Artificial Intelligence: Intellectual Property Litigation Discovery Considerations: Overview

BRIAN WM. HIGGINS, BLANK ROME, LLP, WITH PRACTICAL LAW INTELLECTUAL PROPERTY & TECHNOLOGY

Search the Resource ID numbers in blue on Westlaw for more.

A Practice Note providing an overview of key issues companies and their in-house and outside counsel should consider when preparing for discovery in intellectual property litigation in which an artificial intelligence (AI) technology is at issue. The Note discusses potential claims and defenses, pre-suit considerations, written and third-party discovery, source code review, and depositions in a hypothetical patent infringement and trade secret misappropriation litigation matter involving highly autonomous vehicle (HAV) facial recognition technology.

As more patent, trade secret, and other intellectual property (IP)-related cases involving artificial intelligence (AI) make their way through US courts, lawyers and the judges overseeing those cases need to consider highly-nuanced discovery and other issues specific to AI technologies. Like other software IP cases, AI technology lawsuits present black box challenges, but they also differ from traditional software matters in that AI systems can "learn" as they receive new input data. For example, in the hypothetical facial recognition authentication system discussed below, the system's ability to recognize faces "in the wild" and authenticate users improves as it observes more faces. Other AI systems may make decisions or take actions that are not always foreseeable and may not be explainable by their developers, presenting additional challenges for litigants, their lawyers, and fact finders.

A well-implemented discovery strategy, aimed at gaining insight into the nature of an AI model, can help litigants understand **how** a disputed AI system works, which is crucial to counsel seeking to:

- Present an accurate and compelling story to a judge or jury.
- Mount a technically-accurate defense for a client.
- Negotiate a settlement agreement armed with all the facts.

This Note provides an overview of key discovery issues and practical strategies counsel can use before and during an IP lawsuit involving an AI technology, using hypothetical facial recognition technology in a highly autonomous vehicle (HAV) as a guide, including:

- Background on the factual elements required to prove patent infringement and trade secret claims.
- Pre-suit considerations including:
  - important investigative first steps; and
  - litigation and preservation holds.
- The potential for filing or facing a motion to dismiss.
- Key AI-related discovery issues, including:
  - initial disclosures;
  - written discovery;
  - source code reviews;
  - third-party discovery; and
  - fact and expert witness depositions.

For an overview of legal issues counsel may face concerning AI-related technology, see Practice Note, Artificial Intelligence Key Legal Issues: Overview (w-018-1743).

For a collection of resources concerning legal issues surrounding AI, see Artificial Intelligence Toolkit (w-019-1426).

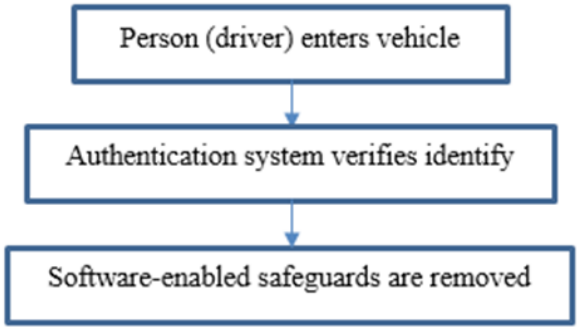
For a collection of resources covering all aspects of patent litigation discovery, see Patent Litigation Discovery Toolkit (8-585-6866).

HYPOTHETICAL AI TECHNOLOGY LAWSUIT: HAV FACIAL RECOGNITION SYSTEM

Company XYZ has patented a new facial recognition system for authenticating HAV users, which uses an in-cabin camera and software to:

- Interrogate passenger faces when they enter the vehicle.
- Authenticate the person who takes the driving position in the vehicle.
- Disengage safeguards allowing the driver to operate the vehicle.

Schematically, the authentication process might look like the following:



XYZ has bet its future on the future of the HAV market, foregoing the existing semi-autonomous passenger vehicle market.

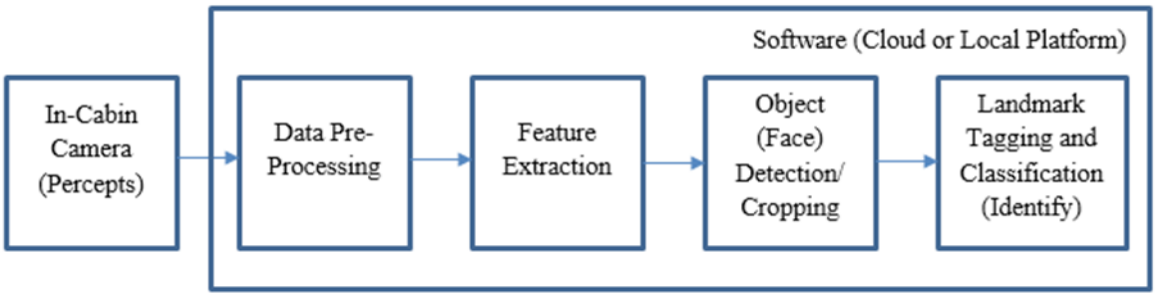
Company ABC has taken a different approach in the market by selling its facial recognition authentication system to vehicle original equipment manufacturers (OEMs) and Tier 1 suppliers for integration into semi-autonomous vehicles. Instead of patenting its system, ABC has chosen to maintain its software, which it calls IVIEWDRIVE®, and certain of its algorithms, as trade secrets.

After XYZ hires some of ABC's machine learning engineers, ABC sues XYZ for misappropriation of its IVIEWDRIVE® software and algorithms in violation of the Defend Trade Secrets Act (DTSA), 18 U.S.C. § 1836 et seq., and analogous state trade secret law, requiring ABC to establish the existence and misappropriation of a trade secret. XYZ denies the allegations and counters with claims for patent infringement under 35 U.S.C. § 271(a), requiring XYZ to establish that ABC made, sold, offered for sale, or used XYZ's patented facial recognition authentication system.

This Note focuses on patent infringement and trade secret misappropriation, although the parties could potentially assert other legal claims, including copyright infringement of source code and non-IP-related claims, such as unfair competition, inducement, and tortious interference.

