Proskauer

Data License Restrictions in the AI Spotlight: Careful Drafting Is More Important Than Ever

New Media and Technology Law Blog on **December 15, 2025**

A recently-filed federal court complaint tests the enforceability of restrictive terms in a data license against the use of licensed data for generative AI purposes. The outcome of this case may turn on interpreting broad terms such as training, internal research, distribution and publication.

On November 26, 2025, legal research platform Fastcase, Inc. ("Fastcase") sued legal AI company Alexi Technologies Inc. ("Alexi") in a District of Columbia district court, alleging that Alexi used licensed "Fastcase Data" (i.e., a sophisticated, extensively tagged legal research database) to train and power a commercial AI legal research product in violation of a 2021 data license agreement.(*Fastcase, Inc. v. Alexi Technologies Inc.*, No. 25-04159 (D.D.C. Filed Nov. 26, 2025)).

The complaint is particularly interesting as it involves a 2021 agreement – entered into before generative AI was a generally known technology – and thus presents the age-old question of interpreting the terms of an agreement to a technology not known at the time the agreement was entered into.

Regardless of the outcome of this case, the complaint highlights the importance of precise license drafting – for both licensors and licensees – in the age of AI. Parties should focus on well-defined understandings of how data can be used, shared and commercialized at a time where every company is seeking to leverage AI. As this complaint illustrates, vague terms can lead to uncertainty and legal disputes.[1]

Of course, we do not know if this litigation will proceed to summary judgment or trial, but if it does, we may have an opportunity to see how a court will interpret and potentially enforce broad and general restrictions in a data license (e.g., no publication, distribution, commercial use, competitive use, use for internal research purposes only) in the context of generative artificial intelligence.[2]

Case Background

Fastcase built a significant online proprietary legal research database with extensive text and metadata tagging to power advanced research tools, a process that Fastcase claims took decades and millions of dollars in investment. According to the complaint, Alexi is a legal services company that formerly operated a service that used a passage-retrieval AI system and human lawyer review to prepare legal memoranda for clients.

In 2021 – well before OpenAI's launch of ChatGPT in November 2022 brought generative AI into the mainstream – the parties entered into a Data License Agreement (the "License") that granted Alexi limited access to Fastcase's legal research database restricted for "internal research purposes."[3] In addition, the License stated: "In no event can the data be used for commercial purposes or for any purpose which is competitive with Fastcase." It further prohibited use "in a way which is inconsistent with the Purpose," and stated the licensee would not "sell, license, publish, copy, or otherwise distribute any part of the Fastcase Data."

Fastcase alleges that Alexi later pivoted its service model to a public, AI-legal research platform, purportedly in competition with Fastcase, and used the licensed database to train and operate its commercial generative AI legal-research product and provide end users with direct access to Fastcase-sourced case law, all in violation of the License restrictions.

The complaint states that Alexi did not give Fastcase notice of its pivot or seek to renegotiate the License to authorize its new uses. In correspondence between counsel prior to the complaint being filed, Alexi purportedly denied any breach or wrongdoing and refused to cease its alleged unauthorized use of Fastcase data. In its reply to Fastcase, Alexi also rejected Fastcase's narrow interpretation of the License ("[T]he intention of the Agreement was never to preclude Alexi from using Fastcase Data as source material for Alexi's generative AI product…"). Fastcase then terminated the License and filed suit, advancing various claims including breach of contract, trade secret misappropriation[4] and Lanham Act claims.[5]

In response, on December 10, 2025, Alexi filed a motion for a temporary restraining order (TRO) and preliminary injunction that would return the parties to the status quo prior to the filing of the lawsuit. The supporting memo of law is redacted, but the relief sought generally seeks to maintain the parties' relationship and Alexi's licensed access to the Fastcase data while the suit is pending. In its supporting memo, Alexi contends that "Fastcase's claims are flawed as a matter of fact and law" and "Alexi's use of the Data to generate legal research memos, using its AI models, is consistent with the terms of the Agreement as well as the parties' course of conduct."

As evidenced by the unredacted portions of Alexi's TRO motion, some of the key questions in the case appear to be whether Alexi's AI training was arguably within the scope of the License's definition of "internal research purposes" and whether Alexi's alleged display of Fastcase data to end users constitutes a "publication" in violation of the License. In addition, given that the License pre-dated the general awareness of generative AI, it is an open question whether the License should be construed narrowly to exclude uses that were not known or contemplated at the time of the agreement.

