ERIN E. MURPHY (*pro hac vice* forthcoming)
MATTHEW D. ROWEN (Cal. Bar #292292)
JAMES Y. XI (*pro hac vice* forthcoming)
MITCHELL K. PALLAKI (*pro hac vice* forthcoming)
ILAN J. POSNER (*pro hac vice* forthcoming)
CLEMENT & MURPHY, PLLC
706 Duke Street
Alexandria, VA 22314
Tel.: (202) 742-8900
erin.murphy@clementmurphy.com
matthew.rowen@clementmurphy.com
james.xi@clementmurphy.com
mitchell.pallaki@clementmurphy.com
ilan.posner@clementmurphy.com

*Attorneys for Plaintiff X.AI LLC*

**UNITED STATES DISTRICT COURT**
**FOR THE CENTRAL DISTRICT OF CALIFORNIA**

| | |
|---|---|
| X.AI LLC,<br><br>*Plaintiff*,<br><br>v.<br><br>ROB BONTA, in his official capacity as Attorney General of the State of California,<br><br>*Defendant*. | Case No. _____<br><br>**COMPLAINT FOR DECLARATORY AND INJUNCTIVE RELIEF**<br><br>**ACTION SEEKING STATEWIDE RELIEF**<br><br>(42 U.S.C. §1983; *Ex parte Young*) |

Plaintiff X.AI LLC ("xAI") respectfully brings this civil action pursuant to 42 U.S.C. §1983 and the United States Constitution for declaratory and injunctive relief against Defendant Rob Bonta, in his official capacity as Attorney General for the State of California, and alleges as follows:

## INTRODUCTION

1. For as long as companies have been innovating, states have provided protection for their trade secrets—information that is valuable only if it is kept secret from others, particularly competitors. The government has long recognized that companies are unlikely to invest the considerable time and money required to develop novel products if they cannot keep secret the methods, processes, and information that make their products innovative.

2. Companies have taken advantage of trade secret protections to develop fertilizers, airplane parts, consumer beverage products, and everything in between. Trade-secrets protections have been applied broadly to all manner of industries and have allowed companies to thrive across all manner of sectors. The burgeoning artificial intelligence ("AI") industry is no different.

3. AI-focused companies like xAI develop, refine, and disseminate AI models, which are software programs that simulate human intelligence and learning to perform a wide range of tasks. A generative AI model ("gen AI" for short) is a particular type of AI model that is capable of creating original content based on user prompts. To teach these models to create accurate and responsive content, companies rely heavily on datasets—i.e., collections of information that are stored

in various formats (text, images, videos, sound files, etc.) and are derived from a variety of sources of human knowledge. Companies input datasets into their AI models so that the model not only can memorize the information it has been given, but also can extrapolate from the information it has learned. By repeating that process over and over to reinforce what the model has learned, AI companies can train their AI models to perform a wide range of tasks—everything from answering trivia questions and summarizing long documents to putting together an action plan and generating images and videos.

4. But those models are only as good as their training. And the training is only as good as the data on which the company relies. For that reason, xAI and other AI developers dedicate substantial resources to identifying high-quality data to train their models. As part of that process, they seek to find data sources that their competitors are not using to train the model to perform its tasks more efficiently. If a developer's AI model has received unique training based on sources of data that other models have not received, then it will likely have a competitive advantage, as it is trained to respond in ways that competitors' models are not.

5. Unsurprisingly, then, one of the keys to an AI developer's success is its ability to find information and sources that its competitors do not have. Accordingly, businesses like xAI make significant efforts to safeguard information about the datasets they have acquired: their sources, the amount of data they hold, the types of data included in those datasets, and their role in the overall process of developing fully functioning AI models. After all, if a competitor had insight into any of those

aspects of the datasets, it could replicate the training process and undercut its rival's competitive edge. For example, if OpenAI (another leading AI company) were to discover that xAI was using an important dataset to train its models that OpenAI was not, OpenAI would almost certainly acquire that dataset to train its own model, and vice versa. At bottom, these datasets are valuable precisely because they are not public. The datasets themselves and their usage are valuable trade secrets.

6. Until recently, California law recognized as much and protected those trade secrets from disclosure. But California Assembly Bill 2013 ("AB 2013") threatens to eviscerate that protection, as well as those afforded by Congress under the federal Defend Trade Secrets Act ("DTSA"). Although billed as a consumer-transparency statute, AB 2013 is actually a trade-secrets-destroying disclosure regime that hands competitors a roadmap to learn how companies like xAI are developing and training their proprietary AI models.

7. AB 2013 requires developers of gen AI systems or services, like xAI, to publicly disclose critical information about the datasets used to develop and train their gen AI models. Though AB 2013's provisions do not explain how detailed those disclosures must be, they cover a wide array of sensitive trade information, including, among other things, the sources of the datasets, a description of how the datasets further the gen AI's intended purpose, and the number of data points included in each dataset. If AB 2013 is allowed to take effect and the state can thereby compel developers to provide detailed and intrusive disclosures, AB 2013 would force xAI to reveal confidential information about how it develops, trains,

and refines its unique gen AI models—all of which are trade secrets that are fundamental to its business and that would otherwise be protected under state and federal law.

8. Perhaps that would be understandable—albeit still unconstitutional—if that information were particularly valuable to consumers. But it is not clear how AB 2013 is supposed to help consumers, who are far more interested in evaluating how an end-product performs the tasks it is given than in obtaining technical details about the datasets and processes companies use to train their models. The parties that are most likely to benefit from the law are thus not consumers, but competitors. While it is unclear what the consumer is even supposed to do with the information AB 2013 requires companies like xAI to disclose, xAI's rivals have both the wherewithal and the motivation to use detailed information about xAI's datasets to replicate xAI's models or improve their own AI systems, thus robbing xAI of a competitive edge in the exponentially growing AI market.

9. AB 2013's disclosure regime is unconstitutional several times over.

10. For starters, AB 2013 runs afoul of the Takings Clause. It forces xAI to give up its trade secrets on developing gen AI without any promise of compensation—a quintessential *per se* taking. Again, information about the datasets xAI uses to improve its gen AI models is a confidential trade secret protected under federal and state law. That information is economically valuable and fiercely protected. By compelling xAI to disclose those trade secrets, AB 2013 would gut their value, and force xAI to surrender a competitive advantage. Depriving xAI of

the right to exclude others from possessing its confidential trade secrets is precisely the kind of *per se* taking that the Supreme Court has reiterated requires just compensation. Because California has offered no compensation *at all* for its forced surrender of that valuable property, it cannot enforce AB 2013's disclosure requirements against xAI consistent with the Takings Clause.

11. Even if disclosure of xAI's trade secrets under AB 2013 did not rise to the level of a *per se* taking, its burdensome regulation works a regulatory taking because it will effectively reduce the value of xAI's trade secrets to zero. AB 2013 imposes a novel and unprecedented requirement on AI developers that wholly disregards the investment-backed expectations on which xAI relied when developing its products. After all, xAI had no reason to anticipate that California would disregard the critical role that trade-secrets protections play in the AI industry by compelling it to hand key details about how it develops its AI models over to its competitors. For similar reasons, by eviscerating xAI's ability to exclude others from accessing that information, AB 2013 will rob xAI of the economic value inherent in its trade secrets. In short, the extraordinary character of AB 2013's obligations confirms that it effectuates a regulatory taking. California cannot constitutionally take xAI's intangible property without compensation.

12. To make matters worse, AB 2013 also violates xAI's free speech rights under the First Amendment. Forcing a company like xAI to disseminate specific information compels speech—which implicates the First Amendment just as much as prohibiting speech. And because AB 2013 requires disclosure of particular

content, it is a content- and viewpoint-based regulation of speech that triggers strict scrutiny. That much is clear from the bill's legislative history, which does not mince words: AB 2013 is designed to "help[] identify and mitigate biases." Cal. S. Rules Comm., AB 2013, *Senate Floor Analysis* 3 (Aug. 20, 2024), *available at* https://perma.cc/7SS2-UUHW. But whatever form of heightened scrutiny ultimately applies, AB 2013 falls short. California has, at best, a minimal interest in forcing AI companies to disclose their valuable trade secrets for all the world to see. On the other hand, the burden AB 2013 imposes on xAI to disclose that information is onerous and economically devastating. The regulation is anything but narrowly tailored. It compels disclosure of information that cannot possibly be helpful to consumers, there are less restrictive alternatives available to meet California's concerns, and the law applies to all AI systems regardless of whether they pose any particular risk and even includes old AI models that are no longer in regular use.

13. On top of that, AB 2013 is also unconstitutionally vague. The law fails to provide fair notice to a person of ordinary intelligence as to what information must be disclosed. The law is unclear as to which AI systems it applies, it is internally inconsistent as to whether it applies only to datasets used during AI training or to all datasets used in development, and it offers no guidance on how much detail developers must disclose about their proprietary information. In short, AB 2013 does not come close to regulating with the "narrow specificity" that the Constitution demands of laws that impinge on free speech (and threatens

trademarks, to boot).

14. In sum, AB 2013 threatens to gut the AI industry, and it violates numerous clauses of the Constitution along the way. The court should not allow this deeply misguided and Constitution-flouting law to take effect.

