AI PATENTING HANDBOOK V3.0

Authors[1]

Frank Chau – Chau IP / F. Chau & Associates, LLC*

Version 3.0 Authors++

Henry Aldridge – Venner Shipley LLP

Thomas Burton – Siemens Corporation

Michael Carey – Chau IP / F. Chau & Associates, LLC

Frank Chau – Chau IP / F. Chau & Associates, LLC

Jennifer Che – Eagle IP

Sumon Dasgupta – Jordan IP Law, LLC**

Christina Huang – Faegre Drinker Biddle & Reath LLP

Dave Kincaid – Cantor Colburn LLP

Jennifer Lacroix – Amundsen Davis LLC

Ryan Phelan – Marshall, Gerstein & Borun LLP

John Pienkos – Amundsen Davis LLC

Mike Jennings – Venner Shipley LLP

Alexander Korenberg – Kilburn & Strode LLP

Edoardo Mirabella – Beckman Coulter, Inc.

---

[1] This article results from a collaboration between members of the IPO's Software Related Inventions Committee and the AI & Other Emerging Technologies Committee. Opinions stated in this article may not be consistent with the firms, corporations, or clients affiliated with the authors.

* Project Chair for version 3.0 and Co-Chair for versions 1.0 and 2.0

** Project Co-Chair for versions 1.0 and 2.0 **++Version 3.0 authors are denoted in bold font in all sections**

AI PATENTING HANDBOOK V3.0

Justin Mullen – Capital One Financial Corporation

Judy Naamat – Troutman Pepper Locke LLP

Nick Transier – Dinsmore & Shohl LLP

Versions 1.0 and 2.0 Authors

Thomas Burton – Siemens Corporation

Michael Carey – Chau IP / F. Chau & Associates, LLC

Frank Chau – Chau IP / F. Chau & Associates, LLC

Sumon Dasgupta – Jordan IP Law, LLC

Christina Huang – Faegre Drinker Biddle & Reath LLP

Dave Kincaid – Cantor Colburn LLP

Jennifer Lacroix – Amundsen Davis LLC

Ryan Phelan – Marshall, Gerstein & Borun LLP

John Pienkos – Amundsen Davis LLC

Mike Jennings – Venner Shipley LLP

Alexander Korenberg – Kilburn & Strode LLP

Edoardo Mirabella – Beckman Coulter, Inc.

Justin Mullen – Capital One Financial Corporation

Judy Naamat – Troutman Pepper Locke LLP

Nick Transier – Dinsmore & Shohl LLP

Shankar Krithivasan – Qualcomm Technologies, Inc.

Nicholas Malden – D Young & Co LLP

AI PATENTING HANDBOOK V3.0

David Pointer – Thomas Horstemeyer LLP

Wen Xie – Global IP Counselors LLP

Table of Contents

I. AI Definitions and Technology Overview

Artificial intelligence (AI) is a rapidly expanding field of technology that has the potential to transform society in multiple ways. Accordingly, many innovative firms and individuals are making improvements in AI technology. Advances in AI technology have been accompanied by explosive growth in the number of AI-related patent applications. The following terms and definitions will be useful in understanding the nature of AI and AI inventions.

A. AI Invention Categories

(Michael Carey; Sumon Dasgupta**)**

AI inventions and AI patents can be categorized as inventions related to 1) core AI technology, 2) applications of AI technology, or 3) the product of AI technology. These categories overlap and patents can include elements of multiple categories. Each category relates to products, designs, processes, computer programs, and other inventions. These categories can influence how a patent or patent application is treated by the United States Patent and Trademark Office (USPTO) as well as the courts.

<u>Improvements to AI</u>

The first category of AI inventions are those that improve AI technology, relate to the operating principles of AI, or have general applicability (i.e., are not limited to specific problem domains). Core AI technology (e.g., AI enablement technology) relates to the building blocks for application-specific tools, including software (e.g., AI training, architectures, and methodologies) and hardware (e.g., computer processors, accelerator chips, and neuromorphic chips).

<u>Applications of AI</u>

Inventions employing core AI Technologies to perform specific tasks in a particular context. Applications of AI technology often integrate AI Enablement Technologies with domain specific systems that provide input data (*e.g.*, sensors) or use the outputs of the AI Technology for a specific end goal (*e.g.*, to predict outcomes or control machines).

Inventions Made Using AI

Inventions that are conceived of or designed by a human with the assistance of an AI technology (AI as a tool for innovation), a human in collaboration with AI (AI as co-inventor), or an AI system without a human inventor (AI as sole inventor).

B. AI Definitions

(Michael Carey; Justin Mullen)

The following terms will be used throughout the document to describe various aspects of machine learning and artificial intelligence:

Architecture

The structure and design of a machine learning (ML) model are encompassed by AI architecture. This includes, for example, how layers are organized and connected within an artificial neural network (ANN). Understanding AI architecture is crucial for developing and optimizing ML models, as it directly impacts their performance and efficiency. Different architectural choices can lead to varying capabilities, making it a key consideration in the design process of intelligent systems.

Artificial Intelligence (AI)

Artificial intelligence (AI) is a wide-ranging branch of computer science focused on developing machines that exhibit intelligence comparable to or surpassing human capabilities. This includes the capacity to perceive, synthesize, and infer information, essentially mimicking cognitive functions. The field encompasses various sub-disciplines, each contributing to the broader goal of creating intelligent systems that can learn, reason, problem-solve, and adapt. AI's applications are diverse, ranging from advanced analytics and automation to natural language processing and robotics, transforming numerous industries and aspects of daily life.

Artificial Neural Network (ANN)

An artificial neural network (ANN) is a machine learning model designed to emulate the operational principles of the human brain. It comprises a network of parametric mathematical functions, known as nodes. These nodes are systematically arranged into interconnected layers. In this configuration, the output generated by the nodes within one layer serves as the input for the nodes situated in the subsequent layer. This hierarchical and interconnected structure enables ANNs to process complex information and learn from data, mirroring the brain's capacity for pattern recognition and decision-making.

Bias

Bias in machine learning refers to systematic errors or prejudices in an ML system's output that disproportionately favor or disfavor certain groups or outcomes. This can arise from biased training data (e.g., data reflecting historical societal biases), flaws in the algorithm design, or even the way human evaluators label data.

Data Label

Data labeling involves annotating data to enhance its interpretability for machine learning (ML) models. A data label is information appended to data, providing context or meaning. Examples include tagging an image of a wolf with "wolf," marking text with "positive sentiment," or flagging transaction data as "fraud." These labels are crucial for training ML models to recognize patterns and make predictions. When data labels offer known, accurate descriptions, they are sometimes called "ground truth" labels.

Explainability

Explainability, often referenced as XAI, refers to methods and techniques that allow humans to understand the reasoning behind a model's decisions or predictions. Unlike "black box" models where the internal workings are opaque, XAI aims to provide transparency and interpretability.

Feature

In machine learning, features are measurable properties or characteristics of data used to train models for predictions. These can be numerical, categorical, or text-based. The quality of features significantly impacts model accuracy. Sometimes, ML models generate or modify features using an encoder, with a decoder then generating output. Feature engineering is the process of making raw data more useful for ML algorithms by cleaning, error removal, filling missing values, and transforming data into an algorithm-friendly format. This crucial step enhances the accuracy and performance of ML models.

Fine-tuning

Fine-tuning adjusts a pre-trained machine learning (ML) model's parameters. Initially trained on vast datasets, the model undergoes further training on a smaller, task-specific dataset. This focused training refines the model, boosting its performance for a designated task. Essentially, it's about specializing a general model to excel in a particular area, making it more efficient and accurate for specific applications.

Hallucination

In the context of generative AI, a "hallucination" occurs when the model produces content that is factually incorrect, nonsensical, or unfaithful to the input data, yet presents it as if it were true or accurate.

Hyper-parameters

Hyper-parameters are crucial settings that define an ML model's architecture and guide its learning process. Unlike parameters learned during training, hyper-parameters are typically set beforehand. Examples include the number of layers in a neural network, which dictates its complexity, and the learning rate, a critical factor determining how quickly the model adjusts to new data. Carefully chosen hyper-parameters significantly impact a model's performance and efficiency, influencing everything from its ability to generalize to unseen data to the speed at which it converges on an optimal solution.

Inference

Inference is the application of a trained machine learning (ML) model to new, unseen data to generate predictions or make decisions. Unlike the training phase, where models learn from existing datasets, inference focuses on leveraging the acquired knowledge to process real-world inputs. This process is crucial for the practical deployment of ML models in various applications, from image recognition and natural language processing to medical diagnosis and financial forecasting. Essentially, inference transforms a learned model into an active tool for interpreting and responding to novel data.

Loss Function

A loss function, also known as a cost or error function, is an objective function in AI and machine learning. It quantifies the difference between a model's predicted and actual outputs, serving as a critical measure of performance. During training, the loss function provides feedback, with larger values indicating significant deviations. The primary goal of optimization is to minimize this loss, typically achieved through algorithms like gradient descent. By iteratively adjusting model parameters based on the calculated loss, the model learns to make more accurate predictions, driving it towards optimal performance.

Machine Learning (ML)

Machine Learning (ML), a subset of Artificial Intelligence (AI), develops mathematical models that learn and improve autonomously from experience, not explicit programming. Though often interchanged with "AI," ML provides the specific methods and algorithms to train machines for AI tasks. This distinction highlights ML as the practical engine driving AI's analytical and adaptive capabilities, enabling systems to discern patterns, make predictions, and adapt behaviors based on data. An ML "model" is a mathematical framework, like those for predictions or human interaction, whose parameters are automatically learned and refined through data during training.

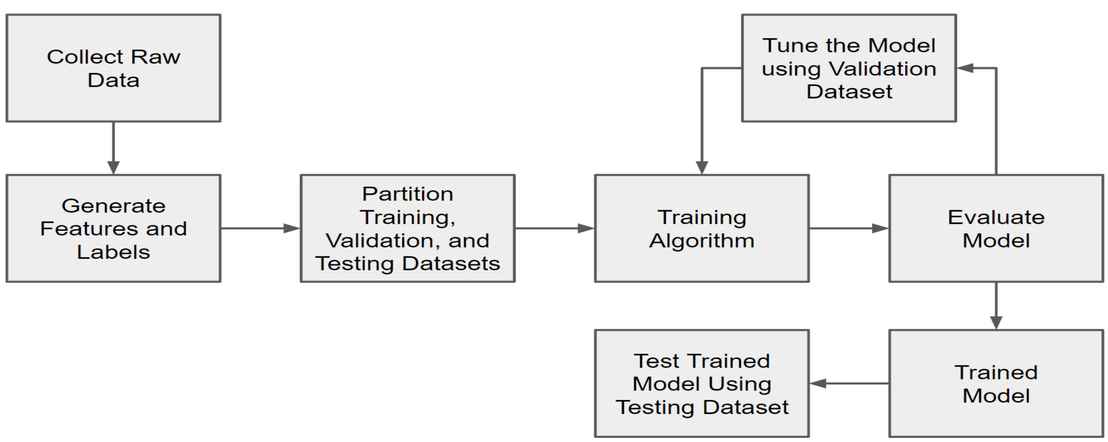
Prompt Engineering

Prompt engineering is the process of defining effective inputs (prompts) for large language models (LLMs) and other generative AI systems to elicit desired outputs. It involves structuring questions, providing context, defining constraints, and specifying formats to guide the AI's generation process.

Training

Training in machine learning involves feeding an algorithm data to learn from. Training may also refer to the broader development of an ML model or specifically to updating its parameters. The goal of training is to find optimal parameter values (e.g., weights) that best represent the training data. This is achieved by iteratively feeding data to the model and minimizing a loss function, thereby allowing the algorithm to learn and refine its internal workings.

A Typical Model Development Process



C. Types of AI

(Michael Carey; Justin Mullen)

AI can be categorized by the architectures and applications it is used for. The following terms are useful in understanding the different types of AI:

<u>Agentic AI</u>

Agentic AI describes an AI system paradigm characterized by autonomous, goal-directed operation within a dynamic environment. These systems architecturally integrate modules for perception, reasoning, planning, and action execution. Crucially, they frequently leverage Large Language Models (LLMs) and other foundation models as sophisticated inference engines for high-level reasoning, task decomposition, and contextual understanding. The system maintains a state, iteratively refines its internal model of the environment, and dynamically adjusts its operational strategy to achieve complex objectives without requiring continuous human intervention or explicit programmatic rules for every contingency. These features distinguish them from purely reactive systems or traditional automation pipelines.

<u>Attention</u>

An attention mechanism is a technique that allows a machine learning model to determine the relationships among input parameters. Attention mechanisms enhance model performance by enabling them to concentrate on pertinent input segments, especially for sequence-to-sequence tasks. This greatly improves their capacity to process extended sequences. Attention can be divided into two main types: cross-attention, which involves two distinct input types, and self-attention, which determines the relevance of different components within a single input.

<u>Backpropagation</u>

Backpropagation is a fundamental algorithm for training neural networks. It optimizes network performance by calculating the gradient of a loss function concerning the network's weights. This calculation allows for efficient adjustment of these weights, iteratively minimizing the difference between predicted and actual outputs. Through this process, backpropagation enables neural networks to learn from data, progressively refining their internal parameters to improve accuracy and generalization capabilities in tasks such as pattern recognition and prediction.

Computer Vision

Computer Vision, an AI subfield, empowers computers to interpret and understand the visual world. It processes images and videos to detect and identify objects or features. Machine Learning (ML) models are integral to this process, as they can be trained to analyze visual data and even generate new images. This technology is crucial for applications ranging from autonomous vehicles to medical imaging, enabling machines to perceive and interact with their environment in increasingly sophisticated ways.

Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a specialized feed-forward neural network primarily used for image processing and computer vision. Its key feature is the use of convolutional layers, which apply filtering operations to input data. These filters are designed to automatically detect and extract spatial features from images, such as edges, corners, and textures, which are crucial for tasks like object recognition and image classification. This hierarchical feature extraction makes CNNs highly effective in understanding and interpreting visual information.

Diffusion Model

Diffusion models create outputs by reversing a noise diffusion process. They start with a simple noise distribution and progressively add detail to achieve a sample from the target distribution. These models are often used in text-to-image generation, where a diffusion network can be trained to produce realistic images conditioned on a text prompt embedding.

Embedding

An embedding converts data (e.g., words, sentences) into a numerical vector in a multi-dimensional space, capturing intrinsic characteristics and relationships. Proximity in this embedding space directly reflects the similarity of original data. Data points with shared features or meanings are mapped to nearby locations, facilitating efficient analysis and processing of complex information.

Feed-forward Neural Network

A feed-forward neural network is a ML model characterized by unidirectional data flow. Information progresses strictly from the input layer, through one or more hidden layers, to the output layer, without forming cycles or loops in its connections between nodes. This architecture ensures that the network processes information in a singular, forward-moving path, making it a foundational element in various machine learning applications.

Foundation Model

A foundation model is a very large AI model, typically a deep learning model, trained on a massive, diverse dataset at scale. These models are designed to be highly versatile and can be adapted (e.g., through fine-tuning) to a wide range of downstream tasks, rather than being built for a single purpose. LLMs are a type of foundation model.