THE PARTIES' SYSTEMS

The figure below summarizes how the parties' systems generally operate:



As shown above, the system:

- Extracts data from the camera's output signal (a frame-by-frame analysis of streaming video data).
- Locates faces in the image data (which is an object detection process typically done using a machine learning algorithm called a convolutional neural network).
- Identifies unique features (for example, facial landmarks) from the faces that can be used to differentiate them.
- Compares those features to the faces of people already known to identify a match.

Identification is possible if stored (known) facial data are matched to incoming feature data. This facial data comes from previously-provided template face scans, which may be stored locally in the vehicle or possibly in an app carried by the driver into the vehicle and transmitted to the local computing platform.

ALGORITHMS AND MODELS

As used in the hypothetical and throughout this Note, an algorithm refers to a series of steps, written in a programming language, such as Python, performed by a computer that when executed perform a specific function. An AI model refers to a set of algorithms used together to replicate a real-world process. For example, the above convolutional neural network used for object detection may be considered an algorithm (or part of an algorithm) and the software aspects of the facial recognition authentication system as a whole could be considered a model (it replicates the human cognitive process of identifying a person by recognizing their facial features).

PRE-SUIT INVESTIGATION CONSIDERATIONS

Counsel should perform a thorough pre-suit investigation before filing a complaint or responsive pleading that includes defenses and counterclaims. Counsel should ensure that all pleadings (and all other court filings) comply with Federal Rule of Civil Procedure (FRCP) 11 and relevant local court rules and case law before filing.

APPLICABLE LAW AND RULES

FRCP 11 sets the standard for all representations made to the court, including those in pleadings. Specifically, FRCP 11(b)(3) requires that, before filing, counsel perform an "inquiry reasonable under the circumstances," including ensuring any factual contentions have evidentiary support or, if specifically so identified, will likely have evidentiary support after a reasonable opportunity for further investigation or discovery. Courts' local civil rules also often include a FRCP 11 analogue, such as the Eastern District of Texas's Local Rule 11(a), which requires designation of a lead attorney deemed responsible for the party's actions. In some respects, these rules impose on counsel a duty to perform some adequate amount of pre-suit investigation before filing suit.

Counsel or a party failing to comply with Rule 11(b) may face sanctions under Rule 11(c). In patent cases, a court may also award attorneys' fees under 35 U.S.C. § 285 if it finds the plaintiff's pre-suit investigation to be inadequate (see *Lumen View Tech. LLC v. Findthebest.com, Inc.*, 811 F.3d 479, 483 (Fed. Cir. 2016) (citing *Octane Fitness LLC v. ICON Health & Fitness, Inc.*, 134 S. Ct. 1749, 1757 ("[A] district court may award fees in the rare case in which a party's unreasonable conduct—while not necessarily independently sanctionable—is nonetheless so 'exceptional' as to justify an award of fees.")).

In trade secret cases, some state laws authorize courts to award attorneys' fees if a plaintiff brings a trade secret claim in bad faith (see, for example, *Microstrategy, Inc. v. Bus. Objects*, 331 F. Supp. 2d 396, 430 (E.D. Va. 2004) (considering Virginia's uniform trade secret act, Va. Code. Ann. 59.1-338.1)). Counsel should therefore carefully consider the applicable local rules and case law when preparing to file suit or a responsive pleading.

HYPOTHETICAL LAWSUIT PRE-SUIT INVESTIGATION

Although the adequacy and reasonableness of a pre-suit investigation in a typical IP-related lawsuit is assessed on a case-by-case basis, counsel in a lawsuit involving an AI technology should consider the pre-suit investigation a critical part of developing and drafting claims, defenses, and counterclaims in view of the complexity, opacity, and learned nature of AI technologies. In the HAV facial recognition authentication hypothetical, counsel's pre-suit investigation should include at least the following:

- Before filing its complaint for trade secret misappropriation, ABC's counsel should:
  - evaluate whether its client has taken reasonable measures to keep its software and unique algorithms secret (18 U.S.C. § 1839(3));
  - consult with an economic expert to assess whether ABC's software and algorithms derive 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 (18 U.S.C. § 1839(3)); and
  - conduct a forensic analysis of ABC's trade secret information to understand the circumstances surrounding its possible theft or disclosure by the former engineers who now work at XYZ, including an in-depth forensic analysis of the employees' email, desktop and laptop computers, handheld electronic devices, and office files.
- Before filing its answer to ABC's complaint, XYZ's counsel should:
  - determine the circumstances surrounding its hiring of ABC's former employees, including evaluating software configuration changes commensurate with those hirings; and
  - assess whether the former ABC employees "knew or had reason to know that the trade secret was acquired through improper means, under circumstances giving rise to a duty to maintain the secrecy of the trade secret" (*Ford Motor Co. v. Launch Tech Co.*, 2018 WL 1089276, at *16-17 (E.D. Mich. Feb. 26, 2018) (citing 18 U.S.C. § 1839(5)).
- Before filing its counterclaim for patent infringement, XYZ's counsel should:
  - engage a technical expert to help test ABC's AI-based facial recognition authentication system in a vehicle (or obtained as a separate system, if available); and
  - determine if the system's operation can shed light on how the AI system works and if it can be reverse engineered.
- Before answering XYZ's patent infringement claims, ABC's counsel should:
  - review the known facts and key documents concerning its client's AI system; and
  - compare them to the asserted patents, prosecution histories, and prior art to assess the strength of XYZ's patent claims.
- Counsel for both parties should engage independent consultants to help them identify evidentiary support for any factual contention asserting in their respective pleadings.

For more on pre-suit considerations generally, see Practice Notes, Patent Litigation Pre-Suit Considerations (6-534-0765) and Trade Secrets Litigation: Preliminary Steps (5-523-8283).

LITIGATION HOLD CONSIDERATIONS

In addition to a pre-suit investigation, counsel must advise their clients about taking steps to avoid spoliation of evidence as soon as the client either:

- Is aware of the possibility of a lawsuit or legal action against it.
- Intends to initiate a lawsuit or take legal action.