## JURISDICTION AND VENUE

15. xAI's causes of action arise under 42 U.S.C. §1983 and the United States Constitution. The Court therefore has subject-matter jurisdiction over these federal claims under 28 U.S.C. §1331. This Court has authority to grant legal and equitable relief under 42 U.S.C. §1983 and *Ex parte Young*, 209 U.S. 123 (1908), injunctive relief under 28 U.S.C. §1651, and declaratory relief under 28 U.S.C. §§2201(a) and 2202.

16. Venue is appropriate in this district under 28 U.S.C. §1391(b) because the Attorney General performs his official duties in the Central District of California and is therefore considered to reside in this district as a matter of law.

## PARTIES & STANDING

17. X.AI LLC is a limited liability company that is organized under the laws of the State of Nevada and has its principal place of business in Palo Alto, California. As part of its services and operations, xAI produces and develops AI models that it intends to share broadly with members of the public. xAI accordingly constitutes a "developer" as defined by AB 2013, because it is a "corporation that designs, codes, produces, or substantially modifies an artificial intelligence system or service for use by members of the public." Cal. Civ. Code §3110(b). As a result,

when AB 2013 takes effect on January 1, 2026, xAI will need to comply with AB 2013's disclosure obligations "regarding the data" it uses "to train the generative artificial intelligence system[s] [and] service[s]" it has made publicly available since January 1, 2022, or will make publicly available again. *Id.* §3111.

18. xAI has standing to bring this action. To allege standing, a plaintiff need allege facts sufficient to establish only that his conduct is "arguably proscribed by the statute [he] wish[es] to challenge." *Susan B. Anthony List v. Driehaus*, 573 U.S. 149, 162 (2014); *see also id.* at 158-59 (explaining that "we do not require a plaintiff to expose himself to liability before bringing suit"); *Virginia v. Am. Booksellers Ass'n*, 484 U.S. 383, 393 (1988) (finding that plaintiffs had standing to bring a pre-enforcement challenge because they "alleged an actual and well-founded fear that the law will be enforced against them").

19. Because xAI is a "developer" under AB 2013, it is directly regulated by the Act and will have disclosure obligations if the law takes effect. And there is an "actual and well-founded fear that the law will be enforced against" xAI if it does not comply. *Am. Booksellers*, 484 U.S. at 393. Indeed, "[t]he State has not suggested that the newly enacted law will not be enforced," and there is "no reason to assume otherwise." *Id.* And complying with those disclosure obligations would violate xAI's constitutional and property rights for the reasons explained below. The deprivations of constitutional and property rights that enforcement of AB 2013's requirements would work are quintessential injuries in fact that demonstrate xAI's standing to bring this pre-enforcement challenge. *E.g.*, *Columbia Basin Apartment*

*Ass'n v. City of Pasco*, 268 F.3d 791, 798 (9th Cir. 2001); *Jacobs v. Clark Cnty. Sch. Dist.*, 526 F.3d 419, 426 (9th Cir. 2008).

20. Defendant Rob Bonta is the California Attorney General. Defendant is a California resident and sued in his official capacity. Attorney General Bonta, as the chief law officer of the state, Cal. Const. art. V, §13, is generally charged with enforcing California's laws, including AB 2013. AB 2013 lacks its own enforcement mechanism. But the California legislature has suggested that its provisions could be enforced through California's Unfair Competition Law, Cal. Bus. & Prof. Code §17200 *et seq.* *See* Cal. Assem. Comm. on Privacy & Consumer Prot., AB 2013, Bill Analysis 9 (Apr. 28, 2024), *available at* https://perma.cc/7SS2-UUHW. Because Attorney General Bonta has authority to enforce California's Unfair Competition Law, Cal. Bus. & Prof. Code §17204, xAI has a credible fear that the Attorney General will attempt to enforce AB 2013 against xAI via that law.

## BACKGROUND

### A. Generative Artificial Intelligence.

21. "Artificial intelligence is the science and engineering of getting machines, typically computer programs, to exhibit intelligent behavior." *United States v. Google LLC*, 747 F.Supp.3d 1, 52 (D.D.C. 2024); *see also* 15 U.S.C §9401(3) ("The term 'artificial intelligence' means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments."). Today, most AI researchers and practitioners focus on a particular type of AI: Generative Artificial Intelligence,

or gen AI for short.[1] The main distinction that sets gen AI apart from traditional forms of AI is that gen AI is able to create something new in response to user prompts, including new content such as audio, code, images, text, simulations, and videos.[2] While traditional forms of AI have existed for decades, gen AI has been around for only a few years.[3] Yet even in that short time, the capabilities of these gen AI systems have drastically progressed, precipitating a corresponding improvement in overall AI performance.[4] Some estimate that gen AI applications stand to add up to $4.4 trillion to the global economy annually.[5]

22. To achieve content-generative ability, gen AI developers use machine learning systems to train their gen AI models to recognize patterns from large datasets and make predictions or decisions based on those patterns.[6] Deep learning is a more intense subset of machine learning that uses artificial, multilayered neural networks (mathematical models that replicate the interconnected, neural structure of the human brain) that can automatically learn complex patterns from raw data,

[1] Cole Stryker & Eda Kavlakoglu, *What Is Artificial Intelligence (AI)?*, IBM, https://perma.cc/FFQ4-4N8K (last visited Dec. 13, 2025).

[2] Sadie O'Connor, *Generative AI*, 8 Geo. L. Tech. Rev. 394, 394-95 (2024).

[3] Harry Surden, *ChatGPT, AI Large Language Models, and Law*, 92 Fordham L. Rev. 1941, 1942 (2024).

[4] *Id.* at 1942-43.

[5] *What Is Generative AI?*, McKinsey & Co., (April 2, 2024), https://perma.cc/GHW4-DZSE.

[6] Stryker & Kavlakoglu, *supra*.

thereby more closely simulating the decision-making power of human intelligence.[7] As a result, deep learning systems usually require less human intervention because the system, through repeated modeling, is able to develop and learn on its own.

23. Datasets are the foundation of this process. Datasets are collections of information that are stored in an organized manner so that specific entries can be located, studied, and utilized.[8] The information within a dataset can vary wildly. It can be as simple as a table containing data points (e.g., a spreadsheet containing client information, phone numbers, addresses, names, and emails) or can include a combination of text, images, and audio (e.g., the combination of notes, photos, and voicemails saved on an individual's cellphone).[9]

24. These datasets can be gathered from a wide variety of sources. They can be acquired by purchasing pre-packaged datasets that others have compiled—as advertisers often do when they purchase datasets of information about potential consumers for a product. They can also be developed by using application programming interfaces (APIs), which allow different software applications to communicate with one another. Or datasets can be compiled from a wide variety of bespoke sources, including by licensing data from the social media platform X.

25. All AI companies seek to acquire datasets to use in their AI model

[7] *Id.*

[8] Annie Badman & Matthew Kosinski, *What Is a Dataset?*, IBM, https://perma.cc/6S3B-2HNV (last visited Dec. 15, 2025).

[9] *Id.*

training process. Many of the datasets contain information from publicly available and widely used resources. Resources like Creative Commons, the Library of Congress's Free to Use and Reuse Sets, and other large repositories of information are prime sources for obtaining data for gen AI models because of both their size and the variation in data available (e.g., text, images, and videos). As a result, most AI companies in the marketplace rely on some of the same data to train and develop their AI models.

26. But AI companies do not rely on *all* the same data. Some AI companies acquire and utilize data that others do not. Indeed, what makes one company's AI model different from another company's model depends in large part on the data a company has acquired and utilized.

27. For example, an AI company may obtain data from sources of information that are public, but not obvious locations to look for data. Take legal data. Most AI companies might use data from obvious sources of information like federal or state courts' websites. But some companies might see a benefit to also obtaining data from some less obvious sources of information, such as Justia or the Government Publishing Office. Because the latter sources might have distinct data that an AI company can use, and because that information might be presented in a different format, an AI company that relies on those specialized sources in crafting its datasets can gain a competitive advantage. Such data enables the owner to provide unique training to its AI model that differentiates it from other AI models.

28. Companies can also rely on different information from their

competitors by acquiring data from non-public sources. For example, a company may purchase test answers from a major quiz bowl competition or from a standardized test company. Or taking the legal example above, a developer may pay a company to assemble briefs addressing common legal questions to be used to train its AI model. To the extent other companies lack access to that data to train their AI models, the company that purchased data from those non-public sources gains a competitive advantage: When the owner of the non-public information trains its AI model on information its competitors are not using, it ensures that its AI model has more versatility than its competitors. In short, datasets—and in particular, the specific sources and kinds of information they contain—are the linchpin in an AI model's development process and success.

29. By using datasets that contain different forms of media, text, images, and other files in hand, AI companies can develop AI systems through the deep learning process. Before any learning occurs, the datasets need to be "cleaned" and then converted into a numeric format that a computer can read and comprehend. The cleaning process requires AI engineers to review datasets to remove unnecessary duplicate entries (e.g., repeated photos) and incomplete entries (e.g., removing phone numbers from a list that do not contain an area code). AI engineers also use this step to identify high-quality data that will be useful for the training process because it will facilitate the most learning—for example, an AI model that is meant to recognize images will initially learn more from clear, crisp photos, than from blurry images.

30. The conversion process that follows is known as "tokenization." Tokenization involves translating data taken from the datasets that have been acquired and cleaned into a list of numerical representations that the model can process called tokens.[10] A token is essentially a component of a larger dataset that represents words, characters, or phrases. *See United States v. Google LLC*, 2025 WL 2523010, at *9 (D.D.C. Sept. 2, 2025). The tokenization process is critical in preparing data for further processing in AI models.