Generative AI

Generative AI is focused on creating synthetic content, including language and images. This content is often produced in response to user-provided prompts or reference data. Its core function lies in its ability to autonomously generate novel outputs that mimic human-created content, demonstrating advanced capabilities in areas such as natural language processing and computer vision.

Large Language Model (LLM)

A large language model (LLM) is a machine learning model meticulously trained on extensive text datasets to produce human-quality text. Functioning by processing input as a sequence of tokens, an LLM iteratively forecasts subsequent tokens to construct coherent and contextually relevant output text. This predictive process enables LLMs to generate diverse forms of content, from articles to code, with remarkable fluency and accuracy.

Natural Language Processing (NLP)

Natural Language Processing (NLP), a key subfield of Artificial Intelligence (AI), empowers machines to interpret and generate human language. This capability enables sophisticated interactions between humans and AI systems, allowing for comprehension of linguistic nuances and generation of coherent, contextually relevant text. NLP's precision is critical for applications ranging from conversational AI to advanced data analysis, marking a significant step towards seamless human-computer communication.

Recurrent Neural Network (RNN)

A Recurrent Neural Network (RNN) is a neural network characterized by directed cycles in its node connections. This unique architecture enables RNNs to leverage an internal state, or "node memory," for processing sequential inputs. Consequently, RNNs are highly effective for tasks involving time-series data and natural language processing. Long Short-Term Memory (LSTM) networks represent a widely adopted variant of RNNs, particularly prevalent in natural language processing applications due to their ability to manage long-term dependencies.

Reinforcement Learning

Reinforcement Learning (RL) is an ML paradigm where an agent learns optimal decision-making through trial and error within an environment, aiming to maximize a cumulative reward. The agent develops a policy by executing actions and receiving feedback (rewards or penalties), iteratively refining its strategy. A significant advancement in RL is Reinforcement Learning from Human Feedback (RLHF), where human preferences guide the learning process.

Instead of predefined reward functions, RLHF leverages human evaluations of the agent's behavior to train a reward model, which then optimizes the agent's policy, aligning its actions more closely with human values and intentions.

Retrieval-Augmented Generation

Retrieval-Augmented Generation (RAG) is a technique that enhances large language models by integrating an information retrieval system with the model. Instead of solely relying on its pre-trained knowledge, a RAG-enabled LLM can query an external knowledge base to retrieve relevant information, which it then uses to generate more accurate, up-to-date, and contextually grounded responses.

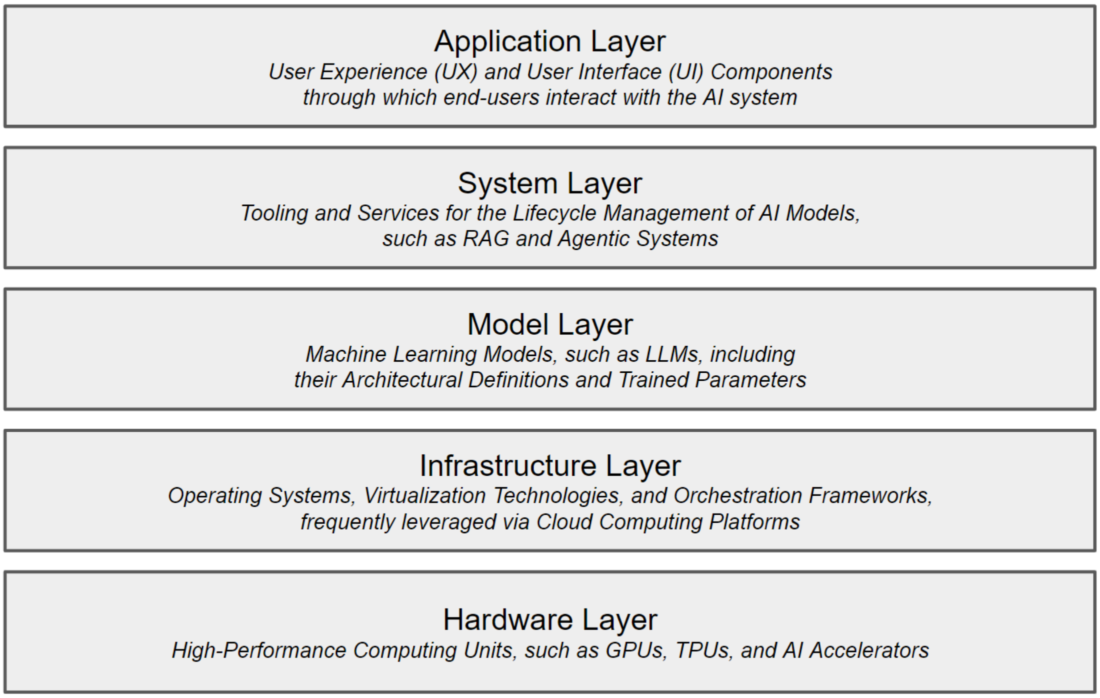
Supervised and Unsupervised Learning

Supervised learning trains models on labeled datasets, mapping inputs to known outputs to predict future outcomes. In contrast, unsupervised learning analyzes unlabeled data to discover hidden patterns or structures without explicit guidance. This method is particularly useful for tasks like clustering or anomaly detection. While supervised learning excels at prediction from historical data, unsupervised learning offers insights into the inherent organization of complex datasets.

Transformer

Transformers are a neural network architecture adept at processing sequential data, notably in natural language processing. Their core innovation lies in sequential processing using attention mechanisms, which enable them to efficiently model internal data dependencies. Unlike traditional networks, transformers often omit recurrent or convolutional layers, allowing for parallel processing of inputs. This architecture is crucial for handling long-range dependencies and has become foundational in advanced AI models due to its scalability and performance.

The AI Tech Stack



II. Drafting AI Patent Applications

Patent preparation involves multiple phases, from identifying initial inventions to drafting comprehensive applications. Certain practitioners, such as in-house counsel, may prioritize inventor engagement, invention harvesting, ranking, selecting invention disclosures, and conducting an initial review of disclosure materials. Others may conduct formal disclosure interviews with inventors and draft patent documents based on their understanding of the invention. Regardless of the preparation phase, a clear framework for understanding inventions helps practitioners distill an invention to its core advancements over prior art. This enables a clear presentation of these advancements in a patent application. To facilitate this process for AI inventions, a technology-specific approach to drafting AI patents is outlined below.

A. Conducting Invention Disclosure Interviews

(Michael Carey; David Kincaid; Shankar Krithivasan; Edoardo Mirabella**; Justin Mullen)**

Patent practitioners often conduct invention disclosure interviews to grasp an invention's details before drafting an application. Incorporating some suggested disclosure questions into an invention disclosure form provided before the interview can be beneficial. This framework, regardless of the practitioner's interaction method, can educate inventors on how to approach their inventions, leading to more comprehensive and valuable patent applications.

1. Setting the Context

AI inventions can span various categories, including data selection, preparation, computing environments/constraints, training methods, inference methods, architecture, and applications. Understanding how the inventive subject matter relates to existing machine learning processes and the AI tech stack is beneficial for patent practitioners, as it facilitates more efficient invention disclosure interviews and discussions with inventors.

For instance, an interview concerning an AI invention related to neural network pruning or reward function tuning would primarily center on the training process and its supporting architecture. Conversely, an interview for a novel application of a conventionally trained neural network would emphasize the inference process and its supporting architecture. Some AI inventions may encompass multiple categories. Sample interview questions are available to help determine the invention's context.

- Disclosure Question: Provide an overview of the invention. What did you do and how did you do it?
- Disclosure Question: Do the advancements of the AI invention relate to data collection, data pre-processing, training, inference, or the underlying AI architecture?
- Disclosure Question: Where within the AI tech stack do the advancements lie?
- Disclosure Question: Does the AI invention rely on any known or open-source

components? If so, which ones and what functionality do they provide?

2. Identifying the Problem

Once the focus of the AI invention is determined, it is helpful to frame the invention in terms of the problem it addresses (i.e., the technical challenge). This approach will aid in addressing both prior art considerations under Sections 102 and 103 as well as subject matter eligibility considerations under Section 101. Like other software technologies, AI inventions often face subject matter eligibility challenges during prosecution and enforcement. Discussing eligibility considerations with the inventors early in the application drafting process enables patent practitioners to draft an application that preempts potential subject-matter eligibility rejections later.

- Disclosure Question: What technical challenges or unexpected problems did you face that led to the AI invention? How did you address them?
- Disclosure Question: Why are conventional or existing approaches inadequate?

3. Identifying Technical Advantages

Once the technical challenges are identified, seek to understand the technical improvements the AI invention provides in that realm of technology. This is useful when attacking a subject matter eligibility rejection using the USPTO's eligibility test (i.e., in determining whether the alleged judicial exception is integrated to a practical application).

- Disclosure Question: What area of technology does your invention improve? Is the functioning of a computer, another technology, or technical field improved?
- Disclosure Question: What makes your solution better than the best existing alternative?
- Disclosure Question: What technical changes lead to these improvements?
- Disclosure Question: Are there any quantifiable conditions or metrics that demonstrate the technical improvement (e.g., reduced processing cycles, memory utilization, power use, etc.)?

In describing the improvement, it should be shown that the functioning of the computer, another technology, or technical field is improved. For example, consider a neural network trained to operate (i.e., perform inference) on a specific device (e.g., an edge processing device). To the extent that the neural network is designed to operate on the device where a conventional neural network could not, such as by compressing its architecture, reducing parameter count, etc., the performance and capability of the device are improved.

Further, even if a conventional neural network could operate on such a device, the inventive neural network may still improve the functioning of the device if the neural network enables the device to operate more efficiently, with fewer resources, faster, etc. To that end, it is helpful to describe and claim architectural aspects of the invention to show non-conventionality (e.g., an RNN used for natural language processing is a specific instance of a neural network that is trained for a specific purpose and causes the device on which it runs to perform a particular task more efficiently than would be possible using conventional computing systems).

4. Understanding Data Collection

During the invention disclosure interview, it's crucial to understand how data is acquired and structured. Is it organized as objects with attributes, or in another format? What are the potential names and values linked to these objects and attributes?

AI inventions differ from conventional technology in how they utilize input data, which can also vary significantly. A crucial discussion point for an invention disclosure interview should be to gain a clear understanding of the invention's data usage. Specifically, it's important to ascertain whether data is being used to develop or train a machine learning (ML) model, or if it's being used by an already-trained model.

- Disclosure Question: What data is used as input to the system? How is it obtained?
- Disclosure Question: What data is used for training the system? How is it obtained?
- Disclosure Question: What is the minimum viable type or amount of data that could be used? Could other types of data be used?

Additionally, it's beneficial to gather information about the data's volume, sources, and acquisition methods. For instance, is the training or input data generated internally or obtained from an external entity? Is the data fed to the neural network in real-time batches or offline?

- Disclosure Question: Are there sensors or other hardware or systems involved in collecting the data?
- Disclosure Question: What is the format of the data when collected?
- Disclosure Question: What challenges did you encounter when obtaining data?

5. Understanding the Preprocessing Steps

Data preprocessing involves cleaning and preparing raw data for analysis. As part of the data preparation, feature engineering may be used to transform raw data into features that make the data more relevant and informative for ML algorithms.

The successful development and implementation of an AI model depends upon both the quantity and quality of data that is employed. Attaining useful or desirable results by way of ML can depend upon whether the data being utilized is accurate, complete, properly formatted, correctly normalized, and the like. Similarly, feature engineering tasks like selecting which portions of data and modifying the data to best fit the model type may support an advancement of an AI invention. Preprocessing or feature engineering may not, in some cases, seem to be the focus of a given invention. Still, even though the addition of "insignificant extra-solution activity" does not amount to an inventive concept, particularly when the activity is well-understood or conventional (*see Parker v. Flook*, 437 U.S. 584, 588-89 (1978)), this general rule should not inhibit the interviewer from gaining an understanding of any preprocessing that is part of the invention.

For example, when conducting an inventor disclosure interview, an effort should be made to obtain an understanding of what types of problems can arise in the data being used for development or implementation of the AI or ML, and how those problems are avoided or alleviated by preprocessing and feature engineering. For example, is raw data preprocessed to correct or add to the data so as to eliminate known problems or deficiencies in the raw data, such as noise that may be impacting the data, or to reorganize or reformulate the data? Does any such preprocessing entail systems or components that involve more than merely a conventional computer processing device that might provide a further basis for contending that the invention is subject matter eligible and constitutes significantly more than any judicial exception? Does the invention rely on transforming the data in a particular way to allow the model to operate more effectively?

The types of preprocessing and feature engineering that can be performed in any given invention can vary significantly depending upon the embodiment, circumstance, or purpose of the AI or ML. For example, if the AI or ML relates to image processing, these operations can include changing various geometric features, rotational orientations, or brightness or color characteristics, performing erosion, or dilation, normalizing features, or performing filtering, image segmentation, or super-resolution.

- Disclosure Question: How is the raw data processed before being provided to the system?
- Disclosure Question: Did you encounter any issues in using the collected data? How does your solution address those issues?
- Disclosure Question: Is the data used for the invention proprietary, publicly available, or both?
- Disclosure Question: Does the nature or format of the data change between being collected and used by the model? If so, how?
- Disclosure Question: Are there any statistical or aggregates of raw data that are used?

6. Understanding Post-processing Steps

In many cases, the output from an AI or ML operation requires further modifications or processing to be useful. The types of post-processing that can be employed in any given embodiment or circumstance can vary widely depending upon the ultimate purpose or use of the results. Such post-processing can be performed by the same processing device(s) that perform the artificial or ML operations, or by other processing devices or other devices or systems.

As with pre-processing, post-processing may seem to be the focus of a given invention. Nevertheless, an effort should be made during an inventor disclosure interview to determine whether any post-processing is performed and, if so, how such post-processing is accomplished and what components, devices, or systems perform such post-processing.

- Disclosure Question: How is the output of the system processed or made usable?
- Disclosure Question: Is there any calibration, normalization, or transformation of the outputs?
- Disclosure Question: Are output data manipulated, transformed, or the like in order to be used by your invention, or by downstream functions utilizing the output of your invention?

7. Understanding the Architecture

A neural network, or parts of a neural network, can often be classified under one of several high-level architectural categories. For example, a convolutional neural network (CNN) includes at least one convolution layer (which finds local features for each data element that take into account neighboring data points) and is often useful for image processing. By contrast, a recurrent neural network (RNN) includes skip connections between different layers, and is often useful for natural language processing, and processing of time series data. In some cases, more than one high- level architecture is used.

- Disclosure Question: What high-level architecture is used for the invention, and how does the architecture relate to the problem being solved?
- Disclosure Question: Is the architecture custom designed or “off-the-shelf”?
- Disclosure Question: Are multiple high-level architectural elements used in the invention?
- Disclosure Question: How do the architectural elements relate to each other?

In some cases, there is an improvement to the functioning of a particular high-level architecture. For example, such an improvement could include the size or number of layers, or the way in which the layers are connected. Furthermore, in many cases, a neural network layer includes a combination of linear (or affine) functions and non-linear activation functions. Thus, an invention could include structure or constraints related to the function of individual layers.

- Disclosure Question: Is there a preexisting similar architecture to the one used in your invention? How does your architecture differ?
- Disclosure Question: What is the structure of the individual layers within the high-level architectural components?
- Disclosure Question: How are the layers within a high-level architectural component connected or arranged?