A litigation hold (or preservation hold) generally applies to electronic documents, but it also implicates hardcopy documents, equipment, and other things. In any IP lawsuit involving AI technology, a prompt litigation or preservation hold is critical because these cases typically involve source code versions and data that may be routinely updated or modified.

For a collection of resources to help counsel implement a litigation hold, see Litigation Hold Toolkit (2-545-9105).

APPLICABLE LAW AND RULES

FRCP 37(e) provides for discretionary sanctions, including dismissal, if "electronically stored information that should have been preserved in the anticipation or conduct of litigation is lost because a party failed to take reasonable steps to preserve it, and it cannot be restored or replaced through additional discovery." A party's duty to preserve evidence "comes into being when the party has notice that the evidence is relevant to the litigation or should have known that the evidence may be relevant." (*Guzman v. Jones*, 804 F. 3d 707, 713 (5th Cir. 2015)). "Spoliation designed to deprive an adversary of the use of evidence in litigation qualifies as bad faith conduct" and may be sanctionable (*see CAT3, LLC v. Black Lineage, Inc.*, 164 F. Supp. 3d 488, 501 (S.D.N.Y 2016)). Depending on the jurisdiction, a challenging party may need to demonstrate clear and convincing evidence of spoliation (*CAT3, LLC,* 164 F. Supp. 3d at 499).

HYPOTHETICAL LAWSUIT LITIGATION HOLD

As plaintiff, ABC's counsel should advise ABC to initiate a preservation hold as soon as it begins evaluating the possibility of taking legal action against XYZ for trade secret misappropriation. ABC's counsel should also consider serving a litigation hold letter on its former engineers who joined XYZ, in case the company believes taking action against them individually is warranted.

If ABC and XYZ are in discussions about the merits of ABC's trade secret allegations before suit, XYZ's counsel should advise its client to initiate a preservation hold even before receiving service of ABC's complaint.

In particular, regarding the disputed AI technology, counsel for both parties should advise their respective clients to:

- Freeze training and testing data sets that were and are being used for AI model development during the period of time at issue in the dispute.
- Identify and preserve any new data points that have been added to training data sets (where data are added, for example, to re-train AI algorithms to improve models).
- Freeze all source code versions created during the disputed time period.

ABC's counsel should similarly implement a preservation hold concerning XYZ's patent infringement counterclaims as soon as it anticipates the counterclaims or promptly after receiving service of XYZ's answer. For a model patent litigation hold notice, see Standard Document, Patent Litigation: Litigation Hold Notice (Accused Infringer) (w-018-0210).

For more on implementing a litigation hold generally, see Practice Note, Implementing a Litigation Hold (8-502-9481).

MOTION TO DISMISS CONSIDERATIONS

At the early stages of a lawsuit, the parties' pre-suit investigation and analysis of the opposing party's pleadings may help counsel assess the merits of and likelihood of prevailing on a motion to dismiss. FRCP 12, local court rules, and relevant case law provide the procedure for and standards by which courts address motions to dismiss.

APPLICABLE LAW AND RULES

FRCP 12 sets out the procedures for moving to dismiss a party's complaint, either before or after the defendant has filed a responsive pleading. FRCP 12(b)(6) is often the basis for a motion **before** the close of pleadings and requires the movant to establish that respondent (typically, the plaintiff) failed to state a claim on which relief can be granted. FRCP 12(c) is used **after** the pleadings are closed and requires the movant to demonstrate that it is entitled to judgment on the facts and evidence set out in the pleadings.

Defendants in patent litigation involving patents directed to computer-implemented technology, including software used in AI algorithms, often move to dismiss on the basis that the asserted patent claims are abstract and therefore not eligible for patent protection under 35 U.S.C. § 101 (see *Alice Corp. Pty Ltd. v. CLS Bank Int'l*, 134 S. Ct. 2347, 2354-55 (2014) (providing a framework for assessing abstract concepts under § 101)). Under *Alice*, a patent claim is ineligible under Section 101 when it is "directed to" one of three general patent-ineligible categories, including "laws of nature, natural phenomena, and abstract ideas." (*Alice*, 134 S. Ct. at 2354). If so, *Alice* requires consideration of whether the particular elements of the claim, evaluated "both individually and 'as an ordered combination,'" add enough to "transform the nature of the claim into a patent-eligible invention" (*Elec. Power Grp., LLC v. Alstom S.A.*, 830 F.3d 1350, 1353 (Fed. Cir. 2016) (quoting *Alice*, 134 S. Ct. at 2355)). For more information on patent-eligible subject matter, see Section 101 Patent Eligibility Toolkit (w-010-2764).

In the trade secret misappropriation context, a defendant may move to dismiss on the basis that the trade secret holder has not described the subject matter of its trade secrets with sufficient particularity "to separate it from matters of general knowledge in the trade or of special persons who are skilled in the trade, and to permit the defendant to ascertain at least the boundaries within which the secret lies." (*Pellerin v. Honeywell Int'l, Inc.*, 877 F. Supp. 2d 983, 988 (S.D. Cal. 2012) (quoting *Diodes, Inc. v. Franzen*, 260 Cal. App. 2d 244, 253 (1968))). However, while some degree of particularity is required, at the pleading stage "plaintiffs can describe trade secret information in general terms" so as not to publicize the secret itself (*see Covenant Aviation Security, LLC v. Berry*, 15 F. Supp. 3d 813, 818 (N.D. Ill. 2014)). For more on trade secret litigation claims and defenses, see Practice Note, Trade Secrets Litigation (5-523-8283).