31. Creating a gen AI system usually proceeds in several stages. The first stage is the pre-training phase, where engineers develop a foundation model—a term used to describe an AI model that is trained using a wide array of datasets aimed at covering a vast range of information and data formats (i.e., text, images, audio, etc.).[11] The goal of this process is to use a broad base of information, so that the AI model can recognize patterns and extrapolate from the universe of information it has been asked to process. The foundation model serves as the building block for more AI models that are fine tuned to perform specialized tasks, such as moving from an AI model that can generally generate images or videos based on user inputs to one that is specialized to perform graphic design for webpages.[12]

32. The most common foundation models are large language models.

[10] Mahesh Babu, *Toeknization in Deep Learning Models*, Medium (August 29, 2025), https://perma.cc/AN5A-JYTR.

[11] Rina Diane Caballar & Cole Stryker, *What Are Foundation Models?*, IBM, https://perma.cc/5MFV-H2L6 (last visited Dec. 15, 2025).

[12] *Id.*

Large language models are typically designed to understand and then generate human language, but they are sometimes designed to understand and then generate image, video, sound, or music. Some models can support several kinds of content.[13] These foundation models take data inputs, convert them to tokens, predict the most likely next token in a sequence, and then convert those predicted tokens back into a language output. *See Google*, 2025 WL 2523010, at *9-10.[14] A model's ability to predict the next token depends on the quality, diversity, and quantity of the input data. *Id.* So to build out effective foundation models, developers must input huge volumes of raw data. This process is computationally intensive (requiring thousands of clustered graphics processing units), time-consuming (taking weeks of processing), and expensive (costing millions of dollars).[15] The goal of this developmental stage is to produce a foundation model with general reasoning and generative capabilities that can then be adapted for many applications. The process is akin to teaching a child to speak.

33. With a foundation model in place, the gen AI system can be trained and fine-tuned to better respond to specific content or applications. Developers may use tools such as reinforcement learning to assess the gen AI's performance and improve its processing, thus transforming the unfiltered and potentially chaotic responses of

[13] Stryker & Kavlakoglu, *supra*.

[14] *See also id*.

[15] *Id.*; Surden, *supra*, at 1967.

a foundation model into more polished and reliable outputs.[16] This typically involves a smaller yet higher quality, focused training dataset to target specific capabilities of an AI system that has already been pretrained.[17] In creating these higher-quality datasets, developers must employ more direct human involvement to label and curate the datasets. If building the foundation model is like teaching a child to speak, this fine-tuning stage is like sending the child to school.

34. Once the model is fine-tuned, it will be regularly assessed, refined, and improved. This can occur through additional fine-tuning, or it may involve analyzing the model's actual outputs to real-world user prompts. This is an ongoing process, akin to an employee continuing to improve at his job with periodic job-performance reviews.

**B. xAI's Artificial Intelligence Systems.**

**1. xAI Develops Its Own Gen AI Systems.**

35. xAI has taken advantage of the powerful capabilities of AI models in developing its own AI systems and refining the datasets used to train those models.

36. Beginning around March and April 2023, xAI started investing in developing its own AI model for public consumption, which later became known as Grok.

37. xAI's engineers invested substantial amounts of time and energy in acquiring datasets from various sources across the Internet to develop and eventually

[16] *See also* Stryker & Kavlakoglu, *supra*.

[17] Surden, *supra*, at 1967.

train the AI models that it has produced. For the reasons outlined above, this process was key to the work xAI has performed to date. After all, acquiring datasets that are varied in source type, data type, and format is vital to xAI's ability to train AI models that are versatile enough to be adapted for consumers to use to perform a variety of tasks—whether that be providing up-to-date responses about news events or trivia, helping plan events a user wants to host, or brainstorming ideas with a user interested in launching a business.

38. As part of its development process, xAI generally used the methodology outlined above. *See supra* ¶¶31-34. xAI first acquired and refined its datasets (including finding multiple ways to tokenize or package the information for it to be processed by xAI's models)[18] to develop a foundation AI model unique to xAI.

39. From there, xAI engineers used specially constructed datasets to tweak and refine the model. xAI engineers adjusted different aspects of the xAI foundation model's code and tested the resulting variations using its datasets to see which were most effective at the tasks they were given. For example, if a variant of the model were asked to write stories, the engineers would review the resulting story for quality, creativity, and readability. Or if the variations were prompted to recognize

[18] By way of an example, a sentence can be tokenized by assigning a numerical signifier to each word, to each letter, each pair of characters, etc. Given there are many ways to represent a given sentence, the differing representations are all equally important to the testing and refining process, as they ensure the AI model can respond to many different inputs and settings and still yield the same result.

handwriting, they would be evaluated for their ability to handle different handwriting styles, fonts, and legibility in the writing samples they were given. This process is akin to pruning tree branches. Each variation of the base model is a branch that is tested by engineers to determine whether it should be discarded or left to grow and develop by repeating the process.

40. In November 2023, after a months-long effort to acquire and use datasets to develop its models, xAI succeeded in an initial limited public release of its flagship AI model, Grok-1. *See Announcing Grok*, *supra*. A few months later, xAI launched the full public release of Grok-1, *see* xAI, *Open Release of Grok-1* (Mar. 17, 2024), https://perma.cc/JN7Y-ZHSD, and shortly thereafter, xAI's engineers began developing Grok-2, which was rolled out to the public in August 2024, *see* xAI, *Grok-2 Beta Release* (Aug. 13, 2024), https://perma.cc/C75D-FG7E. That same process led to the release of Grok-3 in February 2025 and Grok-4 in July 2025. *See* xAI, *Company*, https://x.ai/company (last visited Dec. 14, 2025).

41. While xAI is constantly developing the next version of its AI models, it also releases updates to existing models and adds features to make them more versatile. For example, xAI recently released Grok 4.1, which operates like Grok-4 but has been trained to exhibit more interpersonal skills when a user provides an emotional prompt (like expressing grief over a lost pet), and to be more effective at creative writing. *See* xAI, *Grok 4.1* (Nov. 17, 2025), https://perma.cc/GF8C-E49M.

42. In addition, xAI continues to support prior models even after the next version is released, which gives new versions time to be fully deployed and tested

by the public, and gives those who use xAI's models in their products time to fully integrate the new version. Eventually, xAI discontinues use and support for earlier models, as newer versions innovate on prior versions, adapt to new uses and tasks, and become more refined and effective.

43. xAI seeks to foster consumer confidence in its AI models by continuing to test them against real-world tasks and providing up-to-date information about how they perform. *See, e.g.*, xAI, *Grok 4 Fast Model Card* (last updated Sept. 19, 2025), https://perma.cc/FEK8-CYQV. For example, xAI evaluates whether its models exhibit political bias when asked about controversial topics, as well as how its models fare when being pressured to give the wrong answer to a question. *See id.* at 4. These tests are valuable for consumers, as they demonstrate how well xAI's models can perform the various tasks for which consumers may want to use them (e.g., summarizing text, performing math problems, or answering questions about current events). And because the tests replicate how a consumer would actually interact with a model and the results it would see, they help consumers assess whether using the AI model is likely to be worthwhile.

44. Despite being around only since 2023—significantly later than many of its competitors—xAI has consistently released models that rank at the top of performance benchmarks and leaderboards, demonstrating the success of xAI's development process and the value of the datasets it has obtained and curated.

**2. xAI Holds Trade Secrets in Several Facets of Its Datasets.**

45. Due to the critical role they play in the development process,

information about the datasets and processes AI developers use to train their AI models is a closely protected trade secret under state and federal law.

46. Under California law, a party holds a trade secret in information that (1) "is valuable because it is unknown to others and (2) … the owner has attempted to keep secret." *Amgen Inc. v. Cal. Corr. Health Care Servs*., 260 Cal.Rptr.3d 873, 886 (Cal. Ct. App. 2020); Cal. Civil Code §3426.1(d). Under the parallel federal law, the DTSA, a party has a protected trade-secrets property interest in information that she intends to use in interstate commerce if (1) the information "derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable through proper means by, another person who can obtain economic value from the disclosure or use of the information" and (2) the owner "has taken reasonable measures to keep such information secret." 18 U.S.C. §1839(3).

47. xAI's dataset information, which xAI uses to develop its AI models that are disseminated in interstate commerce, easily satisfies both standards. They thus constitute trade secrets under both the DTSA and California law.

48. First, dataset information is economically valuable precisely because it is unknown to others. The specific sources of data used to construct xAI's gen AI inform both the kind of information xAI's models are trained on and how useful that information is to the model's performance. *See O2 Micro Int'l Ltd. v. Monolithic Power Sys., Inc*., 420 F.Supp.2d 1070, 1089 (N.D. Cal. 2006) ("Combinations of public information from a variety of different sources when combined in a novel

way can be a trade secret."); *accord WeRide Corp. v. Kun Huang*, 379 F.Supp.3d 834, 847 (N.D. Cal. 2019). For example, an AI developer seeking to produce an AI model capable of image generation may rely on datasets of photos from photo hosting sources such as Wikimedia Commons. But another AI company might think to use other sources of photos, such as Openverse or Public Domain Pictures, or instead license datasets containing a wide range of images from sources like Getty Images or Adobe Stock. If one company does not use a dataset used by another, its AI model may not be as well trained. If xAI were to reveal all the datasets that it uses, its competitors would immediately move to acquire those sources to ensure their models were equally as effective.