In many cases, an AI invention is integrated with another device (e.g., a mobile phone, a robot, or a vehicle). In some cases, the training or deployment of the neural network may be distributed across different components, such as in a federated learning model where different layers of a neural network model are implemented in different devices or components. Thus, it is important to understand the context or system architecture of the AI components. For example, an AI system could include various sensors and control systems. Additionally, an AI element can be a part of a service that is connected to user devices, databases, and other computing elements.

- Disclosure Question: What is the computing environment of the AI invention (e.g., are there user devices, databases, or other external elements)?
- Disclosure Question: What alternative architectures could be used?
- Disclosure Question: Which hyperparameters have been set? Why have those values been chosen?

8. Understanding the Training Process

The training process is one of the key elements that differentiates AI inventions from other software inventions. Training claims have both advantages (e.g., it is sometimes easier to show patent eligibility) and disadvantages (e.g., the training claims can be difficult to enforce due to detection and split infringement issues). However, it is important to understand the training process before deciding whether to include training claims.

Before talking to inventors about training, it is useful to have a baseline understanding of a typical training process. ML techniques include supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning, to name a few common examples. There are also variations on each of these methods, such as autoencoder techniques and diffusion processes. Each of these methods typically includes some form of training data, although the training data is used in different ways.

- Disclosure Question: what type of learning algorithm is used for your AI invention?
- Disclosure Question: What training data is used to train the network?
- Disclosure Question: How has the data been obtained or labeled? Are any automated procedures used for labeling?

Supervised learning is a common training technique. It typically involves calculating something called a loss function that determines how well the model has performed at a given task. For example, a simple loss function could include finding the difference between an output of the model and a ground truth value. The gradient of the loss function is then calculated, and from the gradient, an optimization process called gradient descent (or ascent) is used to determine how to update the model parameters. Training hyperparameters (which may be operator adjustable), such as the number of training batches, the learning rate, and others also impact the training.

- Disclosure Question: What loss function(s) are used in the training process? What does each loss function represent?
- Disclosure Question: What optimization algorithm can be used in the training process?
- Disclosure Question: How are hyperparameters set to improve the training process

Sometimes different parts of a model use different training methods. Also, different parts of a model can be trained simultaneously, or some portions of the model can be fixed (or "frozen") while others are trained. Finally, portions of a model can be trained in different stages. For example, a model can be trained, tuned, pruned, and refined.

- Disclosure Question: Are different parts of the model fixed, trained separately, or trained jointly?

- Disclosure Question: Are parts of the network trained in multiple phases?
- Disclosure Question: Are any aspects of the model discarded after training?

9. Understanding the Inference Process

Inference refers to using a trained model (e.g., a trained neural network) to generate an output using unseen data. Consider an example in which an RNN is trained to perform speech recognition. In this example, inference occurs when the trained RNN receives real-world data such as a sound clip of a person speaking and outputs a textual representation of the person's speech. It is important to understand the inputs that the model uses and outputs that the model generates. When describing the inference method in a patent application, it is also useful to describe the underlying architecture that enables the inference and any improvements that result from the inference method.

- Disclosure Question: What inputs are received and what outputs are generated by the trained model?
- Disclosure Question: What processes are used to produce the outputs?
- Disclosure Question: What improvements result from the inference method (i.e., does it address a new problem or to improve performance)?

Furthermore, consider whether the trained model is updated as a result of the inference. This can occur in various ways, such as retraining an existing model or combining predictions from the existing model with a new model created using outputs of the inference. In such cases, the trained model can be improved over time using inference.

- Disclosure Question: Is the trained model updated as a result of the inference? If so, how?
- Disclosure Question: Are the input data used by the trained model expected to change over time?

10. Understanding How Inferences Are Leveraged

An advancement of an invention may come from how the model's inferences are leveraged. For example, in addition to the previously discussed post-processing operations, an AI system may include communication and logical operations to fully leverage the model outputs. Analyzing the operations performed downstream from the model may uncover additional points of novelty and help establish that the invention falls within a statutory category of invention under Section 101.

- Disclosure Question: How are inferences leveraged?
- Disclosure Question: What operations can the model output trigger?
- Disclosure Question: What post-inference operations occur to enable downstream operations?

B. Drafting Claims for AI Inventions

(Michael Carey; Frank Chau; David Kincaid; **Nick Transier;** Wen Xie**)**

The claims of a patent application are of critical importance as they define the precise scope of the legal monopoly granted to the patentee. In the case of inventions involving AI, the process of drafting claims requires careful consideration of both general best practices for computer-implemented inventions and AI-specific nuances that have emerged through recent jurisprudence and patent office guidance.

While AI innovations span diverse technical domains—ranging from natural language processing and computer vision to autonomous control systems and bioinformatics—most AI inventions are implemented in software on conventional computing hardware. As such, guidance from both the USPTO and the Federal Circuit emphasizes that the established principles for claiming software and computer-based inventions apply with equal force to AI-based innovations. This includes drafting claims that emphasize technical improvements to computing performance, rather than abstract functional goals or mere automation of mental processes.

However, simply identifying a claimed method as one involving “artificial intelligence” or “machine learning” is insufficient to satisfy the requirements of patent eligibility under 35 U.S.C. § 101. Courts and examiners have repeatedly found that reciting generic ML techniques (e.g., neural networks, support vector machines) without detailing how they are improved or applied in a non-conventional way renders claims vulnerable to rejection as directed to an abstract idea. This was exemplified most recently in Recentive, where the Federal Circuit held that claims applying standard ML models to a new data environment were ineligible because they lacked a technological improvement or inventive concept.

In addition to subject matter eligibility under § 101, AI-based patent claims must also satisfy the disclosure requirements of 35 U.S.C. § 112, including enablement, written description, and definiteness. These requirements present unique challenges in the context of AI because the internal workings of many machine learning models—particularly deep learning architectures—can be opaque even to their developers. Courts have emphasized that merely disclosing high-level functional outcomes or identifying a known class of algorithms (e.g., “a neural network” or “a support vector machine”) may not be sufficient to demonstrate that the inventor was in possession of the claimed invention or to enable a person of ordinary skill in the art (POSITA) to practice the full scope of the claims without undue experimentation.

Therefore, effective claim drafting for AI inventions must go beyond merely invoking AI as a tool. Practitioners should identify and emphasize any specific improvements to AI models, novel training methodologies, or technical benefits conferred by the AI system on computing functionality—such as reduced latency, improved data compression, enhanced model interpretability, or superior resource efficiency. Moreover, claims should be supported by a well-detailed specification that teaches how the claimed AI model or technique functions in a particular technical context.

1. Subject Matter Eligibility (§101)

To ensure that AI-related patent claims are patent eligible under 35 U.S.C. §101, it is useful to understand the test that examiners use to determine eligibility. Specifically, MPEP 2106 provides the following phases of evaluation: 1) determine if the claims fall into one of the statutorily-defined categories of patentable subject matter; 2) ask if the claims recite an abstract idea (i.e., Alice/Mayo step 2A, Prong 1); 3) consider whether the claim is integrated into a practical application (i.e., Alice/Mayo step 2A, Prong 2); and, 4) determine whether additional elements amount to “significantly more” than the abstract idea (i.e., Alice/Mayo step 2B).

Furthermore, recent developments in *Recentive* and the USPTO have led to important differences between how different AI inventions are handled. Specifically, inventions relating to an improvement to AI can be easier to prosecute and enforce than inventions related to an application of AI. Thus, in some cases it can be desirable to draft claims that recite an improvement to AI technology.

Furthermore, although most AI-related inventions can easily be claimed in terms of one of the statutorily defined categories of eligible inventions (e.g., a process or machine), they may be susceptible to characterization by a patent examiner as an abstract idea, such as a mental process, mathematical concept, or method of organizing human activity.

Thus, there are three basic strategies for claim AI inventions, corresponding to steps 2, 3, and 4 of the process outlined in the MPEP: 1) avoid using language that obviously invokes an abstract idea, 2) include language that integrates the claim into a practical application, and 3) include “additional elements” that are integral to the invention.

- When possible, to avoid characterization as an abstract idea, avoid the following:
- Steps that sound like mental processes (e.g., predict, identify, evaluate, etc.)
- Mathematical equations and notation
- Words that invoke business or legal applications (e.g., contracts, advertising, sales, etc.)

To integrate a claim into a practical application, include steps that are directly related to improving a particular technology, such as:

- Connecting AI results to system level technical improvements (e.g., reduced latency, decreased error rate, more accurate predictions, etc.)
- Reducing the size of training data, the number of parameters, the speed or the number

of layers of a model
- Enabling the use of different kinds of training data
- Enabling lifelong learning (e.g., utilization of previously learned parameters without complete retraining)

To ensure that a claim includes "significantly more" than a judicial exception, include language directed to specific elements, such as:

- Specific architectural elements or unconventional ordering or combination of system components
- Steps that are related to specific ML architectural elements including subcomponents (e.g., self-attention or U-net layers)
- Steps or components that involve physical sensors for collecting data
- Steps or components that result in a physical transformation

2. Disclosure Requirements (§112)

Practitioners must also be mindful of the written description, enablement, and definiteness requirements under 35 U.S.C. §112. District courts and the PTAB have questioned claims that recite broad AI functionality without explaining how the model is trained, what data is used, or how specific results are achieved. As AI technologies often depend heavily on the choice and preprocessing of training data, hyperparameters, and optimization techniques, a specification that fails to sufficiently disclose these aspects may run afoul of § 112(a). Furthermore, claim terms like "intelligently classifying," "learning from data," or "generating optimized outputs" can trigger § 112(b) definiteness issues if they are not clearly defined or tied to concrete implementations.

To mitigate these risks, practitioners should provide detailed algorithmic disclosures where possible, including representative training data, model architectures, training objectives, and example use cases. Even when exact source code is not provided, the specification should teach the POSITA how to construct and apply the claimed model in a reproducible manner. Doing so not only enhances the enforceability of the patent but also aligns with the broader public disclosure goals of the patent system.

Because AI inventions are often claimed functionally, or in such a way as to be performed by a general computing device, Examiners and courts may interpret such claims as styled in Means Plus Function format (MPF), invoking Section 112(f). In such instances, a description of structure corresponding to the claimed nonce or generic term that performs the claimed function must be present in the specification, or the claims can be rejected as failing the definiteness requirement under Section 112(b). Such description must be more than a repeat of the same generic term.

To avoid characterization as MPF, refrain from using non-structural terms that do not have a specific meaning in the art in the body of the claim. Terms considered non-structural include: module, component, unit, element, mechanism, system, and apparatus.

In some cases, more specific terms can be used that have a particular meaning in the field of AI, including terms such as: machine learning model; Artificial Neural Network (ANN); Convolutional Neural Network (CNN); Generative Adversarial Network (GAN); and Reinforcement Learning model

If a non-structural term is used to achieve the desired claim scope, describe in the specification whether the element is a hardware element or if it is implemented using software (or provide a sufficient description of both implementations). If no structure is provided in the specification, then the claim may be rejected during prosecution or later held invalid. This can be the case even in situations where methods of performing the invention are well known. If the claim or the specification is deficient with respect to how the claimed element functions, the claim may be rejected during prosecution or later held invalid under Section 112(a).

3. Training Claims

Many AI inventions are enabled by trained ML models, and the training of these models is often claimable subject matter. In some cases, the training methodology is the primary invention (e.g., federated / distributed learning type inventions), whereas in other cases a conventional training methodology is a means to another inventive end.

While once a safe harbor for training-type claims, Example 39 of the SME Guidance is now routinely ignored by USPTO examiners. Further, new Example 47, regarding Claim 2, presents a much more restrictive view of training claims, and is hard to reconcile with Example 39. Moreover, the *Recentive* case has given examiners yet another lever for rejecting ML claims generally, but especially with respect to training. Furthermore, subsequent examples provided by the USPTOs 2024 guidance call into question the eligibility of claims that recite mathematical training algorithms.

With training claims, as with other types of computer software claims, practitioners should avoid reciting explicit mathematical concepts (e.g., formulas) in the claims. Rather, practitioners should explicitly explain in the specification how the training is unconventional, complex, computationally intense, impractical to perform in the human mind, and how it improves the state of the art with respect to ML training methods. These and other strategies will be further discussed in the sections below.

Training comes in many varieties, including supervised learning, semi-supervised learning, and unsupervised learning to name just a few. Unique aspects of each of these different methods may provide claimable content. Further, in most cases, a training method requires some objective metric to measure the progress of the training, such as a loss or objective function. Because the structure of loss functions can significantly affect the outcome of training, novel loss functions may contribute to claimable subject matter. In some cases, a portion or even an entire model architecture may be dedicated to training and then jettisoned once the model is trained, such as with autoencoder models. In such cases, the unique configurations of the training architecture may provide claimable subject matter.

Finally, training ML models is generally a computationally intensive task. Accordingly, special-purpose hardware (e.g., a GPU, a TPU, and/or a so-called ML or AI “accelerator”) is often used to speed up training. Claiming the special-purpose hardware can help to overcome a rejection based on a “generic computer” implementation.

An emergent trend with USPTO examiners is the requirement for a technical improvement to be recited in the claims, explicitly. In other words, it is often the case that claiming the method or apparatus that conveys the technical benefit is insufficient, and examiners want to see the actual technical improvement in the claim language. This presents several challenges. First, it creates claims for which proving infringement may be difficult. Second, it creates issues with respect to definiteness. For example, if a claim now recites “reducing memory utilization during training,” an examiner or litigant may challenge to what extent the utilization needs to be reduced in order to meet the requirement. Trying to claim some specific percentage to make the claim more definite may then engender written description challenges. Accordingly, reciting a technical improvement in the claims is often a last resort strategy, and requires care to not create downstream validity issues. Practitioners should take care to add language to specifications that will support claiming the technical benefit(s) in a way that avoids Section 112 issues while still being infringeable.

4. Inference Claims

AI claims directed to inferencing (i.e., using a model to generate a useful output, as opposed to training) have several distinct species, including claims directed to: (1) a standard ML model performing a novel task; (2) a novel ML model performing a standard task; and (3) a novel ML model performing a novel task. Note that while “model” is used as a convenient term in the discussion below, it is often the case that an AI-related invention uses a processing architecture that may include one or more models as well as other processing elements.

For claims directed to a standard ML model performing a novel task, the strategy should include a focus on what the novel task is because the ML model is a “black box” unlikely to confer eligibility by itself. In other words, similar to *Alice’s* rationale, the mere recitation of a generic ML model cannot transform a patent-ineligible abstract idea into a patent-eligible invention. *Recentive* seems to reinforce this point. However, citing specific elements of the operation of the ML model can still be a hook for eligibility even if the model is not novel.

For claims directed to a novel ML model performing a standard task, a strategy is to explicitly claim what makes the model novel, such as a specific configuration or combination of layers, activation functions, and/or connection pathways between the layers for a neural network model. Further, the combination of different types of models (e.g., an ensemble architecture) can combine benefits of different model types in novel ways. The section below regarding drafting eligible architecture claims provides additional guidance.

For claims directed to a novel ML model performing a novel task, reciting elements of the novel task may allow for claiming the novel model more flexibly or more broadly. Because explicitly claiming specific model architectures increases the opportunity for design-around and may require disclosure of aspects of a model that might alternatively be kept trade secret, the novel task can be leveraged as a complementary tool for eligibility.