HYPOTHETICAL LAWSUIT MOTION TO DISMISS

In response to XYZ's patent infringement allegations, ABC's counsel may consider moving to dismiss under FRCP 12(b)(6) based on arguments that accused infringers have used in successful *Alice*-type Section 101 challenges, including that XYZ's asserted patent claims:

- Are "directed to a mental process" performed by a computer.
- Are directed to an "abstract concept of using mathematical algorithms to perform predictive analytics" by collecting and analyzing information.
- Are "result-focused" and "so functional, as to effectively cover any solution to an identified problem" (that is, detecting a face of a person inside a vehicle).
- Lack specificity necessary to show how the claimed computer processor's operations differ from prior human methods, and thus

are not directed to a technological improvement but rather are directed to an abstract idea.

- Are directed to unpatentable "mathematical formulas and algorithms, divorced from a particular solution to a technological problem."

(*See Alice*, 134 S. Ct. at 2358 (citations omitted).)

In a DTSA or state trade secret matter, plaintiff's pleading must avoid conclusory and generalized allegations of trade secret misappropriation, including vague timelines of alleged acts of misappropriation and uses of purported trade secrets. XYZ's counsel therefore should consider moving to dismiss where ABC's DTSA or state trade secret misappropriation claims:

- Do not set out, with sufficient particularity, the metes and bounds of the AI-based facial recognition system software that it contends was misappropriated, instead referring only to the software by a trademark name generally (IVIEWDRIVE) and the software's general functionality.
- Do not set out, with sufficient particularity, the specific algorithms that it contends are being used in XYZ's AI-based system, such as the type and characteristics of those algorithms (for example, a specific type of neural network).
- Are factually baseless regarding XYZ's alleged use of ABC's trade secret algorithms because XYZ's algorithms for performing facial recognition and authentication are different than ABC's alleged trade secret algorithms.

In the hypothetical lawsuit, ABC's counsel may also allege future misappropriation based on the "inevitable disclosure doctrine" regarding its former engineers. In response, XYZ's counsel may move to dismiss on several grounds. For more on the inevitable disclosure doctrine, see Practice Note, Trade Secrets Litigation: Common Causes of Action: Inevitable Disclosure of Trade Secrets (5-523-8283).

INITIAL DISCLOSURES CONSIDERATIONS

If the lawsuit survives early motions to dismiss, the parties will need to exchange their initial disclosures under FRCP 26, local court rules, and relevant case law, which provide the scope, content, and procedures for making initial disclosures. In a lawsuit involving an AI technology, counsel should pay close attention to the possibility of having to make its source code available for inspection early in the case.

APPLICABLE LAW AND RULES

FRCP 26(a)(1) requires a party, without awaiting a discovery request, to provide to the other party:

- Preliminary identifications of individuals likely to have discoverable information.
- Types and locations of documents.
- Other information that a party in good faith believes may be relevant to a case, based on each parties' claims, counterclaims, facts, and various demands for relief set out in their respective pleadings.

(See *ADC Ltd. NM, Inc. v. Jamis Software Corp.*, slip op. No. 18-cv-862 (D.N.M. Nov. 5, 2018) (referring to initial disclosures as an "exchange of core information about [a party's] case.").)

A party failing to comply with initial disclosure rules "is not allowed to use" the information or person that was not disclosed "on a motion, at a hearing, or at a trial, unless the failure was substantially justified or is harmless." (*Baker Hughes Inc. v. S&S Chem., LLC*, 836 F. 3d 554 (6th Cir. 2016) (citing FRCP 37(c)(1))).

Courts' local civil rules also often include a FRCP 26(a) analogue, such as Delaware's default initial disclosures in patent infringement matters, which requires additional disclosures of accused products and supporting "core" documentation by specific deadlines and a default standard for analysis and accessibility of source code, available on the court's website (*see also Drone Techs., Inc. v. Parrot SA*, 838 F. 3d 1283, 1295 (Fed. Cir. 2016) (citing US District Court for the Western District of Pennsylvania local rule LPR3.1, requiring, in patent cases, initial disclosure of source code and other documentation... **sufficient to show** the operation of any aspects or elements of each accused apparatus, product, device, process, method or other instrumentality **identified in the claims pled** of the party asserting patent infringement...")) (emphasis added)).

HYPOTHETICAL LAWSUIT INITIAL DISCLOSURE OF PERSONS

FRCP 26(a)(i) requires the identification of individuals likely to have discoverable information, along with the subjects of that information, that the disclosing party may use to support its claims or defenses, unless the use would be solely for impeachment. In the hypothetical AI-based facial recognition authentication lawsuit, these individuals may include:

- Data scientists and engineers.
- Machine learning engineers.
- Software engineers.
- Stack engineers and systems architects.
- Hired consultants (including those employed by a third party).

In particular, the parties should identify:

- Data scientists, data engineers, and machine learning engineers involved in selecting and processing data sets or training, validating, or testing the algorithms at issue in the lawsuit.
- Data scientists involved in developing the final deployed AI model.
- Software engineers involved in writing the machine learning algorithm code, especially any who can explain how parameters and hyperparameters were selected or derived and if any relevant technical standards or measures of accuracy were used.
- Stack engineers and systems architects who know how the hardware and software features of the contested AI systems were put together.
- Task and project managers and other higher-level scientists and engineers involved with the disputed AI system.
- Any other person who has knowledge of any cloud-based or edge computing features of the disputed AI system. For example, if a portion of the party's system is operable on a machine learning as a service (MLaaS) platform, that party may need to disclose the cloud-based platform service where a party's application program interface (API) requests are piped for processing along with a contact person.

HYPOTHETICAL LAWSUIT KEY INITIAL DOCUMENT DISCLOSURES

FRCP 26(a)(ii) requires production of a copy or a description by category and location of all documents, electronically stored information (ESI), and tangible things that the disclosing party has in its possession, custody, or control and may use to support its claims or defenses, unless the use would be solely for impeachment. Depending on the jurisdiction in which the hypothetical lawsuit is pending, a party's initial "core document" disclosure burden may involve identifying or making available:

- By XYZ:
  - documents related to the inventions disclosed in the patents-in-suit, including file histories;
  - prior art;
  - reduction to practice and original invention disclosures (if relevant);
  - documents related to the hiring of ABC's former machine learning engineers;
  - source code (by inspection only) and related software development documentation (see Source Code Review Considerations); and
  - damages-related information including existing licenses, price, units sold, revenues, and profits.
- By ABC:
  - source code (also by inspection only) and related software development documentation;
  - software access control logs;
  - documents evincing alleged theft of software; and
  - damages-related information.