49. The amount of data that xAI uses is also valuable precisely because it is unknown to others. The breadth of information engineers rely on to develop xAI's models reveals details about the model's ability to handle a wide range of tasks. A model trained to recognize images, text, and audio, for example, would have more capabilities than one trained to recognize only text.

50. There is accordingly significant value in xAI keeping not just the content, but also the size, of its datasets secret. Many AI companies will have overlap in the datasets they use. But it is the *differences* between the datasets each company uses that gives xAI a competitive edge. If competitors could see the sources of all of xAI's datasets, or even the size of its datasets, competitors could evaluate both what data xAI has and how much they lack. And from there, they could better understand and try to fill any gap in their own models' knowledge to

ultimately produce a better product.

51. xAI's processes for cleaning, modifying, and refining the datasets it has obtained are economically valuable information too. Those details reveal the methods xAI has used to choose which datasets are appropriate for training its AI models, and which data xAI engineers believe will be most effective for training and refining the AI model's capabilities. Those methods are economically valuable because the specific steps xAI employees take to identify the data that will be used to train xAI's models provide a roadmap for competitors to understand how to replicate xAI's development process. Those steps shed light on the strategic choices xAI engineers have made in advancing the capabilities of its AI models—for example, whether they believe judicial opinions or other sources will be more effective at training a creative-writing AI model to write stories based on user prompts. As a result, other companies could use information about xAI's processes for selecting and modifying datasets to undercut any competitive edge xAI holds in its development process by replicating and improving on those methods.

52. Put simply, dataset information is "valuable because it is unknown." *Amgen Inc.*, 260 Cal.Rptr.3d at 886; *see also Comet Techs. U.S. of Am. Inc. v. Beuerman*, 2018 WL 1990226, at *3 (N.D. Cal. Mar. 15, 2018) (recognizing under the DTSA that "[a] collection of data that allows the holder to recreate one of Plaintiff's top technologies derives its value from not being generally known, because Plaintiff's competitive edge would evaporate if the public and its competitors could easily recreate its products"); *Whyte v. Schlage Lock Co.*, 125

Cal.Rptr.2d 277, 287 (Cal. Ct. App. 2002) (trade secrets have "independent economic value" because the information would be "valuable to a competitor" if disclosed). To be sure, some of that information comes from publicly available sources. But xAI's decisions about how to craft the best mixture of data from public and non-public sources—and which public and non-public sources—are what helps it to develop and train top-notch AI models. *See O2 Micro Int'l*, 420 F.Supp.2d at 1089; *WeRide*, 379 F.Supp.3d at 847 (that plaintiff "derived much of its source code from open source code … does not mean that [plaintiff's complete] source code … was not confidential"). And decisions that xAI makes about things like how to tokenize, and whether to use an entire dataset or only a subset during the training process, are valuable and non-public too, as they are equally essential to xAI's ability to develop unique AI models. *Abba Rubber Co. v. Seaquist*, 286 Cal.Rptr. 518, 527 (Cal. Ct. App. 1991) (noting that information is valuable to competitors "if it indicates to them a fact which they previously did not know"); *accord E. W. Bank v. Shanker*, 2021 WL 3112452, at *9 (N.D. Cal. July 22, 2021) (explaining that parties can hold trade-secrets property rights in "roadmaps related to confidential technology" because such roadmaps would enable competitors to recreate the resulting product). In short, granular information about xAI's datasets—including the sources xAI uses, the volume of data (or tokens) xAI has and uses, the types of data xAI uses, and how xAI uses such data in the training process—is all valuable precisely because it is not known to xAI's competitors.

53. Courts around the country already recognize that AI training data

constitute economically valuable, confidential information that warrants protection against disclosure during litigation and discovery proceedings. *See, e.g.*, Order at 2, *In re OpenAI ChatGPT Litig.*, No. 3:23-cv-3223 (N.D. Cal. Sept. 24, 2024), Dkt.182 (court order designating OpenAI's training data as "Highly Confidential – Attorneys' Eyes Only"); Order at 3, *Dow Jones v. Perplexity AI*, No. 1:24-cv-7984 (S.D.N.Y. Aug, 1, 2025), Dkt.62 (Confidentiality and Protective Order describing training data as "extremely sensitive" and designating it "Highly Confidential").

54. Given the obvious economic importance of this highly confidential data, xAI makes a significant effort to prevent its disclosure and maintain its secrecy. All employees sign confidentiality provisions when they execute their employment agreements to work for xAI, and to work in the AI development process in particular. These confidentiality provisions communicate to employees that all parts of the development process are xAI's non-public, proprietary information, that that information is to be used solely for development of xAI's models, and that none of it should be publicly disclosed. *See Whyte*, 125 Cal.Rptr.2d at 286-87; *Google LLC v. Point Fin., Inc.*, 2025 WL 1616533, at *4 (N.D. Cal. June 6, 2025).

55. xAI has also made sure that its datasets are access gated and accessible only to individuals with appropriate level of access. *See Whyte*, 125 Cal.Rptr.2d at 286-87; *WeRide*, 379 F.Supp.3d at 847; *BladeRoom Grp. Ltd. v. Facebook, Inc.*, 219 F.Supp.3d 984, 986-87 (N.D. Cal. 2017). The location of data storage is known only to the individuals that need access for an approved purpose, and xAI's systems also alert security when certain datasets are accessed or moved. That alert ensures that

the appropriate xAI team can contact the individual whose actions were flagged to ensure the access was authorized and make sure that the information is being used only for an authorized purpose.

56. On top of all that, xAI has made clear to employees that they can and should report any unauthorized access that they observe to xAI's legal or incident response teams so that xAI can actively remedy any potential misuse of its datasets. That reporting structure is critical to safeguarding xAI's datasets, as it allows xAI to rely on its employees' insight on the ground to help protect and safeguard its highly valuable proprietary information.

57. To enhance security even further, xAI has introduced role-based access requirements, which ensure that an employee's access is limited to the datasets they actually need to use. That extra layer of security helps xAI ensure that it can respond quickly to any unauthorized use and preempt any improper disclosure.

58. xAI also has a robust confidentiality policy that underscores the confidentiality provisions in employees' employment contracts. This policy states explicitly that all information, including xAI's datasets and process and methods, is protected and non-public.

59. In addition, xAI will implement time-limited access controls, which will require a user to gain re-approval after a set number of days, so that permission to access data does not extend beyond the needs of a particular project.

60. In short, xAI takes extremely seriously the importance of safeguarding its datasets, information relating to them, and its methods and processes. Indeed,

xAI is constantly refining its security systems to ensure that xAI alone has access to the highly valuable information that is critical to its competitive edge in innovating AI models. xAI thus plainly holds trade secrets in that information under California law and under the DTSA.

**C. California Passes Assembly Bill 2013.**

61. On September 28, 2024, the Governor of California signed into law AB 2013. The bill, entitled "Artificial Intelligence Training Data Transparency," imposes substantial information-disclosure requirements on developers of gen AI. It is scheduled to take effect on January 1, 2026. Cal. Civ. Code §3111.

62. While AB 2013 itself does not contain any statement of purpose, the legislature claimed that its goal is to "provide[] transparency to consumers of AI systems and services by providing important documentation about the data used to train the services and systems they are being offered." *See Senate Floor Analysis*, *supra*, at 6; *accord* Cal. Assem., AB 2013, *Assembly Floor Analysis* 2 (Aug. 27, 2024), *available at* https://perma.cc/7SS2-UUHW. In particular, the law's disclosure obligations purportedly aim to "help[] identify and mitigate biases," *Senate Floor Analysis*, *supra*, at 3, based on the notion that "garbage in" is "garbage out"—the quality of the data going in affects the quality of the ultimate product, *Assembly Floor Analysis*, *supra*, at 2. Yet it is hard to see how AB 2013's requirements do anything to accomplish that goal, as they do not require gen AI companies to disclose the kinds of information that consumers typically find useful—e.g., how well an AI model has performed when given particular tasks.

They instead require companies to disclose information that is highly valuable to their *competitors*—namely, the very proprietary dataset information that xAI works so hard to keep secret.

63. In particular, AB 2013 requires any "developer" of a "generative artificial intelligence system or service, or a substantial modification to a generative artificial intelligence system or service" made "publicly available to Californians for use" after January 1, 2022, to "post on the developer's internet website documentation regarding the data used by the developer to train the generative artificial intelligence system or service." Cal. Civ. Code §3111.

64. Those disclosures must include, but are not limited to, a "high-level summary of the datasets used in the development" of the generative AI. *Id.* §3111(a). And the statute sets out a list of 12 categories of dataset information that developers must also publicly disclose:

- The sources or owners of the datasets.
- A description of how the datasets further the intended purpose of the AI.
- The number of data points included in the datasets.
- A description of the types of data points within the datasets.
- Whether the datasets include data protected by copyright, trademark, or patent.
- Whether the datasets were purchased or licensed by the developer.
- Whether the datasets include personal information.

- Whether the datasets include aggregate consumer information.
- Whether there was any cleaning, processing, or other modification to the datasets, and the intended purpose of those efforts.
- The time period during which the data was collected, including if the data collection is ongoing.
- The dates the datasets were first used during the development AI.
- Whether the AI system or service used or continuously uses synthetic data generation in its development.