As discussed above with respect to training claims, inference claims are also frequently subject to examiner requirements to explicitly recite technical improvements in the claims. Similar strategies as discussed above are equally relevant here.

5. Architecture Claims

Best practices for drafting AI architecture claims are similar to those for drafting training and inferencing-type claims as discussed above. However, whereas claims related to training and inferencing often include elements covering the respective processes end-to-end, an architecture claim might instead focus on a particular portion of the architecture. This is because, for example, a conventional training technique can be applied to a novel architecture to perform a conventional inference task.

In many cases, a complete claim set will include training, inference, and architecture claims that each have a different scope and rely on different patent eligibility strategies. That is, some elements that have been included in other claims may be omitted in architecture claims if other elements are included to ensure eligibility. There are multiple aspects of an AI architecture to consider, including the software (or process) architecture as well as the hardware (or compute) architecture.

In particular, AI architecture claims may be directed to specific components and processing patterns that can be recited at a high level of generality without being considered generic computing components. Examples of these architectural components include a convolutional neural network (CNN), a recurrent neural network (RNN), transformer-based models (e.g., Bidirectional Encoder Representations from Transformers (BERT), and Generative Pre-Trained Transformer (GPT), language models (e.g., Large Language Models (LLMs) and diffusion transformer models), to name but a few). These components are "particular machines" that are not generic because 1) such components are not present in most typical computing systems and 2) such components are often suited for particular applications (e.g., a CNN is well-suited for image processing applications, whereas an RNN may be better suited for natural language processing (NLP)).

Referencing specific AI architectural elements in a claim can prevent a claim from being considered a judicial exception under step 2A, prong 2 of the USPTO's subject eligibility guidelines. For example, using training or architectural elements can provide grounds for showing that the claim is integrated into a practical application. Even where the inventive aspect relates to inference, it can also be useful to describe and claim architectural aspects of the invention to show non-conventionality. For example, an RNN used for NLP is a specific instance of a neural network that is trained for a specific purpose and causes the device on which it runs to perform a particular task not performed by conventional computing systems.

There are also ways to incorporate architectural elements into inference claims. For example, a step of a method claim could recite an architectural element that performs the step, or the step could be phrased in a way that specifies a particular kind of processing. For example, a claim that uses a CNN could claim "performing a neural network convolution" on input data.

Finally, AI inventions often include hardware aspects. For example, self-driving vehicle systems rely on the input of specific types of hardware sensors, such as cameras and ranging sensors, to provide input data to trained ML models, which in turn generate the outputs that drive (or assist with driving) the vehicle.

6. Problem/Solution Claims

Claims that do not recite specific architecture or training elements can still be crafted to avoid a subject matter eligibility rejection. Specifically, a claim can recite a problem, a technical solution, and the resulting improvement over prior art systems.

For example, if an image recognition system is used to navigate through a physical environment, the input to the system could be described in a way that identifies a problem (e.g., "receive a 2D image that depicts an obstacle in an environment"). Intermediate steps can describe the technical solution (e.g., "generate a depth map indicating a distance of the obstacle"). Finally, the result can be described as an action taken based on the ML model (e.g., "navigate through the environment while avoiding the obstacle based on the depth map").

Accordingly, in addition to including architectural and training elements, AI claims can be patent eligible if they include elements directed to improving the functioning of a computer or an ML model. Such claims should include specific terms that correspond to 1) the problem, 2) the technical solution, and 3) the result. The problem, technical solution, and result should be fully described in the specification to provide support for advocating that the claim is not directed to a judicial exception or includes "significantly more" than a judicial exception.

C. Drafting the Specification

(Michael Carey; Frank Chau; Christina Huang; **Judy Naamat;** David Kincaid; **Nick Transier;** Wen Xie**)**

While the claims are important in defining AI inventions, there are a number of AI-specific drafting techniques applicable to the written description to avoid or overcome rejections and challenges based on Sections 101 and 112.

1. Subject Matter Eligibility (§101)

Although subject matter eligibility under 35 U.S.C. § 101 is primarily based on the claims, the specification can impact whether the claims are deemed ineligible. The primary reason for this is that claimed terms are understood in light of the specification. For example, the MPEP 2106.04(a)(2) III(c) states "in evaluating whether a claim that requires a computer recites a mental process, examiners should carefully consider the broadest reasonable interpretation of the claim in light of the specification." The corpus of Federal Circuit cases that have found computer-implemented inventions eligible under Section 101 have frequently cited to the specification in the decision rationale.

The fact that claim terms are understood in light of the specification is particularly important because the terminology used in an AI/ML context can have meanings that differ from conventional usage. For example, the term "predict" can result in the determination that an element of a claim is directed to a mental process. If terms such as "predict", "identify", "observe", "evaluate" are used in a claim, it is important to include a technical description of these terms that describes them in an AI/ML context with a technical definition that can differentiate from a mental process.

Even terms such as "encode" could potentially be performed by a human with the aid of pen and paper, so providing technical details of how encoding is performed can help avoid a claim interpretation that causes a claim to be ineligible. For example, details can be included that clarify that "encoding" includes representing data in a high-dimensional embedding space, or that the encoding refers to a process performed with a particular ML model.

A claim that includes abstract elements can still be patent eligible if the judicial exception is integrated into a practical application. For example, the specification may include descriptions of a problem in the prior art, the technical solution, the benefits derived therefrom. The improvements may be to computer functionality, or to any other existing technology. The specification is also considered when determining whether a claim includes an additional element that amounts to significantly more than an abstract idea. Here, it is important to clearly describe any additional elements that are included in a claim and how they are related to the inventive concept.

2. Written Description, Enablement, and Best Mode (§112(a))

Section 112 of the Patent Act sets forth key disclosure requirements that ensure a patent application sufficiently informs the public about the invention. The three core requirements are: (1) written description under § 112(a); (2) enablement under § 112(a); and (3) definiteness under § 112(b).

Regarding written description, the patent must clearly describe the invention in sufficient detail to demonstrate that the inventor was in possession of the claimed invention at the time of filing. This requires more than functional language or aspirational results; it must show what the invention is and how it works.

Regarding enablement, the application must teach a person of ordinary skill in the art (POSITA) how to make and use the full scope of the claimed invention without undue experimentation. This includes disclosing necessary steps, structures, algorithms, materials, or data, depending on the field. In *Amgen Inc. v. Sanofi*, 598 U.S. 594 (2023), the Supreme Court reiterated that the enablement requirement demands more than disclosure of a general research plan or an invitation to trial-and-error experimentation.

Regarding definiteness, the claims must particularly point out and distinctly claim the subject matter regarded as the invention. A claim is indefinite if it fails to inform a POSITA, with reasonable certainty, what falls within the claim's scope.

These requirements collectively ensure that the patent provides sufficient public notice, supports the scope of the claims, and advances the quid pro quo of disclosure for exclusivity. For many AI-inventions, the written description requirement can be viewed as relating more to architectural elements (including algorithms), whereas the enablement requirement relates more to training (i.e., making AI models) and inference (i.e., use of AI models). Thus, novel details relating to training, inference, and architecture should all be described in virtually every AI patent application.

The enablement requirement states that the specification must describe the claimed invention in a manner understandable to a person of ordinary skill in the art (POSITA).

Unfortunately, the term "a person of ordinary skill in the art" is not perfectly clear in a fast-moving field such as in AI. Terms that seem obvious to inventors might be completely unknown to examiners and judges. Furthermore, architectures, such as transformer models and diffusion models, can go from being virtually unknown to widely used in a matter of months. In sections below, guidance is provided for levels of descriptions of AI inventions in the specification to help avoid written description rejections.

The Supreme Court, in *Amgen Inc. v. Sanofi*, has emphasized that the broader the scope of the claims, the more the applicant must enable with teachings in the specification. The teachings need to provide enough information that a POSITA could make or use the invention without more than a reasonable amount of experimentation. Therefore, when claiming a genus, a description needs to be provided for making or using each member of the genus. It is tempting (and sometimes wise) when applying ML to a field that is new to ML to claim broadly. Such breadth needs commensurate description, with consideration taken of how much experimentation would be needed by a POSITA to design each of the claimed systems or to perform each of the claimed methods.

The Federal Circuit in *Regents of the Univ. of Cal. v. B*road*. Inst., Inc.,* 136 F.4th 1367 (Fed. Cir. 2025) held that the greater the unpredictability and complexity of the invention, the greater the level of detail needed in the specification to satisfy the written description requirement.

From *In re Floyd*, 2025 U.S. App. LEXIS 9504 (Fed. Cir. 2025), we see that written description issues can arise when a design patent claims priority to a utility patent. For example, when the technology in the utility application is simple and quite predictable, the written description requirement may be relatively low compared to that of a design patent, The claimed design is not limited to utilitarian functionality, and may range from the straightforward to the ornate. As an example, a utilitarian application that taught a cooling blanket that had compartments configured as a 6x6 grid and a 6x4 grid did not satisfy the written description requirement for a design application to a cooling blanket that had compartments configured as a 6x5 grid.

Novartis Pharms. Corp. v. Torrent Pharma Inc. (*In re Entresto (Sacubitril/Valsartan))*, 125 F.4th 1090 (Fed. Cir. 2025) teaches to beware of drafting claims based on futuristic discoveries not yet enabled at the time of conception. This can arise not only in pharmaceutical contexts, but also in the context of emerging technologies, such as artificial intelligence.

In a recent ruling about an ingredient in a medication for COVID-19, *In re McLeay*, 2025 U.S. App. LEXIS 3626 (Fed. Cir. 2025) points out that the fatal flaw of the claims was a need for undue experimentation, not lack of utility. Questions about the utilitarian requirement based on efficacy of the ingredient were merely informative, but not dispositive.

While it is imperative to be diligent in providing detailed descriptions of that which is claimed as the "invention," the description requirement for non-invention parts of the product are lower. *In in Teva Branded Pharm. Prods. R&D, Inc. v. Amneal Pharms. of N.Y.*, LLC, 124 F.4th 898, the Federal Circuit illustrated with a hypothetical about a claim to an improved steering wheel intended for use in a car, in which description and enablement of the steering wheel would be required, but not of the car.

As taught in *In re Riggs*, 131 F.4th 1377 (Fed. Cir. 2025), care is needed when drafting a provisional application to provide sufficient description and enablement, since once published, the provisional application will be treated as a published non-provisional application or patent, and can be applied as prior art for all that it teaches as of its filing date in compliance with 35 U.S.C. §112(a).

We learn from *Regeneron Pharms., Inc.* v. *Mylan Pharms. Inc.*, 127 F.4th 896 (Fed. Cir. 2025) that the written description requirement can still be satisfied when a claim element is omitted in a later-filed claim (e.g., relative to originally filed claims of the application or a parent) provided that there is indication that the element can be optional and the claim is still workable, and provided that the specification and prosecution history do not ascribe a needed function or significance to the element.

In re Xencor, Inc., 130 F.4th 1350 (Fed. Cir. 2025) teaches that the preamble of a Jepson claim requires a written description.

Finally, while some would argue that the America Invents Act (AIA) eliminated the consequences for failing to disclose the inventor's best mode, the requirement still exists. For AI applications, this means that a practitioner should describe any preferred forms of training, architecture, and inference.

3. Definiteness (§112(b))

Section 112(b) requires that claims must particularly point out and distinctly claim the invention, i.e., they must be definite. As with the patent eligibility rules, this requirement relates directly to claims, but the specification will help determine whether the claims are considered definite.

The following can be helpful for establishing definiteness under Section 112(b). Care should be taken to describe a function rather than stating desired results. This can include one or more of: provision of an algorithm (e.g., a flowchart, pseudocode); a description of how the function is performed; a description of what elements are used to perform the function; a description of how exceptions are handled; and provision of examples of how exceptions are handled.

Regarding definiteness under Section 112(b), a determination of indefiniteness may be avoided by providing method steps that are performed automatically in a way that can be readily understood by a POSITA, as opposed to inclusion of user actions, which are more likely to render a claim indefinite. In some cases, claim terms for well-known or off-the-shelf ML and/or AI modules may avoid invocation of Section 112(f), whereas specialized ML modules would require sufficient description to provide structure, such as via algorithms, a description of the input, output, and functions performed, a description of how exceptions are handled. Examples are helpful in showing that an inventor had possession of the invention at the patent application's priority date.

In re Shafovaloff, 2025 U.S. App. LEXIS 15890 (Fed. Cir. 2025) provides a reminder that contradictory limitations within a single claim or between different claims can render a claim indefinite.

Use care when using a term of art in a nonconventional way, such as by providing a clear definition in the specification. *Impact Engine, Inc.* v. *Google LLC,* 2024 U.S. App. LEXIS 16254 (Fed. Cir. 2024) points out that when using such a term in a nonconventional way, you will no longer be able to rely on its conventional meaning if the need arises – that meaning will no longer be available for the term.

The definiteness requirement can be satisfied even for a term that could seem facially indefinite. This arose in *In Sci. Applications Int'l Corp. v. United States*, 169 Fed. Cl. 643 (Fed. Cir. 2024), regarding an evaluation of whether a "potential representation" transitioned to a "corresponding region" of an external environment. The definiteness requirement was satisfied because the specification included written description, illustration, and examples with a detailed overview that provided objective criteria and methods for performing the analysis. Multiple available methods to perform the analysis and even the possibility of different results did not render the claim indefinite, nor did involvement of humans in performing the evaluation. Satisfaction of the definiteness requirement was due to the objectivity of the evaluation, lack of dependency on tastes, opinions, preferences, or individual judgments of the human users, and lack of susceptibility to substantial differences in human perception.

The Federal Circuit in *Fintiv, Inc. v. PayPal Holdings, Inc.* (Fed Cir. 2025) held that "payment-handler" terms invoked 112 ¶ 6 (pre-AIA, corresponding to AIA 112((f)) interpretation and the claims are indefinite under 112(b) "if specification fails to disclose adequate corresponding structure to perform claimed function". Fintiv argued that the patents disclose a two-step algorithm: "(1) "wrap[s] APIs of different payment processors, such as, for example banks ..." and (2) "exposes a common API to facilitate interactions with many different kinds of payment processors." The Federal Circuit rejected this argument, finding that "reciting the function of the payment-handler terms using generic terms without providing any details about an algorithm to carry out the functions of using APIs of different payment processors" does not constitute an algorithm. It is recommended to include at least one specific implementation example to avoid such an outcome.

4. Three Levels of Description

The key to describing AI and ML architecture is to become familiar with multiple levels of abstraction. As a rule of thumb, an AI invention should be described on at least three levels: a first level based on the target function, a second level based on high-level technical description, and a third level based on a detailed technical description. By understanding these levels of abstraction, it is possible to describe how the structure performs a claimed function.

For example, when describing architecture, the highest level of description is the functional component description (e.g., an image classification network). On its own, this description may not be enough to satisfy the requirements of Section 112. Furthermore, functional terms might be interpreted as a nonce under Section 112(f). Thus, it is essential to include another level of description based on the high-level architectural paradigms described above (e.g., CNN, RNN, feed-forward network, etc.). The specification should also describe the relationship between the problem to be solved (e.g., image recognition) and the high-level architecture used to solve it (e.g., a CNN).