Lessons

We will watch this litigation closely as it may illuminate many important issues in data licensing in an ever-evolving environment of AI innovation and development. For now, the seed of this dispute recalls many important lessons for both licensors and licensees in negotiating data licenses. These include:

Defining the scope of the license:

- **Enumerating permitted uses**: A license should contain, to the extent possible, separately identified (and, in some cases, defined) concepts (rather than relying on general terms) that contemplate specific AI uses that are potentially included or excluded. Examples include: retrieval augmented generation (RAG), fine-tuning, train, pre-training, inference serving, model, weights, derived data, prompts, output, AI browsers, model context protocol (or MCP), etc. Additional drafting considerations might be made for other autonomous agentic AI uses, such as how data may be used to power agents and whether data may enter long-term or cross-agent shared memory, to name a few issues.
    - **Distinguishing between generative AI and machine learning**: The parties may need to separate two often-confused but very different categories of data use: (1) generative uses that can create or transform

content; and (2) machine-learning or analytic uses that only detect patterns or train classifiers, without enabling reconstruction of the original data. The intent to include or exclude machine learning should be taken into account to ensure that the scope of the license does not, inadvertently or intentionally, go too far.

- **Derivative works**: Often, licensors prefer derived works to remain the licensor's property or remain lightly controlled, with licensees obtaining only a limited internal-use or aggregate license to them. However, in the context of generative AI, the actual desired use case is, in most cases, to create derivative works from the licensed data. Thus, the parties must be careful and precise in drafting derivative work grants and restrictions.

- **What constitutes a "publication" or "distribution" in a user interface?** The Fastcase complaint alleges that Alexi's act of displaying full-text case law sourced from the licensed database to end users was "publishing" in violation of the License. There is additional complexity as, in this case, the "data" at issue is public domain material. This is another example of emerging technology pushing the boundaries of what may have been once well-understood concepts in licensing and intellectual property law. Careful drafting is necessary to ensure the parties' intentions are effectively expressed in the agreement.
- **Address "future technologies"**: In light of evolving technology, contract terms can be unclear, undefined or overbroad in their technological scope (which, perhaps, might be an issue in the Fastcase litigation). Many licenses include a forward-looking express clause addressing whether the use of data in future technologies is permitted or excluded (e.g., using the phrase "technologies whether now known or later developed"). If that type of phrase is used, care must be taken so that it modifies the appropriate grants (e.g., including that phrase after a list of grant of rights could be construed to limit the right granted immediately before the use of that phrase, as opposed to the entire string of granted rights).

  This issue is particularly important in the extremely fast moving world of generative AI where new innovations seem to be launched on a daily basis. So, for example, parties might attempt to ensure their rights remain intact as new AI, agentic systems and other technologies emerge and might seek to define granted or restricted uses in broad, functional terms to demonstrate an intent to cover future technologies. Needless to say, this issue is not a new one in technology contracting, and in addition to writing about the issue from time to time, we have had the opportunity to litigate the issue in the context of NFTs.
- **Understand how different legal issues may be a matter of copyright vs. contract**: In structuring an agreement, both parties should understand the

interplay between contract law and copyright law. Is it possible that aspects of the agreement are subject to copyright preemption? Are data use restrictions enforceable against AI "training" if the underlying data is in the public domain? The Fastcase complaint, for example, frames its claims as a breach of contract and trade secret (the proprietary compilation and organization of data) but avoids copyright claims as the content at issue is not subject to copyright protection.

We expect that there will be many more licensing disputes in the context of the use of data in AI applications. The old forms are likely silent on the issue, so both licensors and licensees should approach data and related agreements with AI-related uses well in mind.

---

[1] One of the plaintiff's claims involves an assertion that its legal research database is a compilation trade secret based on the structure, selection and organization of court decisions and legal sources. While this presents an interesting question of law, a discussion of that issue is beyond the scope of this post.

[2] Of course, data licenses are not the only contractual relationships that generate data that may be useful in generative AI applications. Parties to all types of agreements should keep these concepts in mind as ancillary-related data can be valuable in the context of AI platforms and models.

[3] The License itself was appended to the complaint, but as of the date of this post, the exhibit is under seal.

[4] As to the trade secret claims, Fastcase claims that the database "is not generally known, is not readily ascertainable by proper means, and derives independent economic value specifically from its confidential structure, organization, and metadata—the very elements Alexi exploited to accelerate its competing commercial platform." Complaint, at ¶116.

[5] The Lanham Act claims, which involve alleged use of Fastcase marks or branding on AI output, are beyond the scope of this post.

View original.

Related Professionals

- **Jeffrey D. Neuburger**