*See id.* §3111(a)(1)-(12).

65. These disclosure requirements apply to all AI models, with three narrow exceptions: (1) AI models used solely to help ensure security and integrity (i.e., "to detect security incidents that compromise the availability, authenticity, integrity, and confidentiality of stored or transmitted personal information," *Id.* §1798.140(ac)); (2) AI models used solely in the operation of aircraft in the national airspace; and (3) AI models developed for national security and made available only to a federal entity. *Id.* §3111(b). Those exceptions underscore that California recognizes that the information AB 2013 requires companies to disclose is valuable precisely because it is not public. After all, if there were no value to keeping the information secret, then it is hard to see why California would exempt such models from its requirements.

66. AB 2013 does not define the key term "datasets." Nor does it explain how "high-level" a "summary" must be to satisfy the law. For example, is it

sufficient for a developer to disclose the "Internet" generally as the source of the dataset? Or must it include specific details and sources (e.g., "state and federal court websites" or the "Library of Congress")? And AB 2013 provides no guidance as to whether developers can simply provide a "yes" or "no" answer to questions like whether their datasets contain information that is protected intellectual property, consumer data, or personal information, or whether they must instead identify any such protected information. The same problem plagues AB 2013's requirement to disclose whether datasets are "clean[ed], process[ed], or [subject to] other modifications"; AB 2013 does not make clear whether a "yes" or "no" answer is sufficient, or if a developer must instead flesh out what cleaning or refining was done to each dataset.

67. AB 2013 does, however, define "developer" extremely broadly to include any person or entity that "designs, codes, produces, or substantially modifies an artificial intelligence system or service." *Id.* §3110(b). The law thus will require xAI to locate, collect, summarize, and disclose extensive information about the datasets it uses or has used to train and develop each and every one of its AI models released since 2022 even if they are no longer being used, including every update or modification to an existing model.

68. Those disclosures will not only impose onerous compliance burdens on xAI, but will also force xAI to reveal confidential information about how it develops and trains its AI models. Meanwhile, the requirements will do little to help consumers assess the things they actually care about, like whether a particular AI

model is effective or user-friendly. Consumers measure the value of an AI model based on *outcomes*, by testing its responsiveness to particular types of inquiries. *See, e.g.*, *Grok 4 Fast Model Card*, *supra*. The disclosures that AB 2013 requires tell consumers little, if anything, about that. Moreover, to the extent that California aims to have AB 2013's disclosure requirements "help[] identify and mitigate biases," *Senate Floor Analysis*, *supra*, at 3, based on the notion that "garbage in" is "garbage out," *Assembly Floor Analysis*, *supra*, at 2, that only underscores that AB 2013's provisions target First Amendment protected speech based on its content and viewpoint. Regardless, given xAI already discloses the results of a wide variety of tests, including the "political bias" of its AI models, *see, e.g.*, *Grok 4 Fast Model Card*, *supra*, AB 2013's requirements miss their mark.

69. Indeed, the only thing AB 2013 seems to do is force developers to provide their competitors with a roadmap to mirror their success. It gives competitors invaluable insight into how an AI model is trained, what datasets are used, what data datasets are not, and more—all information that others can exploit for their own competitive advantage. In short, it is hard to see how the copious information AB 2013 would require xAI to disclose has meaning to anyone *but* other developers trying to build their own AI models.

## CLAIM FOR RELIEF

### COUNT ONE
### (*Per Se* Takings)
### Takings Clause of the Fifth Amendment to the U.S. Constitution
### 42 U.S.C. §1983; *Ex parte Young*; 28 U.S.C. §§2201(a) and 2202

70. xAI incorporates by reference the allegations contained in the preceding paragraphs as though fully set forth herein.

71. The Fifth Amendment to the U.S. Constitution provides that "private property [shall not] be taken for public use, without just compensation." U.S. Const. amend. V. That prohibition on taking private property for public use without just compensation applies to the states through the Fourteenth Amendment to the U.S. Constitution. U.S. Const. amend. XIV, §1.

72. As the Supreme Court has long explained, "[b]y requiring the government to pay for what it takes, the Takings Clause saves individual property owners from bearing 'public burdens which, in all fairness and justice, should be borne by the public as a whole.'" *Sheetz v. Cnty. of El Dorado*, 601 U.S. 267, 273-74 (2024) (quoting *Armstrong v. United States*, 364 U.S. 40, 49 (1960)).

73. The Takings Clause covers both *per se* takings and regulatory takings.

74. A *per se* taking occurs when the government effects "a direct government appropriation or physical invasion of private property." *Lingle v. Chevron U.S.A. Inc.*, 544 U.S. 528, 537 (2005). When it comes to *per se* takings, "the Takings Clause imposes a clear and categorical obligation to provide the owner with just compensation." *Cedar Point Nursery v. Hassid*, 594 U.S. 139, 147 (2021).

75. Whether the government has effected a regulatory taking, by contrast, depends on "factors such as the economic impact of the regulation, its interference with reasonable investment-backed expectations, and the character of the government action." *Id.* at 148. But that balancing test has "no place" when the government "appropriates for the enjoyment of third parties" "a fundamental element of the [owner's] property right," *id.* at 149-50, or "otherwise interfere[s] with the owner's right to exclude others from it," *Sheetz*, 601 U.S. at 274. "That sort of intrusion on property rights is a per se taking" that "trigger[s]" the "right to compensation" without regard to any balancing of burdens, expectations, and degree. *Id.*

76. The Supreme Court has long made clear that the Takings Clause protects intangible property rights from government appropriation. *Ruckelshaus v. Monsanto Co.*, 467 U.S. 986, 1003-04 (1984) (collecting cases that "ha[ve] found other kinds of intangible interests to be property for purposes of the Fifth Amendment's Taking[s] Clause"); *cf. City of Cincinnati v. Louisville & Nashville R.R. Co.*, 223 U.S. 390, 400 (1912) ("every description of property … [including] tangibles and intangibles alike" can be subject to appropriation). Trade secrets are one category of intangible property that receives constitutional protection against government appropriation. *Ruckelshaus*, 467 U.S. at 1003-04. Indeed, trade secrets are protected property rights under both the DTSA and California law, *see supra* ¶¶45-60, underscoring their entitlement to the protections of the Takings Clause. *See Tyler v. Hennepin Cnty.*, 598 U.S. 631, 638 (2023) (looking to governing law as

an important source for identifying property rights protected by Takings Clause).

77. The key feature of the trade-secret property right is its secrecy. By definition, a "trade secret" is information that "[d]erives independent economic value … *from not being generally known* to the public or to other persons who can obtain economic value from its disclosure or use." Cal. Civ. Code §3426.1(d) (emphasis added); 18 U.S.C. §1839(3)(B); *accord Quintara Biosciences, Inc. v. Ruifeng Biztech, Inc.*, 149 F.4th 1081, 1085 (9th Cir. 2025) ("By definition, trade secrets derive their value from nondisclosure."). It therefore follows that "[t]he property in a trade secret is the power to make use of it to the exclusion of the world," *Hartley Pen Co. v. U.S. Dist. Ct.*, 287 F.2d 324, 328 (9th Cir. 1961), making the "right to exclude" others from knowing or accessing that secret information the "*sine qua non*" of that property interest, *Cedar Point*, 594 U.S. at 150. In short, "[w]ith respect to a trade secret, the right to exclude others is central to the very definition of the property interest." *Ruckelshaus*, 467 U.S. at 1011.

78. Because the right to exclude is "one of the most treasured rights of property ownership," laws that appropriate the right to exclude work a *per se* takings. *Cedar Point*, 594 U.S. at 149, 155. The categorical duty to pay compensation is accordingly triggered whenever the government prevents a property owner from restricting others from accessing its property. *Id.*

79. By compelling xAI to disclose "how [its] datasets further the intended purpose" of xAI's models, Cal. Civ. Code §3111(a)(2), the number of data points or tokens xAI uses, *id.* §3111(a)(3), and the types of data xAI has culled for developing

its AI models, *id.* §3111(a)(4), AB 2013 effects a *per se* taking, as those obligations appropriate xAI's trade secrets. To the extent AB 2013 compels xAI to reveal the sources of its datasets beyond the Internet writ large, *see* Cal. Civ. Code §3111(a)(1), (6), that disclosure would appropriate xAI's trade secrets in the sources of its datasets, effecting a *per se* taking too. And while it is not at all clear what information xAI must provide to satisfy AB 2013's obligation to disclose whether its datasets contain intellectual property, *id.* §3111(a)(5), personal or aggregate consumer information, *id.* §3111(a)(7)-(8), and whether xAI uses cleaning or modification of its datasets, including the use of synthetic data, *id.* §3111(a)(9), (12), to the extent the law requires anything more than a "yes" or "no" answer, that too would appropriate xAI's trade secrets, as all of that information constitute protected trade secrets too, *see supra* ¶¶45-60. Simply put, AB 2013 would force xAI to disclose to all the world information that it has a protected property interest in keeping secret, thereby eviscerating the very core of its trade-secret property right.