However, to fully satisfy the requirements of Section 112, it is important to include details that go even deeper than a recitation of the basic architecture. Thus, a third level of description can be included that provides technical details about the operation of the network at the level of layers, nodes, and activation functions. This does not require describing the actual parameter. For example, the specification could include a description of how a CNN works at the node level (e.g., describe the role of different filters of the CNN). If possible, the inventive concept should be woven into the description at each level of description.

As with the different levels of architectural description, there are also three levels of description that are useful when describing the training process. Again, at the highest level, a practitioner should include a functional description (e.g., a neural network trained to classify objects in an image). At the next level, the practitioner should provide details broadly descriptive of a high-level training paradigm, such as supervised learning, unsupervised learning, or reinforcement learning. Then, at the level of fine detail, the practitioner should provide specifics related to, e.g., the loss functions of a supervised learning process or a policy model of a reinforcement learning process.

5. Problem/Solution Description

In addition to technical details related to the architecture and training, special attention should be given to providing a description of a technical problem, the technical solution, and the resulting improvement over related technologies.

When describing a problem in the existing technology, a practitioner should provide enough description to motivate the solution provided by the invention without conceding too much background as prior art. However, it is generally useful to identify a field of technology, provide a generic name for some device or task in the field, and then describe a problem faced when implementing such a task or device.

When describing the technical solution, provide a description of the structure that performs each function recited in the claims, including how such structure performs the function. If the structure is implemented in software, then provide an algorithm for how the function is performed. Furthermore, if an AI-specific term is claimed, then provide a non-limiting definition or provide a description of an exemplary use of the claim term in the specification.

Thus, in addition to a description of the problem and a description of the technical solution, it is useful to describe how the technical solution results in an improvement over existing systems and methods. Preferably, the improvement should relate directly to the problem and be specific to a particular area of technology.

III. Prosecution and Enforcement

Many of the considerations for prosecuting and enforcing AI applications are similar to those for other areas of technology (e.g., traditional software inventions). However, several unique considerations for AI applications are described below. Of particular importance is the evolving manner that AI patents are considered under subject matter eligibility analysis.

A. Office Action Rejections

(Michael Carey; Frank Chau; Nick Transier)

1. Patent Eligibility Rejections (§101)

There are three potential strategies for responding to an Office Action that includes a Section 101 rejection. The three strategies are:

1) Arguing that a claim does not recite (or is not directed to) a judicial exception;
2) Arguing that a claim is integrated into a practical application; and
3) Arguing that additional elements (or a combination of elements) add significantly more than conventional solutions.

These strategies parallel the three strategies for *avoiding* a 101 rejection in the first place. Specifically, they correspond to steps 2A, prong 1, 2A, prong 2, and 2B of the subject eligibility test, as described above.

For AI inventions, recent developments have made it more difficult to employ strategy #1. Almost every claim includes some terminology that could be characterized as a mental process, a mathematical concept, or a method of organizing human behavior. However, variants of strategies #2 and #3 remain effective.

Specifically, with respect to strategy #2, it is effective to argue that an invention relates to an improvement to AI technology. With respect to #3, it remains effective to show that an invention includes steps that result in tangible outputs that constitute an improvement to previous technology. However, for inventions that relate to business processes, strategy #3 is not often effective.

Accordingly, The implementation of these strategies differs somewhat from strategies for other technology due to some specifics about AI technology. Furthermore, the manner in which the USPTO implements the four-part test is rapidly changing. The following provides evidence-based strategies based on recently collected data and observations about how AI-related applications are being treated at the patent office.

The rate of Section 101 rejections is consistently on the rise, and is at nearly the highest level since before the SME Guidance was published in 2019. Figure 1, below, shows the rate of Section 101 rejection across all art units from 2009 to present:

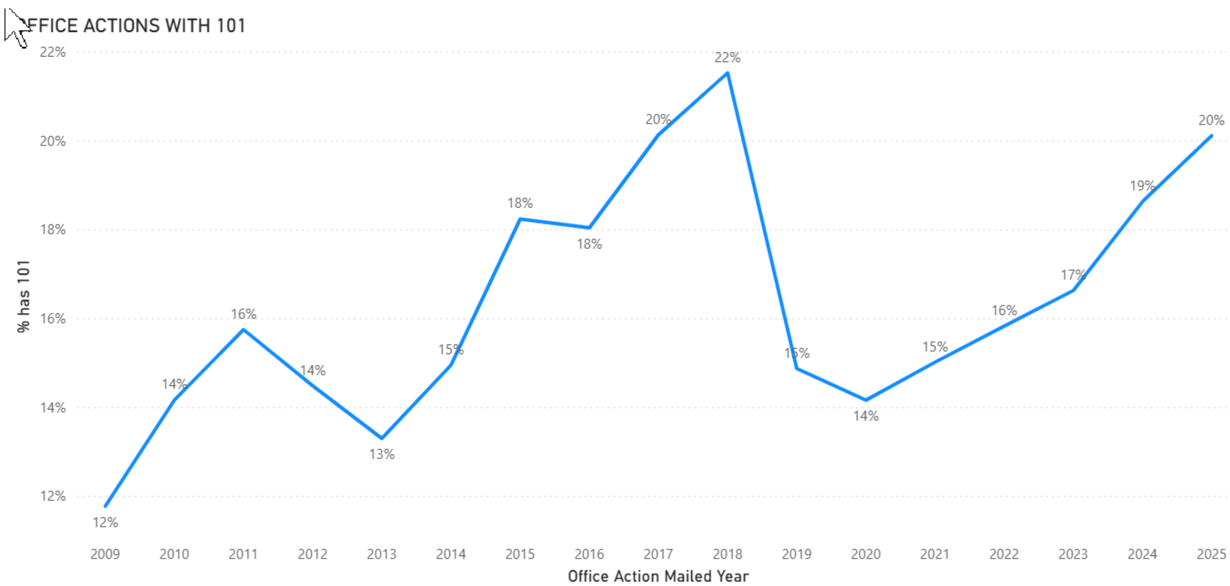


Figure 1: Section 101 Rejection Rate (All Art Units)

According to the USPTO, the Art Units that handle AI inventions most frequently are 2120-2129 and 2140-2148.[2] Thus the following figures depict the increased rate of Section 101 rejections in these art units.

Figure 2, below, shows the Section 101 rejection rate for USPTO Art Units 2121-2129.

[2] *See* https://www.uspto.gov/patents/contact-patents/tc-2100-management-roster

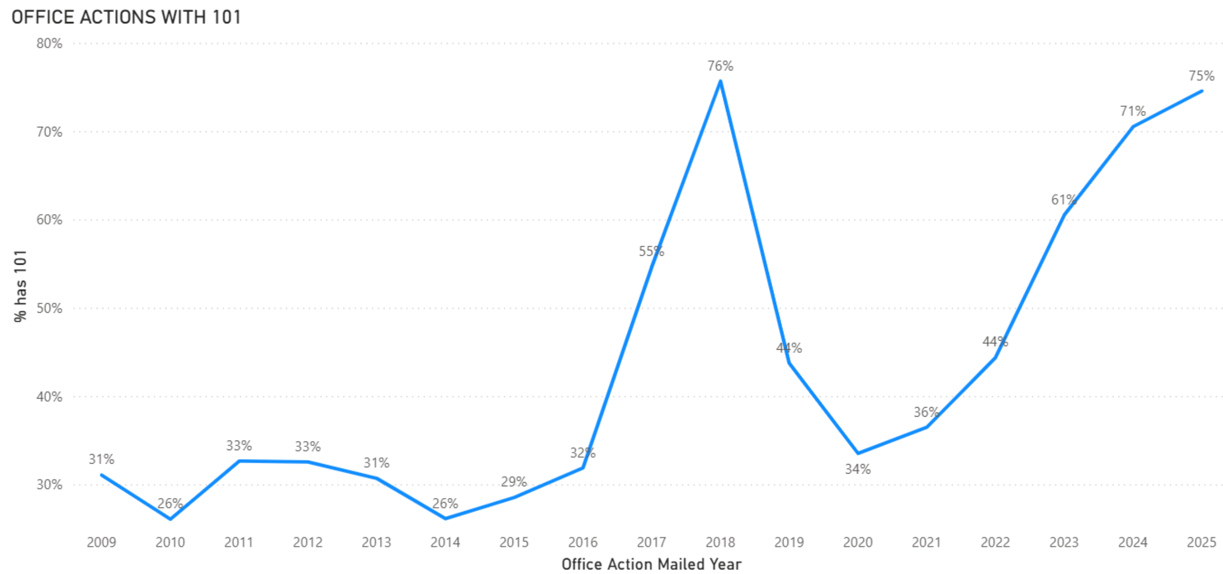


Figure 2: Section 101 Rejection Rate (Art Units 2121-2129)

Figure 3, below, shows the Section 101 rejection rate for USPTO Art Units 2141–2148.

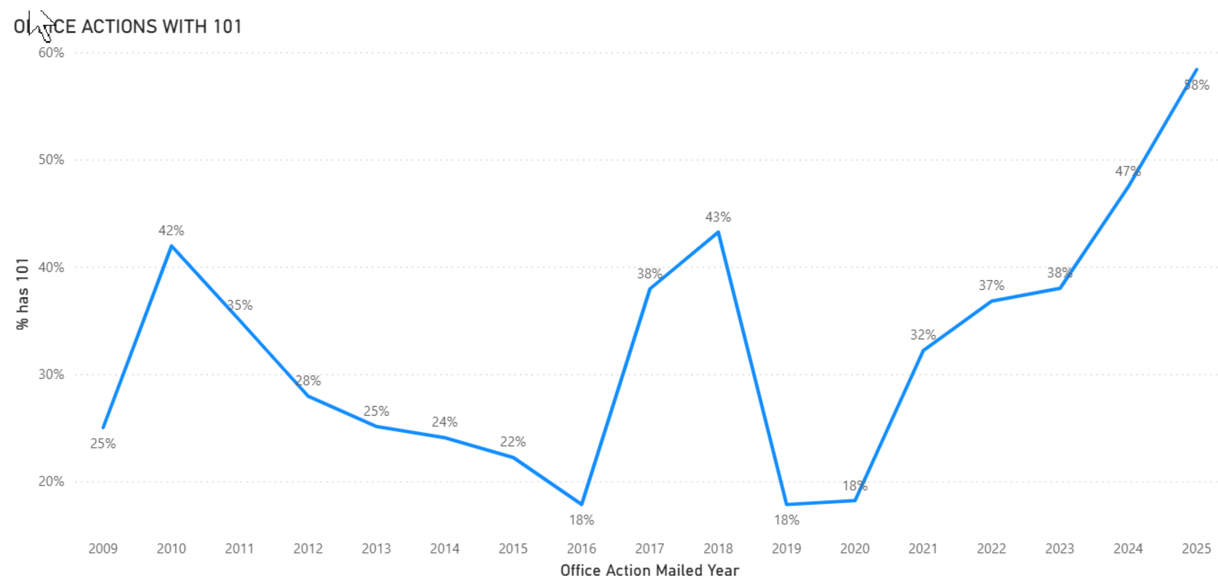


Figure 3: Section 101 Rejection Rate (Art Units 2141–2148)

2. 2023 Patent Eligibility Rejection Study

A review of 200 AI-based patent applications that were rejected under 35 U.S.C. §101 that issued as patents between 1/1/23 to 9/30/23 found that:

2A Prong 1

- 141 out of 200 applications (71%) were rejected under mental process category (37 of these are rejected with human activity category and 23 of these are rejected with mathematical category as well)
- 84 out of the 200 applications (42%) were rejected under the human activity category (4 of these applications were rejected under the mathematical concept category as well)
- 30 out of the 200 applications (15%) were rejected under the mathematical concept category.

2A Prong 2

- 184 out of the 200 applications (92%) were found to use generic computer components to perform the abstract idea. (e.g., "computer system/product, device, memory, processor, non-transitory computer readable medium, ML model")
  - 111 applications included a "processor"
  - 48 applications included a "ML model" or "AI model/technique"; and
  - 11 applications included a "neural network"
- 71 out of the 200 applications (36%) were rejected based on limitation reciting data gathering that amounts to insignificant extra-solution activity.
- 66 out of the 200 applications (33%) were rejected for other reasons. For example, the claim as a whole does not integrate the abstract idea into a practical application or does not amount to an improvement in a technical field, etc.

2B

- 193 out of the 200 applications (97%) were evaluated and rejected under this step.
- 89 out of the 200 applications (45%) were rejected for being well-understood, routine,

conventional activity in the field.

- Of these rejected applications, most of the Examiners reasoned that either data gathering is conventional or using generic computer components is conventional.

3. Mental Process Rejections

Of the three categories of abstract ideas, "mental process" is the most cited in rejections of AI claims, about 71% in 200 rejections. Therefore, "mental process" rejections merit special attention.

One strategy for overcoming a mental process-type rejection is to amend the claims by modifying language that sounds like a mental process (e.g., change the word "predict" to "compute") or adding additional language specifying architecture-specific elements that cannot practically be performed mentally (e.g., "performing a convolution on an image" or "generating an embedding vector representing text"). Alternatively, a practitioner can simply argue that the existing language does not recite a mental process.

The MPEP explains that "[t]he courts consider a mental process (thinking) that "can be performed in the human mind, or by a human using a pen and paper" to be an abstract idea. MPEP § 2106.04(a)(2)(III) (citing *CyberSource Corp. v. Retail Decisions, Inc.*, 654 F.3d 1366, 1372 (Fed. Cir. 2011)). "[E]xamples of mental processes include observations, evaluations, judgments, and opinions." *Id*. And notably, "both product and process claims may recite a "mental process," *id*.; in other words, an apparatus configured to perform a so-called mental process would be rejected as a method claim directed to the mental process.

While claims requiring a computer to perform the alleged mental process run afoul of the rule, *id*., there is an important carve-out. Namely, "[c]laims do not recite a mental process when they do not contain limitations *that can practically be performed in the human mind*, for instance when the human mind is not equipped to perform the claim limitations. MPEP § 2106.04(a)(2)(III)(A) (citing *SRI Int'l, Inc. v. Cisco Systems, Inc.*, 930 F.3d 1295, 1304 (Fed. Cir. 2019)) (emphasis added).

Note that "impractical" and "inconvenient" to perform in the human mind are not the same, and improving convenience and speed by using a computer to assist will not save a claim. *See generally* MPEP § 2106.04(a)(2)(III)(C). The MPEP provides examples of claims that do not recite mental processes because they cannot be practically performed in the human mind, including:

- A claim to a method for calculating an absolute position of a GPS receiver and an absolute time of reception of satellite signals, where the claimed GPS receiver calculated pseudo-ranges that estimated the distance from the GPS receiver to a plurality of satellites.
- A claim directed to detecting suspicious activity by using network monitors and analyzing network packets.
- A claim to a specific data encryption method for computer communication involving a several-step manipulation of data.
- A claim to a method for rendering a halftone image of a digital image by comparing, pixel by pixel, the digital image against a blue noise mask, where the method required the manipulation of computer data structures (e.g., the pixels of a digital image and a two-dimensional array known as a mask) and the output of a modified computer data structure (a halftoned digital image).