DOCUMENT PRODUCTION CONSIDERATIONS

Once fact discovery begins, counsel may exchange requests for the production of documents, interrogatories, and requests for admission.

APPLICABLE LAW AND RULES

FRCP 26(b) governs the procedures and standards by which parties in civil litigation may obtain discovery regarding any non-privileged matter that is relevant to any party's claim or defense and that is proportional to the needs of the case, unless limited by a court, taking into consideration the following factors expressed in Rule 26(b):

- The importance of the issues at stake in the action.
- The amount in controversy.
- The parties' relative access to relevant information.
- The parties' resources.
- The importance of the discovery in resolving the issues.
- Whether the burden or expense of the proposed discovery outweighs its likely benefit.

Evidence is relevant to a party's claim or defense if it tends "to make the existence of any fact that is of consequence to the determination of the action more or less probable that it would be without the evidence" (Fed. R. Evid. 401). Even if the information sought in discovery is relevant and proportional, discovery is not permitted where no need for the information has been shown (*see Standard Inc. v. Pfizer Inc.,* 828 F.2d 734, 743 (Fed. Cir. 1987)). Parties have an obligation to supplement their responses to document requests as relevant information is identified during the lawsuit.

HYPOTHETICAL LAWSUIT DOCUMENT REQUESTS

In lawsuits involving an AI technology, knowing how the AI system makes a decision and takes an action may be highly relevant to a party's case. Assuming counsel for XYZ and ABC can justify a need for the information under FRCP 26(b), they should serve targeted discovery requests seeking the following information regarding the disputed AI systems:

- Data sets considered and used (raw and processed).
- Software, including earlier and later versions of the contested version.
- Software development processes used.
- Sensors for collecting real-time observational data for use by the AI model.
- Source code files.
- Specifications describing what a system is and is supposed to do.
- Schematics illustrating how components of a system work together.
- Flow charts describing how percepts (data) are inputted, processed, and used.
- Formulas, including those forming the basis for algorithms written in code.
- Drawings, especially if they help elucidate how a system works.
- Other AI system documentation.

INTERROGATORIES CONSIDERATIONS

Interrogatories are useful for a party to develop facts, discover the identity of additional witnesses, and identify the location of relevant documents (*see Alston v. Sharpe*, 2015 WL 6395937, at *2 (D. Conn. Oct. 22, 2015); *Harris v. Escamilla*, 2016 WL 1224057, at *6 (E.D. Cal. Mar. 29, 2016) ("the purpose of interrogatories is to expedite trial by narrowing and clarifying the issues and identifying potential witnesses or documents.") (citations omitted)). In an AI lawsuit, counsel should try to use at least some interrogatories to focus on the data, algorithms, model development, and model deployment (software and hardware) necessary to understanding the disputed AI system.

APPLICABLE LAW AND RULES

FRCP 33 governs interrogatory practice and specifies numerical limits (typically 25, including discrete subparts; more may be sought by leave of court or agreed upon by the parties) and scope (they must relate to a matter that may be inquired into under Rule 26(b)). FRCP 33 also governs how a party answers and raises objections, how much time is permitted to respond (30 days by default), their use as evidence, and optional production of business records if the answer to an interrogatory may be determined by examining, auditing, compiling, abstracting, or summarizing a party's business records (including electronically stored information) and if the burden of deriving or ascertaining the answer will be substantially the same for either party. Parties have an obligation to supplement their interrogatory responses as relevant information is identified during the lawsuit.

FRCP 37(d) gives federal courts discretion to sanction a party if the "party, after being properly served with interrogatories under FRCP 33... fails to serve its answers, objections, or written response." (Fed. R. Civ. P. 37(d)(1)(A)(ii)).

HYPOTHETICAL LAWSUIT INTERROGATORIES

XYZ's counsel should consider serving interrogatories seeking the identification of:

- Documents and information describing ABC's facial recognition authentication system so that XYZ may better evaluate its infringement contentions and damages calculations.
- Witnesses knowledgeable about ABC's system.

Likewise, ABC's counsel should consider serving interrogatories to identify:

- Documents and information describing XYZ's facial recognition authentication system and source code to see if they contain any of ABC's code.
- Documents and information relating to XYZ's hiring of ABC's former engineers.
- Witnesses knowledgeable about XYZ's facial recognition system and source code and hiring practices.

Counsel for both parties should also consider serving interrogatories to gather information about documents and witnesses to testify about:

- The nature and extent of the contested AI system's testing conducted before deployment.
- How and why specific algorithms were selected for the facial recognition system.
- The modeled feature space used in developing an AI model and its relationship to the primary decision variables at issue.
- The relevant scientific community for the technology at issue.
- The relevant industry or technical standards applicable to the disputed AI systems and their outputs (including accuracy, precision, recall).

Because source code may be highly relevant in an IP lawsuit involving an AI system, counsel may also want to use interrogatories to identify:

- Relevant software versions so the applicable versions may be inspected.
- Software development documents, which counsel may use in depositions to understand how the software was developed.
- Documentation regarding various model inputs used, algorithm architectures selected and de-selected, the parameters chosen, and the hyperparameters derived for the various algorithms (for example, anything related to system development).

Because how an AI system is deployed in hardware may help reveal the underlying nature of the software, counsel should consider serving interrogatories to identify:

- All hardware used in a disputed AI system, such as in-cabin cameras for providing image data to the parties' respective facial recognition models.
- Documents describing hardware performance, capabilities, and limitations.
- Witnesses with knowledge of why certain hardware was selected and its chosen placement, which could demonstrate copying.

