Cong. (2025), https://www.congress.gov/bill/119th-congress/house-bill/2385/.
16 Crystal Grant, et al., AI (and other) Companies: Quietly Changing Your Terms of Service Could Be Unfair or Deceptive, F.T.C. (Feb. 13, 2024), https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/02/ai-other-companies-quietly-changing-your-terms-service-could-be-unfair-or-deceptive.
17 Common Crawl, https://commoncrawl.org/.
18 Large-scale Artificial Intelligence Open Network, https://laion.ai/.
19 *See, e.g.,* Jill Crosby, *IP recap: Generative AI lawsuits and what they mean for startups,* Engine (Aug. 15, 2024), https://engineadvocacyfoundation.medium.com/ip-recap-generative-ai-lawsuits-and-what-they-mean-for-startups-7f399702a151.
20 *See, e.g.,* Cecily Mauran, *All the media companies that have licensing deals with OpenAI (so far),* Mashable (June 21, 2024), https://mashable.com/article/all-the-media-companies-that-have-licensing-deals-with-openai-so-far.
21 Ilia Shumailov, et al., *AI models collapse when trained on recursively generated data,* Nature (July 24, 2024), https://www.nature.com/articles/s41586-024-07566-y.
22 *See, e.g.,* Jill Crosby, *IP recap: Generative AI lawsuits and what they mean for startups,* Engine (Aug. 15, 2024), https://engineadvocacyfoundation.medium.com/ip-recap-generative-ai-lawsuits-and-what-they-mean-for-startups-7f399702a151.
23 Harry Styles (@Harry_Styles), BRUSHING TEETH MATE., Twitter (June 15, 2013, 12:27 AM), https://x.com/Harry_Styles/status/345699018594201600.
24 Comments of Engine to the U.S. Copyright Office's Notice of Inquiry on Artificial Intelligence and Copyright, Docket No. 2023-6, Engine (Oct. 23, 2023), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/6541224062cc713bb5eea9d7/1698767424782/Engine+comments+to+CO+on+AI+NOI+10.23.pdf.
25 *See, e.g.,* Reply Comments of Engine to USCO's Notice of Inquiry on Artificial Intelligence and Copyright, Docket No. 2023-6 concerning FTC's Comments dated October 30, 2023, Engine (Dec. 6, 2023), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/6570d59997a3a32c3dc05e0b/1701893529701/Engine+reply+comments+to+CO+on+AI+NOI+12.23.pdf.
26 https://www.engine.is/news/startupseverywhere-new-york
27 *Infra, What is open source?.*
28 *Infra, What is fine-tuning?.*
29 *See, e.g.,* CapwaveAI, https://www.capwave.ai/.
30 *See, e.g.,* CivicReach, https://www.civicreach.ai/.
31 *#StartupsEverywhere: Vance Reavie, CEO and Founder, Junction AI,* Engine (Oct. 17, 2024), https://www.engine.is/news/startupseverywhere-bentonville-ark-junction-ai.
32 *See, e.g., LLM Leaderboard,* Artificial Analysis, https://artificialanalysis.ai/leaderboards/models.
33 *See related, Tools to Compete: Lower Costs, More Resources, and the Symbiosis of the Tech Ecosystem,* CCIA Research Center and Engine (Jan. 2023), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/63d2b8d5bec96f502264fd1f/1674754266044/FINAL_CCIA-Engine_Tools-To-Compete.pdf.
34 *See, e.g., Infra, Policy Agenda.*
35 *See, e.g., Letter from Engine to Gavin Newsom, Governor of Cal.,* Engine (Sept. 3, 2024), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/66d9f473ab2cb563476cb8e0/1725559924043/Engine+SB+1047+veto+letter.pdf.
36 *Tools to Compete: Lower Costs, More Resources, and the Symbiosis of the Tech Ecosystem,* CCIA Research Center and Engine (Jan. 2023), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/63d2b8d5bec96f502264fd1f/1674754266044/FINAL_CCIA-Engine_Tools-To-Compete.pdf.
37 *See Open Source Initiative,* https://opensource.org/
38 *Id.* Licenses, https://opensource.org/licenses.
39 *Comments of Engine Advocacy regarding Openness in AI Request for Comment, NTIA–2023–0009,* Engine (Mar. 27, 2024), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/66048124bf1b864932c664bd/1711571237337/NTIA+AI+openness+March+27,+2024.pdf.
40 Alex Gaynor, et al., *On Open-Weights Foundation Models,* F.T.C. (July 10, 2024), https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/07/open-weights-foundation-models.
41 Safe and Secure Innovation for Frontier Artificial Intelligence Models Act, S.B. 1047, 2023–2024 Leg., Reg. Sess. (Cal. 2024) (vetoed Sept. 29, 2024), https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB1047.