See MPEP § 2106.04(a)(2)(III)(A). Notably, the difference between these eligible examples and ineligible examples, such as "a claim to 'collecting information, analyzing it, and displaying certain results of the collection and analysis,'" *id.*, seems to be related at least in part to the level of specificity in the claim. After all, the eligible GPS method (example 1, above) is an example of "collecting information" (e.g., GPS radio signals) and "analyzing it" to generate an output (e.g., a position estimate).

Relatedly, the PTAB has held, on several occasions, that when claim elements are described in the specification as being "complicated", "difficult to scale", "computationally intensive", or "computationally complex", they are thereby "impractical to perform in the human mind." *See*, *e.g.*, *Ex parte Akli Adjaoute*, Appeal 2018-007443, p. 11 ("the 'classifying' steps of claims 1 and 'modules' of claim 8 when read in light of the Specification, recite a method and system difficult and challenging for non-experts due to their computational complexity. As such, we conclude that one of ordinary skill in the art would not find it practical to perform the aforementioned 'classifying' steps recited in claim 1 and function of the 'modules' recited in claim 8 mentally."); *Ex parte Jean-Baptiste Tristan*, Appeal 2018-004459, p. 6 ("when read in light of the Specification, the claimed 'identifying a particular inference algorithm' is difficult and challenging for non-experts due to their computational complexity. As such, we conclude that one of ordinary skill in the art would not find it practical to perform the aforementioned 'identifying' step mentally.")

AI-related inventions tend to be complex and computationally intensive. Furthermore, some AI concepts are not easy to claim directly because the parameters are not interpretable. However, a practitioner should describe the elements of an AI claim as complex and computationally intensive in the specification to address a Section 101 rejection later. In fact, a practitioner can explicitly describe steps of AI-related inventions as impractical to perform in the human mind. By way of example, if an ML model enables an AI-type invention and takes a month to train on a supercomputer, then a practitioner can easily point out the computational complexity and impracticality of performing such a process in a human mind.

4. Organizing Human Behavior Rejections

The second most common category of abstract idea used in rejecting AI-related applications was the 'methods of organizing human behavior', which was cited in about 40% of the cases. This rejection is particularly common in the 3600 "Business Methods" art unit, where many business-related claims are examined.

The MPEP explains that the "methods of organizing human behavior" category includes the following subcategories:

- fundamental economic principles or practices (e.g., hedging, insurance, mitigating risk);
- commercial or legal interactions (e.g., agreements in the form of contracts, legal

obligations, advertising, marketing or sales activities or behaviors, and business relations); and

- managing personal behavior or relationships or interactions between people, (e.g., social activities, teaching, and following rules or instructions).

See MPEP 2106.04(a)(2). However, USPTO guidance explains that "not all methods of organizing human activity are abstract ideas (e.g., 'a defined set of steps for combining particular ingredients to create a drug formulation' is not a certain 'method of organizing human activity'), In re Marco *Guldenaar Holding B.V.*, 911 F.3d 1157, 1160-61, 129 USPQ2d 1008, 1011 (Fed. Cir. 2018). Second, this grouping is limited to activity that falls within the enumerated sub-groupings of fundamental economic principles or practices, commercial or legal interactions, and managing personal behavior and relationships or interactions between people, and is not to be expanded beyond these enumerated sub-groupings except in rare circumstances" *Id.*

In other words, examiners are not free to make up new sub-categories of the "methods of organizing human behavior" category of abstract idea. Accordingly, the primary method of overcoming an organizing human activity rejection is to avoid explicitly reciting words known to relate to these recognized methods of organizing human behavior. Many practitioners avoid using these terms in both the specification and claims to avoid getting applications sent to Art Unit 3600 in the first place.

Although many AI applications solve problems related to the types of activities considered "methods of organizing human behavior," it is usually possible to describe claims in a way that does not directly invoke these activities. For example, claims can focus on the processes performed directly by the AI model (i.e., computing vectors, text and images), as opposed to the final objective these outputs will be used for (e.g., creating a legal agreement).

5. Mathematical Concept Rejections

Surprisingly, the least common category of abstract idea used in rejecting AI-related applications was the "mathematical concept" category. It was cited in about 15% of rejections. AI-inventions are, of course, mathematical concepts in the broadest sense, however, claims covering such inventions need not–and indeed should not–explicitly recite mathematical formulas to avoid abstract idea rejections based on the mathematical concept group. For example, in an AI-related claim directed at training an ML model, consider reciting elements of a loss function descriptively, rather than the formula representing the exact loss function.

As another example, in an AI-related claim directed at inferencing, consider claiming the specifics of the input data, the transformation to the data made by the model, and the specifics of the output data, as well as any input data preparation steps if applicable, rather than the mathematical operations that affect such transformations of input to output. As yet another example, for an AI-related claim directed at a particular ML model architecture, consider claiming the inclusion of particular types of layers and connections therebetween, as well as the function of those layers, without reciting specific mathematical operations (e.g., dot products, summations, etc.) performed by those layers.

The MPEP notes, when discussing the seminal case *Diamond v Diehr*, that "[t]he Court's rationale for identifying these 'mathematical concepts' as judicial exceptions is that a "*mathematical formula* as such is not accorded the protection of our patent laws,'" MPEP § 2106.04(a)(2)(I)) (quoting *Diamond v. Diehr*, 450 U.S. 175, 191(1981)) (emphasis added). Note the emphasis on "mathematical formula" in that rationale. Consistent with this, the MPEP explains that:

> *When determining whether a claim recites a mathematical concept (i.e., mathematical relationships, mathematical formulas or equations, and mathematical calculations), examiners should consider whether the claim recites a mathematical concept or merely limitations that are based on or involve a mathematical concept. A claim does not recite a mathematical concept (i.e., the claim limitations do not fall within the mathematical concept grouping), if it is only based on or involves a mathematical concept.*

Id. (emphasis added). While the MPEP further adds that "a mathematical concept need not be expressed in mathematical symbols" to be recited in a claim, the PTAB has held on multiple occasions that a claim is not directed to a mathematical concept where the claims did not recite a specific algorithm or formula. *See*, *e.g.*, *Ex parte Akli Adjaoute*, Appeal 2018-007443 (October 10, 2019) (reversing a Section 101 and noting: "the specific mathematical algorithm or formula is not explicitly recited in the claims. As such, under the recent Revised 101 Guidance, the claims do not recite a mathematical concept."); *Ex parte Jean-Baptiste Tristan*, Appeal 2018-004459 (June 25, 2019) ("while the Specification identifies 'implement[ing] an inference algorithm' … the specific mathematical algorithm or formula is not explicitly recited in the claims. As such, under the recent Memorandum, the claims do not recite a mathematical concept."); *see also* PEG Example 38 ("The claim does not recite a mathematical relationship, formula, or calculation. While some of the limitations may be based on mathematical concepts, the mathematical concepts are not recited in the claims.").

6. Specific Structures and Practical Applications

As noted above, over 90% of the recent AI applications were rejected under step 3 (2A, Prong 2) for using generic computer components to perform the abstract idea. In many cases, this is because examiners treat this prong superficially.

As mentioned above, the Federal Circuit in *Recentive* has made it clear that "patents that do no more than claim the application of generic machine learning to new data environments, without disclosing improvements to the machine learning models to be applied, are patent ineligible under § 101." Thus, it is important to characterize claims as including something other than "generic machine learning". That is, it should be argued that claims represent an improvement to AI technology. Furthermore, it is often possible to overcome 101 rejections on this basis by:

- Arguing that the claim recites specific structures or processes that cannot be considered "generic computer components" or, more specifically, "generic AI" or
- Arguing that the claim is integrated into a practical application.

That is, in step 3 of the eligibility test it is determined whether any additional claimed elements other than the abstract amount to significantly more or integrate the claim into a practical application. Claims stating use of “a processor” and “memory” to execute AI functions will likely draw this rejection. If such a claim is coupled with the specification that describes the claimed functions can be performed in a general-purpose computer or processing device, a rejection under this Prong is almost inevitable.

It is noteworthy that in the many recent rejections, “neural networks”, “ML models”, and “AI models/techniques” were considered generic computer components. Therefore, to argue that claims recite a structure specific enough to overcome this element of a Section 101 rejection, a practitioner will likely be required to recite more specific architectural elements. There is some evidence that some more specific architectural elements (e.g., deep neural networks, recurrent neural networks) are starting to be considered generic. However, practitioners have found success citing even more specific architectural elements such as self-attention layers, layers with specific activation functions, U-nets, etc.

Alternatively, a practitioner can focus on the relationship between technical elements of a claim and improvements to a particular technology, Limitations that have been held to amount to significantly more or integrate a claim into a practical application include limitations that:

- Improve the functioning of a computer
- Include an improvement to the functioning of a computer can include an improvement to the functioning of an ML model. Thus, if an invention improves an ML model, the improvement should be included in the claims.
- Implement the judicial exception with a particular machine
- Include specific elements that perform the relevant steps in the claim. The USPTO has trained examiners to reject claims that appear to implement otherwise ineligible methods
- using generic AI or ML models. However, specific architectures should still be considered “particular machines”.
- Effect a transformation or reduction of a particular article to a different state
- Transform a particular thing. For example, ML models that transform or generate images, documents, or audio files can be considered to transform a particular article.

- Apply the judicial exception in some other meaningful way beyond generally linking the use of the judicial exception to a particular technological environment

Ultimately, the key to successfully arguing that a claim is applied to a particular technology in a meaningful way is whether the claims include technical elements that link a judicial exception to the technology in question. For example, RNNs and transformer models may utilize a sequence of ordered tokens or embeddings specific to a particular technological environment. Alternatively, citing specific sensor data, or specific outputs unique to a technological environment, along with architectural or procedural steps specific to a particular ML technology can help apply the invention to a particular technological environment.

7. Combinations of Additional Elements and Unconventional Solutions

Of the recently evaluated applications, 97% were evaluated and rejected under Step 2B of the patent eligibility test. However, as with Step 2A, Prong 2, many of these were initially superficial arguments by the examiner.

The fundamental question of Step 2B of the patent eligibility framework is whether *additional elements* (i.e., not those elements indicated by an Examiner as directed to the abstract idea) amount to an inventive concept that is significantly more than the judicial exception. Some elements of the Step 2B analysis are similar to Step 2A, Prong 2. For example, the identification of the “additional elements” carries over from the step 2 analysis, so it is critical to explicitly recite such additional elements when responding to a rejection.

However, there are two key differences that can be important in the context of AI/ML claims. First, Step 2B takes into account whether an element is routine, well-understood, or conventional. Second, Step 2B emphasizes consideration of the claim as a whole, including *combinations* of elements (including elements that individually could fall under a judicial exception).

The question of whether an element of routine or well-understood has been used to exclude additional elements from consideration in terms of integrating a claim into a practical application. However, it is important to incorporate unconventional elements into a claim even if they might fall under a judicial exception (e.g., unconventional mathematical techniques). Finally, an examiner is required to provide evidence if they claim that an element is conventional. *See*, PTO Memorandum, *Changes in Examination Procedure Pertaining to Subject Matter Eligibility, Recent Subject Matter Eligibility Decision* (*Berkheimer v. HP, Inc.*) (Apr. 19, 2018) at 3-4.

Unconventional elements can be considered in combination with other elements such that when combined, integrate the claim into a practical application. In other words, when responding to Step 2B rejections, it is important to explicitly point out combinations of elements that amount to an inventive concept, even if some of the elements could be considered judicial exceptions. Furthermore, it is useful to cite passages from the specification that show how the technical solution is unconventional.

For example, if an invention uses an AI architectural element in an unconventional way (i.e., to solve a task that usually involves a different architecture), the combination of the architectural element and the output can be considered an additional element that can be used to argue that the claim includes something more than an abstract idea.

8. Recent Changes Regarding Training Claims

The manner in which the patent eligibility steps are applied has been changing rapidly. For example, as discussed above with reference to drafting patent eligible claims, the USPTO guidance provides a training claim as an example of a claim that is not directed to an abstract idea at step 2 (2A, Prong 1). However, current practice in the USPTO has begun to change and more training claims are being rejected under Section 101.

Specifically, MPEP 2106.04(a)(1), example *vii* (based on Example 39 of the USPTO's "Subject Matter Eligibility Examples: Abstract Ideas") is directed to an eligible method of training a neural network. The example claim is reproduced below:

A method of training a neural network for facial detection comprising:

collecting a set of digital facial images, applying one or more transformations to the digital images, creating a first training set including the modified set of digital facial images;

training the neural network in a first stage using the first training set;

creating a second training set including digital non-facial images that are incorrectly detected as facial images in the first stage of training; and

training the neural network in a second stage using the second training set.

The USPTO's original guidance explained that the above claim is not directed to a judicial exception under Step 2 because, inter alia, "the claim does not recite any mathematical relationships, formulas, or calculations" and "the claim does not recite a mental process because the steps are not practically performed in the human mind."

This approach has been solidified based on the more recent July 2024 guidance provided by the USPTO. For example, the analysis of Example 47 states "The training algorithm is a backpropagation algorithm and a gradient descent algorithm. When given their broadest reasonable interpretation in light of the background, the backpropagation algorithm and gradient descent algorithm are mathematical calculations." That is, common AI training process are now being treated as mathematical and therefore will not be viewed as additional elements that can constitute additional elements in a patent eligibility analysis.

Accordingly, examiners at the USPTO continue to reject claims directed to training with a structure similar to MPEP Example *vii* (Example 39). From the study of 200 recent rejections, 30 applications that include claims directed to training were rejected. 24 out of the 30 applications were rejected under Step 2, because the training processes were claimed at a high level of generality or failed to provide improvement to computer functionality or technology. Therefore, claims directed to training should be drafted with the same considerations given above to be patent eligible as in other AI claims.

9. Written Description and Definiteness Rejections (§112)

As discussed above with reference to drafting the specification, 35 U.S.C. §112(a) includes three distinct requirements: 1) the written description requirement; 2) the enablement requirement; and 3) the best mode requirement. However, "best mode" rejections are very rare post AIA so emphasis should be placed on the first two.

The MPEP provides specific guidance for evaluating the written description requirement for computer-implemented inventions: "When examining computer-implemented functional claims, examiners should determine whether the specification discloses the computer and the algorithm (e.g., the necessary steps and/or flowcharts) that perform the claimed function … An algorithm is defined, for example, as 'a finite sequence of steps for solving a logical or mathematical problem or performing a task.' Microsoft Computer Dictionary (5th ed., 2002)." MPEP 2161.01(I).

These elements of Section 112 should be interpreted from the perspective of a person having ordinary skill in the art (POSITA), which may not be clear in a fast-moving field such as in AI. In general, the POSITA generally does not have the same level of expertise in the field as the inventors themselves. Therefore, it is important to differentiate between concepts that would be understood by an average person working in the field of AI and someone with the particular expertise of the inventors.

The primary strategy for responding to Section 112(a) rejections of AI claims is to point out portions of the specification that provide a description of the architecture of a model or the training paradigm. See the three levels of description of architecture and training recommended to be included in the specification above.

Regarding Section 112(b), an application is required to "particularly point out and distinctly claim" the subject matter which the inventor regards as the invention. In many cases, rejections under Section 112(b) relate to informalities such as mistakes regarding proper antecedent basis in a claim. These mistakes are often easily corrected.

Since AI claims are often expressed with a model performing a function, a common reason for rejections under Section 112(b) is the citation of generic structures which have been interpreted as means-plus-function elements under Section 112(f), with a subsequent determination that the specification does not provide sufficient corresponding structure. Guidance on establishing definiteness in the written description can be seen in section II.C.3.