SOURCE CODE REVIEW CONSIDERATIONS

If permitted by an opposing party or a court, either through the court's local rules or after granting a requesting party's discovery motion, counsel may obtain source code from the other party to understand the data, algorithms, and models behind a disputed AI system. In a patent infringement lawsuit, depending on the nature of the asserted claims, counsel typically will want to conduct a source code review to help reveal how an AI system actually works, as opposed to relying on assumptions about the black box based solely on the system's output or results. In a trade secret case involving allegations of theft and use of secret algorithms, counsel for the trade secret owner should review defendant's source code to see if it includes the same or similar algorithms (for example, the same neural network and related parameters and hyperparameters, or at least similar enough to account for a different training or testing data set split).

APPLICABLE LAW AND RULES

FRCP 34 governs source code reviews. In particular, the rule states that a party may seek to inspect any designated documents or electronically stored information and designated tangible things or request entry onto property possessed or controlled by the responding party, so that the requesting party may inspect, measure, survey, photograph, test, or sample the property or any designated object or operation on it. Rule 37(d) gives federal courts discretion to sanction a party if the "party, after being properly served with ... a request for inspection under Rule 34, fails to serve its answers, objections, or written response."

Because of source code's proprietary and trade secret nature, parties requested to produce their code may resist inspection over concerns about the code getting out into the wild. In patent cases, source code reviews are typically provided for in discovery, scheduling, and protective orders, which presume the source code is needed by one or both of the parties as evidence in their respective cases. Even so, the burden falls to the requestor to establish a need to inspect source code (*Cochran Consulting, Inc. v. Uwatec USA, Inc.*, 102 F.3d 1224, 1231 (Fed. Cir. 1996) (vacating discovery order under FRCP 26(b) requiring the production of computer-programming code because the party seeking discovery had not shown that the code was necessary to the case); *People v. Superior Court of San Diego County*, 28 Cal. App. 5th 223, 241 (Cal. App. 4th 2018) (concluding that the "black box" nature of software is not itself sufficient to warrant its production)).

In trade secret cases, procedures for protecting source code are typically set out in a protective order (*see* FRCP 26(c)(1)(G) (a court may impose a protective order for trade secrets specifying how they are revealed)). Regardless of the burden, some courts, like the District of Delaware, have default standards for handling source code.

If a party seeks to introduce source code into evidence, counsel must specifically identify the code, though it may not be necessary in most case to identify pinpoint citations to specific lines of code to establish

one's infringement contentions. For example, the Eastern District of Texas Patent Rules 3-1 and 3-2 do not require the patent owner to disclose specific evidence or prove its infringement case through its contentions (*see Rapid Completions LLC v. Baker Hughes Inc.*, 2016 WL 3407688, at *5 (E.D. Tex. June 21, 2016)). Counsel should examine the court's local patent rules to determine the applicable infringement contention requirements. For more, see Local Patent Rules for Key Jurisdictions Toolkit (2-563-3765).

SOURCE CODE REVIEW PROCEDURES

A protective order (or other order) applicable to a case may define what constitutes source code and how source code reviews are to be conducted. Counsel should therefore ensure that the definitions in the order accurately reflect the source code that a party wishes to review. For example, a protective order may:

- Define source code.
- Describe the terms and conditions for disclosure and review of source code, as well as special confidentiality markings, including an "Outside Counsel Only - Source Code/Algorithm" designation.
- Require production at a secure facility using non-networked, standalone computers.
- Exclude recording media and devices by inspectors during review.
- Govern the handling of source code as deposition exhibits.

(See, for example, *Vidillion, Inc. v. Pixalate, Inc.*, No. 2:18-cv-07270, D.I. 50 (C.D. Cal. Mar. 22, 2019).)

Other relevant terms of a protective order regarding source code may include:

- The amount of time permitted for all source code reviews by a party (for example, 20 hours, 60 hours, or more).
- The permitted location(s) for source code reviews.
- Hours during the day that inspectors may conduct source code reviews (for example, between 10 am and 3 pm local time).

Counsel commonly seeks to define source code broadly, relying on principles of trade secret law, to include things that the producing party believes in good faith are not generally known to others and have significant competitive value. A producing party may take this approach where unrestricted disclosure to others would harm it, or where the producing party would not normally reveal to third parties except in confidence or where it has undertaken with others to maintain in confidence. Under this approach, a party may define source code to include:

- Computer instructions (reflected in, for example, .jupyter or .py source code files).
- Data structures that define feature sets used in an algorithm.
- Data schema.
- Data definitions (that can be sharable or expressed in a form suitable for input to a data processing module).
- Graphical and design elements (for example, SQL, HTML, XML, XSL, and SGML files).

Counsel should keep in mind that the computer used to inspect source code will likely not have access to any network, and no recordable media or recording devices will be allowed at the inspection location. Therefore, the individuals performing the review should ensure they have requested all the resources installed locally to facilitate inspection and testing, including applications to create virtual servers to simulate remote API calls, if that is an element of the lawsuit. Reviewers should therefore consider requesting in advance that the inspection machine be loaded with:

- The above-listed files.
- Relevant data sets.
- A development environment, such as a Jupyter notebook or similar application, to facilitate opening Python or other source code files and data sets.

In some cases, it may be reasonable to request a GPU-based machine to create a run-time environment for instances of the AI model to explore how the code operates and how the model handles inputs, makes decisions, and takes actions.

Depending on the nature of the disputed AI system, the relevant source code may be embedded on hardware devices (such as sensors) and a party could therefore request this device for inspection along with the software. If neither party has access to a piece of hardware important to their respective cases, counsel may need to obtain it from an outside source.

HYPOTHETICAL LAWSUIT SOURCE CODE REVIEW

In the hypothetical lawsuit, counsel for XYZ and ABC should consider the following while negotiating the terms of the source code review:

- Where will the review be conducted?
- Who will be permitted to review the source code?
- What expertise and experience should counsel's source code reviewing expert have?
- Will all of the source code be reviewed, or can reasonable limits be placed on how much code is to be made available?
- How will specific portions be flagged for later use in evidence (for example, to be made into exhibits for future depositions)?
- What data sets will need to be reviewed?
- Will some of the source code or hardware need to be obtained from a third party?
- How will disputes be brought to the court's attention should disputes arise during the review?
- How should counsel allocate its allowed on-premises review time in an efficient manner to accomplish what it needs?