42 *See, e.g.,* Letter from Engine to Cal. Sen. Scott Wiener, (Aug. 6, 2024), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/66b2788607345f3950ab57d9/1722972294380/Engine+letter+Opposing+SB+1047.pdf.
43 *See, e.g., Coalition letter to Gina Raimondo, Sec'y of Com.,* (Mar. 25, 2024), https://www.rstreet.org/wp-content/uploads/2024/03/Civil-Society-Letter-on-Openness-for-NTIA-Process-March-25-2024.pdf.
44 *#StartupsEverywhere: Lauren McCullough, CEO and Co-Founder, Tromml,* Engine (Feb. 14, 2024), https://www.engine.is/news/startupseverywhere-durham-nc-tromml.

45 *Artificial Intelligence (AI) Legislation,* MultiState, https://www.multistate.ai/artificial-intelligence-ai-legislation.
46 "EU AI Act," Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on Artificial Intelligence, (2024), https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.
47 Consumer Protections for Artificial Intelligence, S.B. 24-205, 74th Gen. Assemb., Reg. Sess. (Colo. 2024) (enacted), https://leg.colorado.gov/sites/default/files/2024a_205_signed.pdf.
48 *See related, Comments of Engine Advocacy on CCPA Updates, Cyber, Risk, ADMT, and Insurance Regulations*, Engine (Jan. 14, 2025), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/67869fcfec4abe3e981bb785/1736875984100/Engine+CPPA+ADMT+Comments.pdf.
49 42 U.S.C. § 2000e-2.
50 Safe and Secure Innovation for Frontier Artificial Intelligence Models Act, S.B. 1047, 2023–2024 Leg., Reg. Sess. (Cal. 2024) (vetoed Sept. 29, 2024), https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB1047.
51 *#StartupsEverywhere Profile: Laura Truncellito, Founder, Enployable,* Engine (Apr. 22, 2022), https://www.engine.is/news/startupseverywhere-tysons-va-enployable.
52 *See related,* Letter from Engine et al. to Cal. Cmte. on Jud., (Apr. 1, 2025), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/67ed3e667d0ab649e04d0972/1743601254984/ReCreate+CA+AB+412+Letter.pdf.
53 *See, e.g.,* NO FAKES Act of 2024, S. 4875, 118th Cong. (2024), https://www.congress.gov/bill/118th-congress/senate-bill/4875/text.
54 *See, e.g.,* COPIED ACT of 2024, S. 4674, 118th Cong. (2024), https://www.congress.gov/bill/118th-congress/senate-bill/4674/text.
55 Jill Crosby, *The push for a new federal law regulating digital replicas and what it means for startups,* Engine (Aug. 22, 2024), https://engineadvocacyfoundation.medium.com/the-push-for-a-new-federal-law-regulating-digital-replicas-and-what-it-means-for-startups-ff4b1e9d2557.
56 #StartupsEverywhere Profile: Debarshi Chaudhuri, Co-Founder, Maverick, Engine (Sept. 6, 2024), https://www.engine.is/news/startupseverywhere-san-francisco-maverick.
57 TAKE IT DOWN Act, S. 146, 119th Cong. (2025), https://www.congress.gov/bill/119th-congress/senate-bill/146.
58 *Framework for Artificial Intelligence Diffusion,* 90 Fed. Reg. 4544-84 (Jan. 15, 2025), https://www.govinfo.gov/content/pkg/FR-2025-01-15/pdf/2025-00636.pdf.
59 *See, e.g.,Dual-Use Foundation Models with Widely Available Model Weights,* Nat'l Telecomm. & Info. Admin (July 2024), https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.
60 *See, e.g.,* John Villasenor, *The New AI Diffusion Export Control Rule Will Undermine U.S. AI Leadership,* Brookings Inst. (Jan. 23, 2025), https://www.brookings.edu/articles/the-new-ai-diffusion-export-control-rule-will-undermine-us-ai-leadership/.
61 *See, e.g.,* National AI Research Resource Pilot, https://nairrpilot.org; Empire AI, https://www.empireai.tech/.
62 *See, e.g.,* Small Business Artificial Intelligence Training Act of 2024, S. 4487, 118th Cong. (2024), https://www.congress.gov/bill/118th-congress/senate-bill/4487/text.
63 *See, e.g.,* AI Risk Management Framework, https://www.nist.gov/itl/ai-risk-management-framework.
64 *See, e.g.,* Adam Thierer, *Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom,* Mercatus (Mar. 2016), https://www.mercatus.org/research/books/permissionless-innovation-continuing-case-comprehensive-technological-freedom.
65 *#StartupsEverywhere: Darryl Keeton, Founder and President, Sensagrate,* Engine (Sept. 19, 2024), https://www.engine.is/news/startupseverywhere-scottsdale-ariz-sensagrate.
66 *See, e.g.,* Krista Chavez, The U.S. Does Not Need Bigger Government to Deal With Concerns About AI Tools, NetChoice (Aug. 10, 2023), https://netchoice.org/the-u-s-does-not-need-bigger-government-to-deal-with-concerns-about-ai-tools/.