In particular, it is worth repeating this tip: for every claimed function, provide an identification of the structure performing that function, a description of an algorithm; a description of what structural elements are used to perform the function; and a description of how exceptions are handled.

10. 2025 Patent Eligibility Study

(Frank Chau)

A study was made on 150 AI-based patent applications that were rejected under 35 U.S.C. § 101, and are allowed/issued in 2025. The study, along with the results of a similar study from 2023, confirms a clear pattern: examiners frequently categorize AI claims as mental processes, generic computer implementations, or as directed to well-understood, routine, and conventional (WURC) activities. Despite this, a meaningful number of AI-related patents have ultimately issued when applicants reframed their claims to emphasize concrete technical improvements.

<u>Step 2A Prong 1 – Abstract Idea Categories</u>

The most common category of rejection remains the mental process. Between 2023 and 2025, more than seventy percent of AI applications faced rejections on this basis. Specifically, 114 out of 150 applications (76%) were rejected under the mental process category. 36 out of 150 applications (24%) were rejected under the human activity category, and 31 out of 150 applications (21%) were rejected under the mathematical concept category.

- **Mental Process:**
    - 2023: 71% (141/200)

    - 2025: 76% (114/150)

      → **Trend:** Mental process remains the leading rejection category under this step and prong.

- **Human Activity:**
    - 2023: 42% (84/200)
    - 2025: 24% (36/150)

      → **Trend:** Human activity rejections dropped significantly. This suggests successful applicants are framing claims away from business methods or user activity steps.

- **Mathematical Concepts:**
    - 2023: 15% (30/200)
    - 2025: 21% (31/150)

      → **Trend:** Mathematical rejections increased slightly.

To respond effectively to mental process rejections, practitioners must avoid drafting claims that resemble mental acts such as predicting, identifying, or evaluating. Instead, these should be recast as technical operations. It is also effective to emphasize the computational complexity involved, making clear that the process cannot be performed in the human mind. Linking outputs to tangible system-level effects, such as improvements in latency, bandwidth, or accuracy, further strengthens eligibility.

Rejections based on human activity decreased substantially during the same period, falling from more than forty percent to just under one quarter. This change reflects a shift in applicant strategy: successful claims avoided framing inventions as business or administrative processes. Rather than presenting AI as a tool for advertising, contracts, or sales, practitioners described inventions as technical solutions. By anchoring claims in fields such as network optimization, device control, or dynamic user interface rendering, applicants were able to show that outputs directly modified technical systems rather than guiding human behavior.

Mathematical concept rejections, although less frequent than mental process rejections, increased in prevalence over the study period. Here, the lesson is clear. Explicit mathematical formulas and equations should generally be avoided in claim language. Instead, mathematical operations can be described functionally, and always tied to a technical benefit. For example, rather than reciting the explicit form of a loss function, a claim can describe computing loss values to update weights, while explaining how this improves training efficiency or accelerates inference.

Step 2A Prong 2 – Integration into a Practical Application

In 2025, 148 out of the 150 applications (99%) were rejected for having or using generic computer components to perform the abstract idea. For example, these generic computer components include "computer system/product, device, memory, processor, non-transitory computer readable medium, machine learning model." In some cases, computer components such as "encoder, classifier, decoder, transformer, neural network" are also rejected as generic computer components. 47 out of the 150 applications (31%) were rejected based on limitations reciting data gathering that amounts to insignificant extra-solution activity.

- **Generic Computer Components:**
    - 2023: 92% (184/200)
    - 2025: 99% (148/150)

    → **Trend:** Nearly all applications rely on "generic" computing (processors, ML models, neural networks). However, despite this, patents issued in 2025 show that **applicants successfully argued technical integration, or examiners accepted improvements in context**.
- **AI/ML Components are Not Enough by Themselves:**
    - 2023: 48 ML/AI mentions, 11 neural network mentions.
    - 2025: Broader spread: encoder, classifier, decoder, transformer, neural network — all still labeled "generic."

→ **Trend:** USPTO is treating even advanced AI architectures as conventional unless tied to a **specific technical improvement**.

- **Data Gathering = Insignificant Extra-Solution Activity:**
  - 2023: 36% (71/200)
  - 2025: 31% (47/150)

  → **Trend:** Slightly fewer "insignificant extra-solution" rejections survived to issuance, suggesting applicants re-drafted claims to avoid mere data collection and show **functional use of the data**.

Almost every AI application confronted rejections for relying on generic computer components. Even terms such as neural network, encoder, decoder, or transformer were often treated as generic. The path to allowance required practitioners to go beyond reciting a processor or computer. Successful claims specified how the processor was configured, for example by distributing updates across multiple memory banks or synchronizing them to reduce training latency. References to specific architectures such as CNNs, RNNs, or transformers—down to their subcomponents, like self-attention or U-net layers—also proved valuable. Crucially, claims were tied to system-level technical improvements, such as denoised images, reduced latency, or actuator control, rather than being framed as abstract computation.

A smaller but still significant portion of rejections stemmed from claims that were characterized as insignificant extra-solution activity, often because they appeared to focus on data gathering. The successful response was to demonstrate how the collected data was transformed and applied downstream. A claim that merely collected, analyzed, and displayed data was vulnerable; one that showed sensor data being used to modify actuator control signals or trigger real-time system responses was much stronger.

<u>Step 2B – Additional elements that amount to "significantly more" than the abstract idea</u>

In 2025, 132 out of the 150 applications (88%) were rejected under Step 2B. 110 out of the 150 applications (73%) were rejected for including elements that were deemed well-understood, routine, and conventional in the field.

- **High Rejection Rate Persists:**
  - 2023: 97% (193/200)
  - 2025: 88% (132/150)

    → **Trend:** Most AI applications continue to fail Step 2B initially, but the lower rate in the allowed set shows that some overcame §101 by showing a non-conventional technical implementation.
- **Well-Understood, Routine, Conventional Activity:**
  - 2023: 45% (89/200)
  - 2025: 73% of those rejected under 2B (110/150 total)

    → **Trend:** "WURC" findings remain a central hurdle. The fact that many of these still issued indicates **applicants successfully argued novelty of application context** or persuaded examiners with technical distinctions.

The WURC category of rejections became increasingly common, with nearly three-quarters of applications facing such challenges by 2025. Nevertheless, many patents still issued after applicants demonstrated that their inventions combined conventional elements in an unconventional way. Here, specification support was essential. Practitioners highlighted unique configurations or adaptations, and argued that ordered combinations of elements yielded new system effects. Some applicants cited benchmarks, technical publications, or standards to demonstrate non-routine implementations.

11. Key insights from the 2023 and 2025 studies

Several key insights result from considering results and trends from the 2023 and 2025 studies. Applicants should write claims to avoid framing as mental processes or human activity. Instead, they should anchor claims in the technical aspects of AI/ML models or in the specific technical details of another technical field. As described below, Applicants should also be careful to draft claims that show an inventive concept and that integrate any abstract ideas into a practical application.

Additionally, the specification itself continues to play a critical role, not only for subject matter eligibility but also for §112 compliance. Best practice is to provide three levels of description: a functional level, a high-level architectural description, and detailed technical explanation. Risky terms such as "predict," "encode," or "evaluate" should be defined with precision in a technical context. Flowcharts, pseudocode, training objectives, and concrete examples strengthen both eligibility arguments and enablement support.

In prosecution, practitioners should emphasize practical application when addressing Step 2A prong 2 rejections, and inventive concepts through ordered combinations at Step 2B. Examiner interviews are valuable for drawing attention to system-level improvements. Citing USPTO eligibility examples under MPEP §2106 and PTAB precedent can help distinguish claims from those deemed ineligible.

The overarching lesson from the rejection study is that AI applicants succeed when they reframe claims away from abstract ideas and toward technical integration. Applicants overcame §101 rejections not by convincing examiners that AI is inherently inventive, but by grounding claims in concrete technical improvements, moving away from abstract human or business activity framing, and tying AI models to a measurable change in system performance. Accordingly, inventions should be presented as improvements to computer systems, networks, or devices, supported by a specification that provides layered detail. By adopting these strategies, practitioners can significantly improve the chances of allowance for AI-related patents in the current USPTO landscape.

Eligibility Strategies

Framing Claims Away from "Human Activity"

In 2023, 42% of rejections were tagged as *human activity*. However, by 2025 that number dropped to 24%. We can infer that Applicants succeeded by removing business/administrative framing and re-drafting claims in terms of *technical system operation* rather than user behaviors or commercial processes.

Example Strategy:
Instead of “tracking user purchases,” claims were reframed as “configuring a distributed system to dynamically allocate storage based on transaction data” — shifting the locus of innovation into computing infrastructure.

<u>Anchoring AI/ML Models in a Technical Field</u>

Both 2023 and 2025 show that ML models, neural networks, encoders, and transformers were consistently treated as “generic computer components.” The way through was not to argue that “AI is novel,” but to show *how* the AI model improves a specific technical domain (e.g., image compression, chip layout design, network throughput, sensor fusion in robotics).

Example Strategy:
Showing improved accuracy, efficiency, or speed at a system level, demonstrating a hardware/software co-design (e.g., model implemented on GPU/ASIC in a way that reduces power draw), or claiming data structures and transformations that change the way the computer itself operates.

- Include **specific hardware or system elements** (sensor, actuator, GPU, communication interface).
- Show **interaction between components** (e.g., sensor → processor → neural network → actuator).
- Avoid phrasing such as “Using a model to classify data” (abstract mental process).
- Instead, try “Using a trained model to modify device operation or system output” (anchored in technology).
- **Practice Tip:** S*how the model as part of a technical workflow, not the end goal itself.*

<u>Integrating Into a Practical Application</u>

In 2023, 36% of applications were rejected for “insignificant extra-solution activity.” In 2025, almost as many (31%) were in the allowed set, showing fewer survived with data-gathering-only claims. We can infer that Applicants overcame this by tying data gathering to a downstream technical effect.

Example Strategy:
Not: “collecting sensor data.”
But: “collecting sensor data and adaptively modifying a control signal to reduce latency in a robotic actuator.”

- Avoid claims that look like mere **data gathering or reporting.**
- If data is gathered, specify **how it is transformed and used** to drive a technical result.
- Tie AI results to **system-level changes**: UI adaptation, actuator control, network configuration, image compression, etc.
- Explicitly describe how **data processing alters device behavior** or improves computer functioning.
- **Practice Tip**: S*hift the narrative from “data in/data out” to “system improved.” So “collecting sensor data” is weak; “collecting sensor data to reduce latency in actuator response” is strong.*

Showing an Inventive Concept

In 2023, 97% faced 2B rejections. By 2025, 88% did — but many still issued. Insight: The turning point was arguing against “well-understood, routine, conventional” (WURC) findings by citing technical publications or standards to show the claimed architecture is not routine; arguing the ordered combination of elements produces a new result, even if each element alone was conventional; and including specific system-level improvements (e.g., reduced network bandwidth, improved model training efficiency, faster inference).

- Preempt examiner assertions by **describing non-routine configurations** (e.g., parallelizing updates, adapting neural nets to hardware constraints).
- Argue **ordered combination** — even if parts are generic, together they yield a **new system effect.**

- Practice Tip: *Support with technical literature or benchmarks to show non-conventionality.*

Drafting Techniques That Made a Difference

Based on the results of the study, we also see that several drafting techniques were effective in providing material for overcoming patent eligibility rejections. First, draft independent claims that emphasize technical effect first, with dependent claims to capture variations. Second,. explicitly recite how generic components are configured differently (e.g., a "processor configured to parallelize model updates across memory banks" vs. just "a processor"). Finally, highlight integration into an existing technical field (e.g., medical imaging, autonomous driving, telecommunications) rather than standing alone as "an AI method."

Before and After

In accordance with the guidance described above, the following tables provide a side-by-side "Before vs. After" illustration that highlights how applicants successfully reframed their claims to overcome Abstract idea and "Generic Computer / Processor / Neural Network" rejections under §101.

Table 1. Claim Amendments to Overcome §101, Abstract Idea Rejections

Case	Before (Abstract/ Rejected)	After (Amended / Allowed)	Key Takeaway
Example 1	"Receiving input and classifying it using a model." (flagged as a mental process)	"A method at a computing device comprising: receiving input data; processing the input using a configured model; generating an output comprising a classification result; and adjusting device	Moves from abstract classification to device operation control, showing technical effect.

Case	Before (Abstract/ Rejected)	After (Amended / Allowed)	Key Takeaway
		operation based on the classification result.”	
Example 2	“Analyzing data to generate recommendations for a user.” (flagged as mental process / human activity)	“A method for providing a user interface in a communication system comprising: receiving sensor data; generating predictive recommendations; dynamically altering the interface configuration on a display device based on the recommendations; and transmitting updated control signals.”	Links abstract recommendations to UI/communication system performance, integrating into a technical field.

Table 2. Overcoming “Generic Computer / Processor / Neural Network” Rejections

Case	Before (Generic / Rejected)	After (Amended / Allowed)	Key Takeaway
Example 3	“A non-transitory computer-readable medium with instructions executed by a processor to process data.”	“…receiving sensor data from a device; encoding the data using a trained encoder; and outputting a transformed representation used to control downstream system operation.”	Converted generic storage + processor into a sensor-driven control pipeline with a real-world effect.

Case	Before (Generic / Rejected)	After (Amended / Allowed)	Key Takeaway
Example 4	"A method performed by a processor to process image data."	"…receiving image data; applying a neural network with specialized convolutional filters to reduce noise; generating a denoised image; and transmitting the processed image to a display device for rendering."	Elevated from generic processing to a technical imaging improvement tied to display hardware.
Example 5	"A computing platform comprising processor and memory for training a model."	"…a processor configured to distribute machine learning model training across multiple memory banks; dynamically synchronizing updates; and reducing latency in model convergence."	Changed generic processor into a non-routine processor-memory configuration improving ML training speed.

B. PTAB Decisions

(Ryan Phelan; Nick Transier)

In addition to Office Actions, it is also important to consider how AI is treated by the Patent Trial and appeal Board (PTAB). While there have been a number of PTAB decisions related to AI, there have been no recent decisions designated as precedential or informative. Recently, Examiners are refusing to consider PTAB decisions cited in Office Action responses unless those decisions are designated as precedential or informative, and there are very few such decisions to cite. *Ex parte Hannun*, 2018-003323 (April 1, 2019), is the latest decision designated as "informative". In *Hannun*, an AI claim was found patent eligible based on reciting a practical application of human speech-to-text translation using deep learning.

It is also notable that among recent decisions, when the PTAB is confronted with a combination of apparatus, method, and computer program claims, the PTAB has chosen to select the method claim as exemplary. As method claims are more susceptible to a determination of being directed to mental processes in Step 2 (2A Prong 1), a selection of the method claim as exemplary renders all claims more vulnerable to patent ineligibility despite the applicant undergoing the careful practice of presenting apparatus and device claims. A practitioner should argue against this as it allows the adjudicating body to bypass the default requirement of a claim-by-claim analysis.[3]

For example, as shown below in *Ex parte Philip E. Vasey* (Appeal 2022-001109), Rule 41.37(c)(1)(iv) allows for this practice by stipulating that, when an applicant does not provide separate arguments for different patent claims, the PTAB may select a single claim from a group and decide the appeal on the basis of the selected claim alone. Therefore, to avoid this type of blanket interpretation by the PTAB, the burden is on the applicant to argue the patentability of each claim type separately with regard to the matter of judicial exception during prosecution and appeal practice.