80. By eliminating xAI's ability to exercise its right to exclude others from its trade secrets—the defining feature of the trade-secret property right—AB 2013 effects a *per se* taking. *See Armstrong*, 364 U.S. at 48-49 (government action that eviscerates the value of a mechanic's lien effects a categorical taking because, "[b]efore the liens were destroyed, the lienholders admittedly had compensable property," and "[i]mmediately afterwards, they had none"); *cf. Dolan v. City of Tigard*, 512 U.S. 374, 394 (1994) (government action preventing a property owner from controlling when members of the public could access her property would

"eviscerate[]" her "right to exclude"). If xAI must comply with AB 2013, then its trade secrets will be "disclosed to others," and xAI will "ha[ve] lost [its] property interest" in those trade secrets entirely. *Ruckelshaus*, 467 U.S. at 1011.

81. Because AB 2013 works a *per se* taking, California can impose its requirements (if at all) only if it provides just compensation for that which it takes. Yet AB 2013 does not contemplate compensating xAI for its lost property rights, let alone provide any mechanism to do so. California accordingly cannot enforce AB 2013 against xAI consistent with the Takings Clause.

82. xAI is therefore entitled to an order declaring that AB 2013 effects uncompensated *per se* takings of xAI's trade secrets in violation of the Takings Clause and enjoining the Attorney General from enforcing it.

**COUNT TWO**
**(Regulatory Takings)**
**Takings Clause of the Fifth Amendment to the U.S. Constitution**
**42 U.S.C. §1983; *Ex parte Young*; 28 U.S.C. §§2201(a) and 2202**

83. xAI incorporates by reference the allegations contained in the preceding paragraphs as though fully set forth herein.

84. Because application of AB 2013's disclosure obligations to xAI would accomplish the functional equivalent of a classic taking, they constitute a regulatory taking under the Takings Clause as well.

85. As explained, to determine whether a regulatory taking has occurred, courts balance the factors set forth in *Penn Central Transportation Co. v. City of New York*: (1) "[t]he economic impact of the regulation on the claimant;" (2) "the

extent to which the regulation has interfered with distinct investment-backed expectations;" and (3) "the character of the governmental action." 438 U.S. 104, 124 (1978).[19] Here, all three factors support finding that AB 2013 would effect all manner of regulatory takings.

86. First, xAI invested in and developed its trade secrets starting in March and April of 2023 (*i.e.*, well before AB 2013 was first introduced in January 2024) and continuing to today with certain investment-backed expectations. California law has protected trade secrets for decades. Indeed, its Uniform Trade Secrets Act ensures that individuals who have an interest in preserving the secrecy of their trade secrets can prevent others from unlawfully accessing them. *See* Cal. Civ. Code §§3426 *et seq.* The same is true of federal law. Private parties can invoke the DTSA's private cause of action to sue third parties who have misappropriated their trade secrets. 18 U.S.C. §1836. Federal law also authorizes the federal government to pursue criminal penalties against any party that has knowingly stolen such rights. *See id.* §§1831 *et seq.* Accordingly, xAI had a reasonable expectation that the datasets it used to train AI models, the value of which is derived from their non-disclosure, would be protected under both federal and state law. After all, California has certainly never suggested that AI developers and the models they create are somehow excluded from ordinary property rules.

[19] xAI reserves the right to ask the Supreme Court to overrule *Penn Central*.

87. xAI accordingly had no reason to know or expect that it would not be able to reap the value of its trade secrets in developing specialized datasets to train its AI models. That is especially so because the provisions at issue were not introduced until the calendar year after xAI first began acquiring and developing datasets for training its models, and several months after xAI released its first AI model to the public.

88. The Supreme Court has made clear that the government cannot "manipulate[]" or "extinguish a property interest that it recognizes everywhere else to avoid paying just compensation when it is the one doing the taking." *Tyler*, 598 U.S. at 645. Yet AB 2013's novel disclosure regime does precisely that. By compelling disclosure of xAI's valuable trade secrets, *see supra* ¶¶45-60, AB 2013 bucks the longstanding protections of a trade-secrets owner's right to keep such information secret under both federal law and California property law.

89. California has compounded that disruption of settled expectations by rendering its law retroactive to January 2022. *See, e.g.*, *INS v. St. Cyr*, 533 U.S. 289, 315 (2001) ("Retroactive statutes raise special concerns."). That retroactive aspect of AB 2013 is particularly troubling, as it targets xAI's investments made in developing those trade secrets at a time when xAI had no notice whatsoever that the highly confidential information it was developing could become subject to sweeping disclosure obligations.

90. In short, xAI could not reasonably expect California to destroy this longstanding property interest by legislative fiat, as it has done with AB 2013. *See*

*Murr v. Wisconsin*, 582 U.S. 383, 396 (2017) ("States do not have the unfettered authority to 'shape and define property rights and reasonable investment-backed expectations,' leaving [property-]owners without recourse against unreasonable regulations.").

91. Second, AB 2013's enforcement would substantially interfere with the economic value of xAI's interest in the datasets and processes it has developed to assist with training its AI models. Those datasets and processes derive their core value from their secrecy; after all, that is why they receive protection as trade secrets under California and federal law. *See* Cal. Civ. Code §3426.1(d); 18 U.S.C. §1839(3). By preserving the secrecy of this information, xAI maintains a competitive advantage in the AI marketplace, as the quality of its datasets and their sources and size all shape how xAI trains its AI models, which in turn informs the models' effectiveness and value to the consumer. *See Andersen v. Stability AI Ltd.*, 2025 WL 1927796, *2 (N.D. Cal. July 14, 2025) (recognizing that forcing an AI developer to disclose AI training data to a competitor "raises serious competitive concerns" and "poses a risk of harm"). AB 2013 eviscerates the value of that information, as it forces xAI to tell other developers exactly what they need to do to improve the quality of their own AI models.

92. For those reasons, compelled disclosure under AB 2013 would have severe economic impacts on xAI. As explained, xAI's interest in its datasets (including their size, contents, and sources) hold value precisely because they are not publicly known. "[D]isclosure of that data [thus] entirely extinguish[es] the

value of the trade secret[s]." *Pharm. Rsch. & Mfrs. of Am. v. Stolfi*, 153 F.4th 795, 839 (9th Cir. 2025). After all, even if the data can still be used to train the AI systems, the entire reason xAI keeps that information secret is so that it can produce AI models that are distinct and more effective relative to its competitors' models. *See Ruckelshaus*, 467 U.S. at 1012 ("That the data retain usefulness … even after they are disclosed—for example, as bases from which to develop new products or refine old products … —is irrelevant to the determination of the economic impact[.] … The economic value … lies in the competitive advantage over others that [xAI] enjoys by virtue of its exclusive access to the data, and disclosure or use by others of the data would destroy that competitive edge."). That is all the more important given that AI is a nascent area rife for substantial growth and development, making *any* competitive advantage key to xAI's long-term success in the space. In short, disclosure would eviscerate the value of xAI's trade secrets; its competitors cannot replicate xAI's models without the information AB 2013 would compel xAI to disclose.

93. Third, the character of the government action weighs heavily in favor of finding that AB 2013's disclosure regime would effect a regulatory taking of xAI's trade secrets. As explained, *see supra* ¶¶48-53, xAI's ability to keep dataset information private is core to its property interest in those trade secrets. Disclosure of that information therefore would eviscerate that property right. Even if AB 2013 does not amount to a *per se* taking, the fact that xAI's "property right is extinguished" once it is forced to comply with AB 2013's disclosure requirements,

*Ruckelshaus*, 467 U.S. at 1002, confirms that AB 2013 accomplishes the "functional[] equivalent" of a classic *per se* taking, *Lingle*, 544 U.S. at 539; *Hodel v. Irving*, 481 U.S. 704, 715-17 (1987) (holding that the total abrogation of the right to devise amounted to a taking of that property right based solely on the "extraordinary" character of the government action).

94. The categorical obligation to disclose information to the public at large—without *any* protection for the proprietary nature of that information—reinforces the conclusion that the character of California's action weighs in favor of a taking. Unlike more narrowly focused disclosure statutes that are triggered only once the state has made a case-specific determination that there is a public need for disclosure, AB 2013 broadly compels disclosure of trade secrets regardless of whether there is any identifiable need to make that information public. *Cf. Stolfi*, 153 F.4th at 839-40.

95. That is more troubling because it is not at all clear how the public is supposed to benefit from the trade secrets that AB 2013 would compel xAI to share. After all, AB 2013 does not compel the disclosure of truthful, factual information to combat misleading claims about how AI products work, *see Zauderer v. Off. of Disciplinary Couns.*, 471 U.S. 626, 638 (1985), or alert the public of potential risks associated with them. Indeed, it is unclear what the public is even supposed to do with the information AB 2013 would force xAI to disclose.

96. For instance, why would consumers care how many data points are in a given training dataset? Without the requisite technical expertise, there is no way

for a consumer to know whether, for example, an AI model related to improving driving directions that is trained on a dataset containing a thousand different road maps is better than one trained on ten thousand road maps. The consumer's best metric is the end product, not the amount of data used.

97. It is equally unclear how learning the sources of the data used to train an AI model would help a consumer evaluate its quality. A lawyer may think that a research AI model that draws from state and federal court websites and one that draws from Cornell's Legal Information Institute could yield similar results. But she may not appreciate that slight distinctions in how each source stores and displays statutory text or judicial opinions can affect an AI model's training process and, as a result, the quality of the research it can produce. Again, the best metric for assessing the relative quality of such models—and the one the lawyer would actually care about—is how each performs when given actual research tasks.