67 *See, e.g., Statutes,* U.S. Equal Employment Opportunity Commission, https://www.eeoc.gov/statutes/title-vii-civil-rights-act-1964.
68 *See related, Privacy Patchwork Problem: Costs, Burdens, and Barriers Encountered by Startups,* Engine (Mar. 2023), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/6414a45f5001941e519492ff/1679074400513/Privacy+Patchwork+Problem+Report.pdf.
69 *See Letter to U.S. Trade Policymakers* (Feb. 7, 2024), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/65c3906e36cbbb45ba281205/1707315310372/Startup+Digital+Trade+Open+Letter.pdf.
70 *#StartupsEverywhere: Remy Meraz, Co-Founder and CEO, Zella Life,* Engine (Oct. 31, 2024), https://www.engine.is/news/startupseverywhere-los-angeles-calif-zella-life.
71 Min Jun Jung, *AI Essentials: What is fine-tuning?,* Engine (Oct. 17, 2024), https://engineadvocacyfoundation.medium.com/ai-essentials-what-is-fine-tuning-16cf31e40b3d.
72 Min Jun Jung, *AI Essentials: What is open-source?,* Engine (Oct. 3, 2024), https://engineadvocacyfoundation.medium.com/ai-essentials-what-is-open-source-fe53d0395d21.
73 *See, e.g., Letter to The Honorable Gavin Newsom,* Engine (Sept. 4, 2024), https://engine.is/s/Engine-SB-1047-veto-letter.pdf.
74 *#StartupsEverywhere: Paul Ehlinger, Co-Founder, Flamel.ai,* Engine (Feb.7, 2025), https://www.engine.is/news/startupseverywhere-covington-ky-flamel-ai.
75 *the State of the Startup Ecosystem,* Engine (Apr. 2021), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/60819983b7f8be1a2a99972d/1619106194054/The+State+of+the+Startup+Ecosystem.pdf.
76 Kyrylo Bakumenko and Min Jun Jung, *AI Essentials: The role of data and where it comes from*, Engine (Dec. 6, 2024), https://engineadvocacyfoundation.medium.com/ai-essentials-the-role-of-data-and-where-it-comes-from-bbf1e69eaca1.
77 Jill Crosby, *Generative AI lawsuits and what they mean for startups,* Engine (Aug. 15, 2024), https://www.engine.is/news/category/generative-ai-lawsuits-and-what-they-mean-for-startups.
78 *See related, Privacy Patchwork Problem: Costs, Burdens, and Barriers Encountered by Startups,* Engine (Mar. 2023), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/6414a45f5001941e519492ff/1679074400513/Privacy+Patchwork+Problem+Report.pdf.
79 *See, e.g.,* Carlos Gaitan, AI-Ready Open Data Assets RFI, Benchmark Labs (July 2024), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/6695422d489f5e3a8bd641f7/1721057837134/Data+availability+RFI_+Benchmark+Labs.pdf.
80 *#StartupsEverywhere Profile: Carlos Gaitan, Co-Founder & CEO, Benchmark Labs,* Engine (June 28, 2024), https://www.engine.is/news/startupseverywhere-sandiego-ca-benchmark.
81 Anirban Ghoshal, *Court ban on Google AI stakes would hurt Anthropic clients, say analysts,* ComputerWorld (Feb. 17, 2025), https://www.computerworld.com/article/3826080/court-ban-on-google-ai-stakes-would-hurt-anthropic-clients-say-analysts.html; *FTC Launches Inquiry into Generative AI Investments and Partnerships,* Fed. Trade Comm. (Jan. 25, 2024), https://www.ftc.gov/news-events/news/press-releases/2024/01/ftc-launches-inquiry-generative-ai-investments-partnerships; See generally http://startupexits.engine.is.
82 Jennifer Weinhart, *The Small Business Innovation Research (SBIR) Program and what it means for startups,* Engine (Aug. 31, 2022), https://engineadvocacyfoundation.medium.com/the-small-business-innovation-research-sbir-program-and-what-it-means-for-startups-330152532779.
83 *E.g.,* The National AI Research Resource, https://nairrpilot.org/; *see related Letter for the hearing record,* Engine (Feb. 16, 2024), https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/65d7ba36040b19288b1465cb/1708636727046/Engine+Statement+re+Feb+6+SST+Hearing+.pdf.
84 *#StartupsEverywhere: Dr. Cara Wells, Founder, EMGenisys,* Engine (Feb. 21, 2025), https://www.engine.is/news/startupseverywhere-driftwood-texas.
85 *#StartupsEverywhere: Chip Kennedy, CEO, and Lindsay Avagliano, COO, CivicReach.AI,* Engine (Apr. 11, 2025), https://www.engine.is/news/startupseverywhere-raleigh-nc-civicreach-ai.