Counsel in the hypothetical should request that the court issue a discovery protective order that requires source code reviews be conducted at one of the following locations:

- The producing party's respective litigation counsel's offices.
- On-premises at XYZ's or ABC's facilities (whichever one is the producing party).
- At a location in the district where the lawsuit was filed (a neutral site).
- Some other convenient location.

Counsel will likely want to designate a person to monitor adverse counsel's review of its client's source code and respond to questions or last-minute requests.

Counsel will also likely need to arrange for a testifying expert to accompany counsel and perform the actual review (presumably, an expert familiar with multiple programming languages and developer's tools related to the AI system at issue, which counsel would presumably know in advance from reviewing the opposing party's document production and interrogatory discovery responses).

Scope of Review

As part of the hypothetical HAV facial recognition authentication system source code review, counsel should also consider the need to review:

- The disputed AI systems' underlying machine learning algorithms.
- Data training and testing sets.
- Facial feature landmarks data used for identification (taking into consideration applicable data privacy issues).

In addition to the alleged trade secret nature of the underlying code as a whole, the hypothetical assumes that portions of the underlying model's secret machine learning algorithms have inherent trade secret value to ABC because they are, by themselves or collectively, significantly more accurate than competitors' systems at identifying faces. ABC's counsel should therefore review XYZ's source code to determine if XYZ copied those algorithms, which may support a separate legal claim beyond the trade secret misappropriation claim based on ABC's former engineers improperly copying the entirety of the IVIEWDRIVE source code before leaving the company.

XYZ's counsel should also evaluate the alleged trade secret code portions to compare them to its own algorithms to demonstrate that its facial recognition authentication algorithms are different than ABC's.

In an AI case, an algorithm expressed in code (such as Python) may itself not be sufficient to establish patent infringement (or possibly even trade secret misappropriation) if the deployed model based on the algorithms depends on the data set used for training and testing. Theoretically, a model algorithm trained with one data set may produce an infringing result, whereas the same algorithm trained with a different data set may not infringe (depending, of course, on the language of the asserted patent claims). Counsel's and the experts' source code review should therefore include the data sets used by the disputed system. XYZ's counsel should also consider whether ABC's system is equivalent to the claimed system under the doctrine of equivalents, even if it uses a different algorithm, assuming it can be shown that the parties' different machine learning algorithms essentially work in the same or similar way and can produce the same or similar results once fully and adequately trained.

Finally, in the hypothetical source code review, the parties should also seek to inspect the facial features landmarks used for identification purposes, which could affect the model's output. This type of feature set analysis could be used to show that a party's algorithm works in a different way compared to the other party's algorithm, and thus would not infringe under the doctrine of equivalents and would not be considered misappropriation under trade secret law.

THIRD-PARTY DISCOVERY CONSIDERATIONS

If source code is relevant to a lawsuit and neither party has access to it, one or both parties may need to seek the code and related documents from a third-party software developer or authorized seller by issuing a subpoena *duces tecum*. Counsel should expect third parties to resist production of source code on the basis that doing so would be unduly burdensome and because disclosing it to others would put their business interests (proprietary and trade secret information) at risk. The party seeking access to the source code in a contested AI lawsuit should therefore be prepared for discovery motions (presumably, brought by the third party to quash or modify the subpoena) in the district issuing the subpoena. For more on subpoena's and third-party discovery, see Practice Note, Patent Litigation: Third-Party Discovery Considerations (w-000-8584).

APPLICABLE LAW AND RULES

FRCP 45(d)(3)(B) provides that:

> "...[t]o protect a person subject to or affected by a subpoena, the court for the district where compliance is required may, on motion, quash or modify the subpoena if it requires: (i) disclosing a trade secret or other confidential research, development, or commercial information."

However,

> "...the court may, instead of quashing or modifying a subpoena, order appearance or production under specified conditions if the serving party: (i) shows a substantial need for the testimony or material that cannot be otherwise met without undue hardship; and (ii) ensures that the subpoenaed person will be reasonably compensated."

(FRCP 45(d)(3)(C).)

A court "may find that a subpoena presents an undue burden when the subpoena is facially overbroad." (*Wiwa v. Royal Dutch Petroleum Co.*, 392 F.3d 812, 818 (5th Cir. 2004)). Courts have found that a subpoena for documents from a non-party is facially overbroad where the subpoena's document requests:

- "...seek all documents concerning the parties to [the underlying] action, regardless of whether those documents relate to that action and regardless of date."
- "[t]he requests are not particularized."
- "[t]he period covered by the requests is unlimited."

(*In re O'Hare,* Misc. A. No. H-11-0539, 2012 WL 1377891 at *2 (S.D. Tex. Apr. 19, 2012).)

HYPOTHETICAL LAWSUIT THIRD-PARTY DISCOVERY APPROACHES

In the hypothetical lawsuit, if any portion of ABC's facial recognition authentication system includes a cloud-based component or was developed using a third-party cloud-based machine learning as a service (MLaaS) platform or if it resides on a proprietary, purpose-built edge computing device (such as the in-cabin camera hardware/software described in the hypothetical), XYZ's counsel should issue a subpoena, appropriately tailored, to the third-party vendor to obtain information about ABC's use of that platform system if counsel can

demonstrate a substantial need for the information (for example, to establish infringement and the information is not available from ABC). To alleviate the responding party's proprietary information and trade secret concerns, counsel should work with the third party to seek a protective order to limit the scope of the production and implement appropriate confidentiality safeguards, such as limiting disclosure to outside litigation counsel and experts under the protective order.

DEPOSITIONS CONSIDERATIONS

After counsel have obtained the AI-specific written discovery from the adverse party and generally understands the disputed AI system's source code, counsel should prepare to take depositions to help fill gaps in their understanding of the facts relevant to the parties' contentions in the case. In a lawsuit involving a disputed AI technology, counsel should expect to depose not only general fact and expert witnesses, but one or more "source code custodians" and other individuals involved in developing the disputed technology.