98. So too with information about how xAI intends to use a certain dataset to train its AI models. For example, if a consumer were to know that xAI intended to use a dataset drawn in part from bar trivia questions and answers a certain number of times to help train an AI model that would not give consumers insight into how valuable the model's functionality will be. After all, the consumer could not predict whether that dataset is sufficiently comprehensive, whether xAI engineers provided enough reinforcement for the AI model to learn which questions it has answered correctly, or if the AI model adequately learned how to frame its answers in a manner

that is easy to comprehend. Again, the user would glean more from seeing the results of tests performed on the AI model.

99. The principal beneficiary of these disclosure requirements thus are not consumers, *but rather competitors*, who will use the information to bolster their own products. Namely, competitors will use the disclosures to pinpoint deficiencies in the size of their training sets, identify source data they do not have, and spot issues in how they are using datasets to train their AI models. xAI would thus bear the entire burden of that boon to competitors through the loss of its trade-secrets property rights. *Cienega Gardens v. United States*, 331 F.3d 1319, 1338-39 (Fed. Cir. 2003) (character-of-government-action prong weighed in favor of a taking because placing the burdens of a public program on individual property owners was "the kind of expense-shifting to a few persons that amounts to a taking"). The final prong of the *Penn Central* analysis accordingly also weighs heavily in favor of finding that AB 2013's disclosure regime accomplishes a taking.

100. Moreover, because AB 2013 provides no benefit to the public at large, but transfers the entire value of xAI's trade secrets to its private-company competitors, the law cannot even satisfy the "public use" requirement of the Takings Clause. *See Kelo v. City of New London*, 545 U.S. 469, 477 (2005) ("[I]t has long been accepted that the sovereign may not take the property of *A* for the sole purpose of transferring it to another private party *B,* even though *A* is paid just compensation."). For that reason too, AB 2013 effects an unconstitutional taking.

101. At bare minimum, AB 2013's mandated disclosure of xAI's trade secrets would effect an uncompensated regulatory taking. xAI is therefore entitled to an order declaring that AB 2013 violates the Takings Clause and enjoining the Attorney General from enforcing it.

**COUNT THREE**
**(Compelled Speech)**
**First Amendment to the U.S. Constitution**
**42 U.S.C. §1983; *Ex parte Young*; 28 U.S.C. §§2201(a) and 2202**

102. xAI incorporates by reference the allegations contained in the preceding paragraphs as though fully set forth herein.

103. Not only does AB 2013 effectuate an unconstitutional taking; it also violates the First Amendment. By forcing xAI to disclose information about its datasets—including its confidential trade secrets, *see supra* ¶¶45-60—AB 2013 compels xAI to speak in violation of its right to free speech.

104. The Supreme Court has long recognized that the First Amendment's guarantee of free speech "includes both the right to speak freely and the right to refrain from speaking at all." *Wooley v. Maynard*, 430 U.S. 705, 714 (1977); *see also X Corp. v. Bonta*, 116 F.4th 888 (9th Cir. 2024). That protection extends "not only to expressions of value, opinion, or endorsement, but equally to statements of fact the speaker would rather avoid." *Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Boston*, 515 U.S. 557, 573 (1995).

105. Laws compelling speech are generally treated no differently from laws restricting speech, *see, e.g.*, *Nat'l Inst. of Fam. & Life Advocs. v. Becerra* (*NIFLA*),

585 U.S. 755, 766-67 (2018), even when the government does not compel a speaker to express any particular message, *Riley v. Nat'l Fed'n of the Blind of N.C., Inc.*, 487 U.S. 781, 795 (1988). As the Supreme Court has explained, the standard First Amendment analysis applies equally when the government compels speakers to convey purely factual information. *See Hurley*, 515 U.S. at 573.

106. There is a narrow limited "exception [to the] compelled speech" doctrine for government-compelled speech that aims to combat misleading advertisements. *CTIA - The Wireless Ass'n v. City of Berkeley*, 928 F.3d 832, 843 (9th Cir. 2019); *see Zauderer*, 471 U.S. 626. But the Supreme Court has never applied the principles set forth in *Zauderer* outside the context of misleading advertising. *See NIFLA*, 585 U.S. at 768. To the contrary, the Court has consistently reaffirmed that *Zauderer* was focused only on "combat[ing] the problem of inherently misleading commercial advertisements." *Milavetz, Gallop & Milavetz P.A. v. United States*, 559 U.S. 229, 250 (2010); *see also, e.g.*, *Hurley*, 515 U.S. at 573; *United States v. United Foods*, 533 U.S. 405, 416 (2001).

107. On top of that, *Zauderer* applies only "within the class of commercial speech." *X Corp*, 116 F.4th at 900; *see also Am. Beverage Ass'n v. City & Cnty. of San Francisco*, 916 F.3d 749, 756 (9th Cir. 2019) (*Zauderer* only applies to "a First Amendment claim involving compelled commercial speech"). In other words, it applies only when the government is regulating speech that "propose[s] a commercial transaction." *NetChoice, LLC v. Bonta*, 113 F.4th 1101, 1119 (9th Cir. 2024); *see also Stolfi*, 153 F.4th at 821 (finding commercial speech where compelled

information "communicate[d] the terms of potential commercial transaction[s]" and was "tethered to commercial transactions."). *X Corp. v. Bonta* is instructive. Because the compelled-disclosure reports at issue there did not "communicate[] the terms of an actual or potential transactions," and instead "require[d] a company to recast its content-moderation practices in language prescribed by the State," they had "few indicia of commercial speech." 116 F.4th at 901.

108. *Zauderer* plainly has no application here. AB 2013 requires xAI to disclose its dataset gathering practices in language prescribed by California. None of that information has anything to do with "propos[ing] a commercial transaction," let alone with addressing any potentially misleading advertising or claims about a product or service. *Id.* at 900. Indeed, the law applies to all AI systems available for public use "regardless of whether the terms of that use include compensation." Cal. Civ. Code §3111. This case thus does not involve the kind of compelled-speech mandate that might get lesser scrutiny under *Zauderer*.

109. To the contrary, AB 2013 is a content-based regulation that triggers strict scrutiny, as it compels xAI to disclose specific *content* related to its AI models. *X Corp*, 116 F.4th at 900; *NIFLA*, 585 U.S. at 766. Strict scrutiny applies whenever the government forces companies to disclose particular content outside the context of commercial speech, whether it be factual information about a company's service and products, *see X Corp.* 116 F.4th at 902, risks those services or products may pose, *NetChoice*, 113 F.4th at 1119, or information about where to obtain other types of products or services, *see NIFLA*, 585 U.S. at 766. Like those content-based

disclosures, AB 2013 forces xAI to speak a particular message it does not want to convey to the public.

110. AB 2013 compounds the First Amendment problems by discriminating based on viewpoint.

111. AB 2013's viewpoint discrimination is clear from its face. AB 2013 exempts from its requirements gen AI models with certain favored "purpose[s]." Cal. Civ. Code §3111(b). In particular, California has extended favorable treatment to AI models "whose sole purpose is to help ensure security and integrity" or "the operation of aircraft in the national airspace," or were "developed for national security, military, or defense purposes [and are] … made available only to a federal entity" by exempting them from the disclosure obligations that would otherwise apply. *Id.* §3111(b)(1)-(3). Those "purpose"-based distinctions, however, are proxies for viewpoint discrimination. *See City of Austin v. Reagan Nat'l Advert. of Austin, LLC*, 596 U.S. 61, 74 (2022). It compels speech "based on the ideas or opinions it conveys," but allows other speakers with different views to stay silent. *Iancu v. Brunetti*, 588 U.S. 388, 393 (2019) (ban on registering "immoral" or "scandalous" trademarks was impermissibly viewpoint-based); *see, e.g.*, *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 565 (2011) (restrictions on speech "promot[ing] brand-name drugs" were impermissibly "aimed at a particular viewpoint").

112. The First Amendment does not permit California to compel private speech based on its perception that certain ideas (i.e., information regarding data developers use to train gen AI models focused on security, aviation, or military uses)

are important enough to be kept secret, and that other, less-favored ideas (e.g., information regarding data developers use to train gen AI models focused on creative writing) are not, in the state's eyes, valuable enough to be kept confidential. The Supreme Court has "emphatically rejected" the notion that a legislature may "weigh[] the value of a particular category of speech against its [perceived] social costs" and compel the disclosure of speech that it deems low value over the speaker's objection. *Brown v. Ent. Merchs. Ass'n*, 564 U.S. 786, 792 (2011).

113. Even if the above were not enough to confirm that AB 2013 triggers strict scrutiny, AB 2013's asserted purpose confirms as much. As the California legislature made clear, AB 2013 compels speech for the purpose of "identify[ing] and mitigate[ing] biases"—particular ideas and messages that the state disfavors. *Senate Floor Analysis*, *supra*, at 3. "Given the legislature's expressed statement of purpose, it is apparent that [AB 2013] imposes burdens that are based on the content of speech and that are aimed at a particular viewpoint," triggering strict scrutiny several times over. *Sorrell*, 564 U.S. at 565.