APPLICABLE LAW AND RULES

FRCP 30, local rules, case management orders, judges' individual practices, and generally-accepted norms in a particular jurisdiction govern depositions by oral examination. These rules of practice provide the procedures counsel must follow at a deposition, including:

- Giving reasonable written notice to opposing counsel before the deposition.
- Whether documents should be produced by a witness.
- How the deposition is to be recorded.
- How examination and cross-examination and other matters (including raising objections on the record) will be handled.

Strategies for taking and defending depositions are numerous and beyond the scope of this Note. For detailed resources on taking depositions in patent litigation, see Patent Litigation Discovery Toolkit (8-585-6866).

Party Witnesses

Under Rule 30(b)(6):

> "In a Rule 30(b)(6) deposition, there is no distinction between the corporate representative and the corporation. The Rule 30(b)(6) designee does not give his personal opinion. Rather, he presents the corporation's 'position' on the topic. The designee testifies on behalf of the corporation and thus holds it accountable."

(*Sprint Commc'ns Co. v. Theglobe.com, Inc.*, 236 F.R.D. 524, 527 (D. Kan. 2006) (quotation marks and footnotes omitted).)

Also under this rule:

> "...companies have a duty to make a conscientious, good-faith effort to designate knowledgeable persons for Rule 30(b)(6) depositions and to prepare them to fully and unevasively answer questions about the designated subject matter."

(*Sprint*, 236 F.R.D. 524, 527 (quotation marks and footnotes omitted).)

Fact Witnesses

In a technical deposition of a fact witness, such as a data scientist, machine learning engineer, software engineer, or stack developer, investigating the algorithm behind an AI model will help answer questions about how and why a particular system caused a particular result that is material to the litigation. Counsel should therefore ask about the following issues:

- Which algorithms were considered?
- Were they separately tested?
- How were they tested?
- Why was the final algorithm chosen?
- Did an independent third-party review the algorithm and model output?

Regarding the data set used to create the AI model, the deposition taker will want to explore:

- What data sets were used for training, validation, and testing of the algorithm?
- How was testing and validation conducted and were alternatives considered?
- What sort of exploratory data analysis was performed on the data set (or sets) to assess usability, quality, and implicit bias?
- Was the data adequate for the domain that the developer was trying to model and could other data have been used?

Regarding the final model, the deposition taker may want to explore the following issues:

- How old is the model (that is, is it based on old or incomplete data sets)?
- If it models a time-series (for example, a model based on historical data that tends to increase over time), has the underlying distribution shifted enough such that the model is now outdated?
- If newer data were not considered, why?
- How accurate is the model and how is accuracy measured?

Finally, if written discovery revealed an independent third party reviewed the model before a party deployed it, the deposition taker should explore the testing and its results. If sensors are used as the source for new observational data fed to an AI model, counsel should explore why those sensors were chosen, how they operate, their limitations, and what alternative sensors could have been used instead but were not selected.

Expert Witnesses

In an expert deposition, the goal of the deposition shifts to exploring the expert's assumptions, inputs, applications, outputs, and conclusions for weaknesses. If an expert prepared an adversarial or counterfactual model to dispute the contested AI system or an opposing expert's model, a litigator should keep in mind the factors in *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993) and FRE 702, when deposing the expert. For example, the following issues may need to be explored during the deposition:

- Was an adversarial or counterfactual modeled developed, and why?
- Can the expert's analysis be challenged objectively for reliability?
- Was the technique used subject to peer review or publication, or both?

- What was the model's known or potential rate of error when applied to facts relevant to the lawsuit?
- What technical standards were applied to the analysis?
- Was the analysis based on techniques or theories that have been generally accepted in the scientific community?

HYPOTHETICAL LAWSUIT DEPOSITIONS

ABC's and XYZ's counsel should explore the above issues as a starting point during depositions of the parties' respective corporate representatives, other fact witnesses, and expert witnesses identified under FRCP 26 and 30. Similarly, if necessary, counsel should elicit deposition testimony on the above issues from third-party witnesses commanded to testify by subpoena.

For example, XYZ's counsel should seek deposition testimony with at least the following goals in mind:

- Demonstrating that ABC's facial recognition authentication system (both hardware and software), on a claim-element by claim-element basis:
  - performs every element of the asserted patent claims; and
  - works in the same way and produces the same result.
- Demonstrating that the asserted claimed inventions were not disclosed in the prior art and would not have been obvious to a person of skill in the art.
- Establishing that ABC did not take reasonable measures to protect its trade secret algorithms, and that the algorithms:
  - are in the public domain; or
  - could easily be reverse engineered using common knowledge.
- Showing that ABC's trade secret algorithms are not present in XYZ's system (using a source code comparison).
- Demonstrating that ABC's former engineers did not provide the IVIEWDRIVE® software to XYZ.
- Establishing that ABC made and sold an amount of infringing facial recognition authentication systems to OEMs and Tier 1 manufacturers and should therefore pay a reasonable royalty.

ABC's counsel should seek testimony with at least the following goals in mind:

- Showing that XYZ's expert's testing of ABC's facial recognition authentication system for infringement purposes was not reliable.
- Demonstrating that the asserted patent claims:
  - were disclosed in the prior art and were therefore anticipated or would have been obvious, on a claim-element by claim-element basis; or
  - cover abstract ideas.
- Establishing XYZ's source code was modified around the time of its hiring of ABC's former engineers to incorporate specific algorithms that ABC has maintained as trade secrets.
- Showing that XYZ moved its in-cabin camera to a location inside vehicles that would optimize its use of ABC's trade secret algorithms.
- Demonstrating that ABC's former engineers that now work for XYZ will inevitably disclose to XYZ the trade secret algorithms in the IVIEWDRIVE® software that they misappropriated from ABC.
- Establishing that XYZ could use different algorithms in its system and therefore would not be irreparably harmed by a permanent injunction.