114. AB 2013 cannot survive any level of heightened scrutiny, let alone strict scrutiny. Under strict scrutiny, California must show that the statute "furthers a compelling governmental interest and is narrowly tailored to that end." *X Corp*, 116 F.4th at 903. But even under intermediate scrutiny California must show that AB 2013 is "narrowly tailored to serve a significant governmental interest." *McCullen v. Coakley*, 573 U.S. 464, 486 (2014). AB 2013 is plainly not narrowly tailored to advance any legitimate interest the state could assert. Its disclosure

obligations "are more extensive than necessary" to help "consumers … make informed decisions" about AI models. *X Corp.*, 116 F.4th at 903. Indeed, as explained, it is far from clear how the trade secrets AB 2013 would force xAI to disclose would be of any value to consumers at all. *See supra* ¶¶95-99. Reports from developers or outside certifiers who test and assess an AI model's effectiveness at performing key tasks are far more useful metrics of transparency about the effectiveness of a product. xAI already releases "Model Cards" for each of its AI models for exactly that reason. *See Grok 4 Fast Model Card*, *supra*. Those disclosures are appropriately focused on the AI model's outputs—the information actually relayed to users. By contrast, AB 2013 is directed at the model's inputs—i.e., the data used during development. An exhaustive list of the raw datasets (which can number in the hundreds or thousands) and data points (which far exceed even the datasets) used to develop and refine AI models does not give consumers any meaningful way to evaluate an AI model's effectiveness; all it does is enable competitors to replicate a highly valuable model's success. AB 2013 thus imposes onerous disclosure requirements while providing no meaningful benefit to consumers.

115. The fact that AB 2013 compels disclosures regarding all AI models released since 2022, even if they are no longer regularly used by consumers, underscores the disconnect between the law's obligations and the legislature's consumer-transparency goal.

116. In short, because AB 2013 imposes onerous disclosure obligations that

do nothing to advance its professed consumer-protection interests, it cannot survive any level of First Amendment scrutiny. For that reason too, AB 2013 should be declared unconstitutional, and the Attorney General should be enjoined from enforcing it.

**COUNT FOUR**
**(Unconstitutional Vagueness)**
**Due Process Clause of the Fourteenth Amendment to the U.S. Constitution**
**42 U.S.C. §1983; *Ex parte Young*; 28 U.S.C. §§2201(a) and 2202**

117. xAI incorporates by reference the allegations contained in the preceding paragraphs as though fully set forth herein.

118. AB 2013 is unconstitutionally vague both on its face and as applied to xAI.

119. "A fundamental principle in our legal system is that laws which regulate persons or entities must give fair notice of conduct that is forbidden or required." *FCC v. Fox TV Stations, Inc.*, 567 U.S. 239, 253 (2012). A law is unconstitutionally vague if it "fails to provide a person of ordinary intelligence fair notice of what is prohibited, or is so standardless that it authorizes or encourages seriously discriminatory enforcement." *United States v. Williams*, 553 U.S. 285, 304 (2008). "When speech is involved, rigorous adherence to those requirements is necessary to ensure that ambiguity does not chill protected speech." *Fox*, 567 U.S. at 253-54. After all, vague laws risk chilling would-be speakers by forcing them "to 'steer far wider of the unlawful zone'" than they would "if the boundaries of the forbidden areas were clearly marked." *Baggett v. Bullitt*, 377 U.S. 360, 372 (1964).

For that reason, laws touching on speech must themselves speak "only with narrow specificity." *NAACP v. Button*, 371 U.S. 415, 433 (1963).

120. AB 2013 is not written "with narrow specificity," and it fails to provide fair notice to a person of ordinary intelligence as to what it requires. To begin, it fails to define the key terms "dataset" or "data point." Do these refer to each individual set of data a developer might retrieve from a broad source (for example, particular websites like Creative Commons) and each particular byte of information available there? Or does it refer to a broad category of sources, like publicly available Internet data? Or does AB 2013 instead cover only the curated datasets a developer actually inputs into an AI model or all those datasets that a developer sourced before curating the particular datasets ultimately used for training? AB 2013 does not say.

121. The statute also requires developers to provide a "high-level summary" of its datasets but provides no guidance on *how much* information these summaries must disclose. For example, subsection (a)(2) requires a description of how the datasets further "the intended purpose of the artificial intelligence system or service." Cal. Civ. Code §3111(a)(2). Does this require xAI to disclose its internal strategies as to how it values each individual dataset? Or does it suffice to simply note that the dataset helps improve the AI model's effectiveness? Subsection (a)(5) suffers from the same flaw. That subsection requires developers to disclose "[w]hether the datasets include any data protected by copyright, trademark, or patent, or whether the datasets are entirely in the public domain." *Id*. §3111(a)(5).

Can xAI satisfy this with a simple "yes" or "no"? There is no way for xAI or any other developer to know whether its "high-level summar[ies]" must be 100 words or 100 pages. AB 2013 is thus far afield from regulating "with narrow specificity." *Button*, 371 U.S. at 433.

122. Making matters worse, AB 2013's disclosure obligation makes clear that even the enumerated list of information is not sufficiently comprehensive to satisfy the "high-level summary" mandate. *See* Cal. Civ. Code §3111(a) (explaining that the list must "includ[e], but [is] not limited to," that enumerated list). There is no way of knowing what additional information must be provided to fully comply with that obligation. That open-ended mandate plainly "invite[s] arbitrary enforcement" based on California's own assessment about whether enough information has been produced. *Kashem v. Barr*, 941 F.3d 358, 364 (9th Cir. 2019). Due process does not permit such standardless rules, especially when First Amendment rights are at stake.

123. Compounding the vagueness problem, AB 2013 is internally inconsistent as to what datasets it covers. In its operative provision, AB 2013 requires developers to disclose information "regarding the data used by the developer to train the generative artificial intelligence system or service," Cal. Civ. Code §3111, which the law defines as the "testing, validating, or fine tuning … of the artificial intelligence system or service," *id*. §3110(f). But the law later refers to "the datasets used in the development of the generative artificial intelligence." *id*. §3111(a). The datasets used to *train* an AI model are different from—and are far

fewer than—the datasets used to *develop* that AI model. But AB 2013 leaves developers guessing whether they must provide information only on training datasets, or whether they must disclose the broader universe of datasets that they might have sourced. Because the boundaries of AB 2013's disclosure requirements are not "clearly marked," *Baggett*, 377 U.S. at 372, xAI and other developers may feel compelled to disclose more information than AB 2013 requires, thereby exacerbating AB 2013's infringement on free speech—and its evisceration of xAI's trade secrets, to boot.

124. Finally, AB 2013 fails to provide fair notice as to which AI systems it covers. The law covers "a generative artificial intelligence system or service … made publicly available to Californians for use." Cal. Civ. Code §3111. But what constitutes being "made publicly available to Californians for use"? If xAI licenses an AI model it has developed to an online retailer, who in turn uses that model to provide consumers with an automated chatbot that incorporates generative AI, must xAI disclose information about the datasets used to train that AI model, even though it is only privately licensed? Does any company that incorporates and optimizes someone else's generative AI system for its own business have to disclose details on how it did so? AB 2013 does not say.

125. In short, AB 2013 is plagued with vagueness problems on everything from which AI systems it covers, to what datasets it covers, to the level of detail developers must provide to comply with its mandates. That lack of "narrow specificity" will force companies to over-disclose confidential information to ensure

compliance—a result that is exceptionally problematic in the First Amendment and trade-secret context. AB 2013 should therefore be declared unconstitutionally vague, and the Attorney General should be enjoined from enforcing it.

**RELIEF REQUESTED**

For the foregoing reasons, xAI respectfully requests from the Court:

a. A declaration, pursuant to 28 U.S.C. §§2201(a) and 2202, that the provisions of AB 2013 that are codified at Cal. Civ. Code §3111(a)(1)-(9), (12) effect an uncompensated taking of trade secrets owned by xAI in violation of the Takings Clause of the U.S. Constitution;

b. A declaration, pursuant to 28 U.S.C. §§2201(a) and 2202, that AB 2013 unconstitutionally compels xAI's speech in violation of the First Amendment of the U.S. Constitution;

c. A declaration, pursuant to 28 U.S.C. §§2201(a) and 2202, that AB 2013 is unconstitutionally vague in violation of the Due Process Clause of the U.S. Constitution;

d. An order preliminarily enjoining Attorney General Bonta, as well as all officers, agents, and employees subject to his supervision, direction, or control, from enforcing the provisions of AB 2013 against xAI;

e. An order permanently enjoining Attorney General Bonta, as well as all officers, agents, and employees subject to his supervision, direction, or control, from enforcing the provisions of AB 2013 against xAI;

f. Award xAI its costs and reasonable attorney's fees incurred in this action pursuant to 42 U.S.C. §1988 and other applicable law; and

g. Grant xAI all other such relief as the Court may deem just and proper.

Respectfully submitted,

s/Matthew D. Rowen
ERIN E. MURPHY (*pro hac vice* forthcoming)
MATTHEW D. ROWEN (Cal. Bar #292292)
JAMES Y. XI (*pro hac vice* forthcoming)
MITCHELL K. PALLAKI (*pro hac vice* forthcoming)
ILAN J. POSNER (*pro hac vice* forthcoming)
CLEMENT & MURPHY, PLLC
706 Duke Street
Alexandria, VA 22314
Tel.: (202) 742-8900
erin.murphy@clementmurphy.com
matthew.rowen@clementmurphy.com
james.xi@clementmurphy.com
mitchell.pallaki@clementmurphy.com
ilan.posner@clementmurphy.com

*Attorneys for Plaintiff X.AI LLC*

December 29, 2025