When it comes to Step 2A, Prong 2, recent PTAB decisions remain consistent with prior decisions such as *Ex parte Wataru Kimizuka* (Appeal 2018-001081), in that technological improvements must relate directly to the functioning of a computer or device to make a finding of integration into a practical application. Specifically, user benefits are not sufficient to make a finding of technological improvements. Rather, the test requires a technological improvement. This is a high standard for many applications and may render technological improvements a more difficult hurdle to meet at the PTAB compared to the Federal Circuit, where the court has not followed the same standard.

The selection of decisions below also illustrates how the PTAB addresses the question of what is well-understood, routine or conventional via assessment of the patent disclosure. For example, *Ex parte Akira Harada* (Appeal 2022-003628) illustrates the danger of using "black

[3] See, e.g., *A House Divided: Is the PTAB Ignoring the USPTO's Section 101 Guidance?*, available at https://ipwatchdog.com/2020/04/13/house-divided-ptab-ignoring-usptos-section-101-guidance/id=120589/

boxes" to describe computer parts. An insufficiency in disclosure with respect to computer components as shown in the drawings will not only pose a problem for the applicant with regard to Section 112, but also subjects the claimed invention to being deemed as well-understood, routine or conventional.

1. Ex parte Hannun (2018-003323)

In *Hannun*, the patent at issue related to "systems and methods for improving the transcription of speech into text." The claims included several AI-related elements, including "a set of training samples used to train a trained neural network model" as used to interpret a string of characters for speech translation. Applying the two-part Alice test, the Examiner had rejected the claims finding them patent-ineligible as merely abstract ideas (i.e., mathematical concepts and certain methods of organizing human activity without significantly more.)

The PTAB disagreed. While they generally agreed that the patent specification included mathematical formulas, such mathematical formulas were "not recited in the claims." (original emphasis). Nor did the claims recite "organizing human activity," at least because the claims were directed to a specific implementation comprising technical elements, including AI and computer speech recognition. Finally, and importantly, the PTAB noted the importance of the specification describing how the claimed invention provides an improvement to the technical field of speech recognition, with the PTAB specifically noting that "the Specification describes that using DeepSpeech learning, i.e., a trained neural network, along with a language model 'achieves higher performance than traditional methods on hard speech recognition tasks while also being much simpler.'"

2. Ex parte Aurélien Coquard and Christopher Bourez (2022-004679)

Regarding step 2 (2A Prong 1), the PTAB held that processing contract documents is considered a part of a legal interaction. Specifically, the PTAB concluded that claim 1 recites a legal interaction, which falls under the category of "methods of organizing human activity". This suggests that using language such as "contract document" in a claim makes it more vulnerable to a subject matter ineligibility rejection.

Regarding Step 2A, Prong 2, the PTAB found arguments with respect to "an improvement in computing" unpersuasive because the Specification describes using generic computing devices to process contract documents and the Appellant did not persuasively explain why using generic computing devices to process data makes claim 1 patent eligible. In other words, the claims did not represent an improvement in the computers themselves, but merely used generic computers as tools for performing a method of organizing human activity.

3. Ex parte Philip E. Vasey (2022-001109)

As mentioned above, the PTAB took a method claim as representative, despite the presence of apparatus claims. The claimed steps of "evaluating the first rule based on first inputted information to generate a partially customized document comprising the compulsory content elements, the first symbol element, the second rule, and the second symbol element" and "subsequently generating a fully customized document from the partially customized document, without reference to the document template, by evaluating the second rule based on second inputted information" can be performed in the human mind.

Under step 2 (2A Prong 1), the PTAB concluded the claims recite the judicial exception of mental processes. Under Step 2A, Prong 2 the PTAB considered the Appellant's purported improvement. Appellant argued that its claims integrate the judicial exception into a practical application by enabling a customized document to be generated on a remote computer that does not have access to the template documents. However, the PTAB found that solution was not necessarily rooted in computer technology as Appellant argued.

4. Ex parte Robert Kerr (2023-000284)

Under step 2 (2A Prong 1), the Appellant contended that requiring a process be stored electronically and performed by a processor overcomes the rejection because the claim excludes processing by the human mind. The PTAB found that to the extent some limitations cannot be performed mentally, this does not demonstrate error or demonstrate eligibility when at least some of the limitations recite steps that could be performed mentally. If a claim under its broadest reasonable interpretation covers performance in the mind but for the recitation of generic computer components, then it is still in the mental processes category unless the claim cannot practically be performed in the mind.

The PTAB stated that the Appellant provided insufficient detail of how the additional elements are implemented. Appellant's lack of a detailed disclosure of computer hardware or functional requirements and the lack of details describing a computer-specific implementation of the recited ML model and other functions (such as might have been indicated by inclusion of a detailed flow chart depicting unconventional computer operations and/or routines for performing each of the claimed steps), meant the details were well-understood, routine, and conventional.

5. Ex parte Martin Rupp (2023-000022)

In analyzing step 2 (2A Prong 1), the PTAB stated the claims merely recite performing certain calculations without detailing any particular hardware circuitry performing mathematical operations. Manipulation of the input data relates to the pre-Internet activity of performing mathematical computations to convert input data into equivalent output performance data. Accordingly, the PTAB found the cited steps are mathematical steps that can be performed with pen and paper by an operator to reduce the amount of data stored in memory as set forth in the claim. Furthermore, the claims, under their broadest reasonable interpretation, recite a mental process for organizing information through mathematical.

Under Step 2A, Prong 2 the PTAB reasoned that the Specification does not provide additional details about the general-purpose computer that would transform it into a specific computing device for converting input data from one form to another. Further, Appellant's identified improvements were held to relate to the abstract idea itself, not improvements to a technology or computer functionality. Under step 4 (Step 2B), the PTAB held that the Appellant failed to establish how converting input data into equivalent output performance data is distinguished from the conventional processor-implemented calculation of data.

6. Ex parte Akira Harada (2022-003628)

Under step 2 (2A Prong 1), the PTAB held that the claims were directed to a mathematical formula. Specifically, "a numerical data acquisition step of receiving numerical data derived from . . ." and "a standard error calculation step of calculating a standard error of numerical data at each data acquisition point or the data based on the numerical data" relate to mathematical relationships for processing data by receiving data and calculating the associated standard error.

Under Step 2A, Prong 2, the PTAB held that the additional element of claim 1 is a detector in a chromatography analysis device. However, the Specification did not refer to any chromatography analysis device, and the closest device to a chromatography analysis device appeared to be the analysis device in the Specification. Because the Specification merely illustrated the analysis device as a generic box and merely referred to—but did not describe in any detail— the device, then the claimed "chromatography analysis device" was well understood, routine, or conventional in the field.

C. Appeals Court Decisions

(Ryan Phelan, Nick Transier)

The Federal Circuit has not reviewed many cases involving AI. However, in one recent case, the Federal Circuit found that a machine learning claim element lacked sufficient enablement under 35 U.S.C. § 112 because both the claim itself and the written description failed to describe "how" the claimed invention implemented this element. *See In re Starrett*, (2023). While the decision is nonprecedential, the case reveals the Federal Circuit's analysis and treatment of AI-type claims with respect to Section 112.

An even more recent case, *Recentive Analytics, Inc. v. Fox Corp*. (2025)., relates to 35 U.S.C. § 101. It, along with *Realtime Data v. Array Networks*, (2023), adds to a growing list of Federal Circuit patent eligibility cases. The following table outlines a number of key Federal circuit cases related to patent eligibility for AI inventions:

Table 3: Selection of Federal Circuit Cases Applicable to AI

Case	Technology Claimed	Step 1: Abstract Idea?	Step 2: Inventive Concept?	Outcome	Key Distinctions
Recentive v. Fox (2025)[4]	Use of conventional ML models (e.g., neural nets, SVMs) to optimize event schedules and network maps	**Yes** – Applying generic ML to new data environments (scheduling, broadcasting) is abstract	**No** – ML usage was generic; no improvement to ML itself or to computing	Ineligible	Merely applying ML to a field (event planning) without improving ML or computing is not enough
Realtime Data v. Array Networks, (2023)[5]	Computer-implemented digital data compression methods and systems.	**Yes -** Manipulating information through generic data compression	**No -** claims lacked specific techniques, only functional results	Ineligible	No concrete technical improvement, functional claims, reliance on existing techniques
McRO v. Bandai (2016)[6]	Rules-based automation of facial animation for speech	**No** – Claimed a specific improvement to animation using data-driven rules	N/A	Eligible	Claimed how to achieve a result (specific rules, inputs, outputs), not just the result

[4] *Recentive Analytics, Inc. v. Fox Corp.*, 134 F.4th 1205 (Fed. Cir. Apr. 18, 2025).

[5] *Realtime Data v. Array Networks*, 2023 U.S.P.Q.2D (BNA) 901 (Fed. Cir., Aug. 2, 2023).

[6] *McRO, Inc. v. Bandai Namco Games America Inc.*, 837 F.3d 1299 (Fed. Cir. 2016).

Case	Technology Claimed	Step 1: Abstract Idea?	Step 2: Inventive Concept?	Outcome	Key Distinctions
Enfish v. Microsoft (2016)[7]	Self-referential data table for improved database performance	**No** – Claimed a specific, non-generic improvement to database architecture	N/A	Eligible	Claimed a new data structure that improved computer functionality
Koninklijke KPN v. Gemalto (2019)[8]	Detecting errors in data transmission using improved parity check method	**No** – Specific improvements to error checking techniques	N/A	Eligible	Technological improvement to error detection; not a generic application
Stanford v. Chinese Univ. of Hong Kong (2021)[9]	Haplotype phase prediction using statistical models	**Yes** – claims are directed to the use of mathematical calculations and statistical modeling	**No** – Recited steps were well-known, conventional, and routine math on a generic computer.	Ineligible	No new ML method; just applying known techniques to a new field (genomics)

[7] *Enfish, LLC v. Microsoft Corp.*, 822 F.3d 1327 (Fed. Cir. 2016)

[8] *Koninklijke KPN N.V. v. Gemalto M2M GmbH*, 942 F.3d 1143 (Fed. Cir. 2019)

[9] *In re Board of Trustees of the Leland Stanford Junior University*, 991 F.3d 1245 (Fed. Cir. 2021)

Case	Technology Claimed	Step 1: Abstract Idea?	Step 2: Inventive Concept?	Outcome	Key Distinctions
SAP America v. InvestPic (2018)[10]	Financial data analysis using statistical models	**Yes** – Abstract idea of analyzing data sets	**No** – Math + generic computing; no specific improvement	Ineligible	Similar to *Recentive* – using conventional math/ML tools on financial data
Customedia v. Dish (2020)[11]	Custom advertising system for digital video recorders	**Yes** – Data delivery and targeting	**No** – Claimed improvement was speed via generic computing	Ineligible	No technological innovation; just faster use of generic systems

1. *Recentive Analytics, Inc. v. Fox Corp.*, 134 F.4th 1205 (Fed. Cir. Apr. 18, 2025)

Recently, in *Recentive v. Fox*, the Federal Circuit addressed the eligibility of AI-related claims for the first time. *See Recentive Analytics, Inc. v. Fox Corp.*, 134 F.4th 1205 (Fed. Cir. Apr. 18, 2025). There, the Federal Circuit affirmed a district court decision holding that four patents, which applied machine learning techniques to optimize event schedules and television network maps, were invalid under 35 U.S.C. § 101 as directed to abstract ideas and lacking an inventive concept under the Alice/Mayo two-step framework. As the Court summarized:

[10] *SAP Am., Inc. v. InvestPic, LLC*, 898 F.3d 1161 (Fed. Cir. 2018)

[11] *Customedia Techs., LLC v. Dish Network Corp.*, 951 F.3d 1359 (Fed. Cir. 2020)

> *This case presents a question of first impression: whether claims that do no more than apply established methods of machine learning to a new data environment are patent eligible. We hold that they are not.*

Specifically, the Court found that the claims did not improve machine learning itself, nor introduce a new ML architecture or training method. Instead, the claims used standard ML techniques (e.g., neural networks, decision trees, support vector machines) in conventional ways to process inputs and generate outputs in a new domain. The asserted innovation—using ML to automate and optimize scheduling and mapping—was an abstract idea, akin to longstanding human planning activities.

The Federal Circuit's decision in *Recentive Analytics* was a first decision from the Federal Circuit surrounding artificial intelligence (AI) and patent eligibility under 35 U.S.C. § 101. The decision reinforces the principle that merely applying AI to a specific use case, without disclosing a technological improvement, does not rise to the level of patent-eligible subject matter.

This ruling offers valuable guidance to legal practitioners and inventors navigating AI-based patent applications. While the court acknowledged the growing importance of machine learning, it emphasized that innovations must offer more than the application of established techniques to a new data environment. As the court clarified: "patents that do no more than claim the application of generic machine learning to new data environments, without disclosing improvements to the machine learning models to be applied, are patent ineligible under § 101." *Recentive Analytics, 134 F.4th* at *1216.

The case originated from Recentive Analytics' claims that Fox Corporation used infringing software to schedule regional broadcasts, including NFL games. Recentive, known for deploying machine learning to assist the NFL in scheduling, held four patents across two families. These included machine learning training patents focused on optimizing schedules using historical data, and "network map" patents that determined how content should be displayed across geographic markets. *Id.* at *1208. Despite their claimed innovation, the district court—and ultimately the Federal Circuit—found the patents invalid under § 101.

The Federal Circuit's analysis was grounded in five main deficiencies:

1. **Generic Use of Machine Learning**: The court found that the claims did not protect any novel algorithm or AI architecture. Instead, they applied broadly defined machine learning models to existing scheduling problems, offering no advancement over conventional techniques.
2. **Lack of Technological Improvement**: Although Recentive asserted their inventions addressed technical challenges, the court held that dynamic updates and iterative training are fundamental features of machine learning and did not, in this context, amount to a technological breakthrough.
3. **Insufficient Implementation Details**: The patents failed to provide meaningful implementation guidance. Without specific algorithms or concrete steps, the claims were seen as aspirational rather than technically instructive.
4. **Field-of-Use Limitation**: The application of AI to television scheduling did not save the claims. Courts have consistently held that limiting an abstract idea to a specific field of use does not make it patentable under § 101.
5. **Performance Improvements Are Not Enough**: Gains in speed or efficiency, absent a corresponding technical innovation, were insufficient to convert the claims into patent-eligible subject matter.

In trying to salvage their claims, Recentive analogized their patents to past cases such as *Enfish*, *McRO*, and *KPN*, where software claims were upheld due to specific technological improvements. *Id.* at *1212. However, the Federal Circuit rejected these comparisons, finding that unlike those cases, Recentive's claims lacked detailed implementation and demonstrable technical benefit. Instead, the court found the claims more akin to *Electric Power Group* and *SAP v. InvestPic*, where data collection and analysis without technological improvement were held to be abstract. *Id.* at *1214.

Under Alice step two, which examines whether a claim includes an "inventive concept" sufficient to transform it into patentable subject matter, Recentive argued that real-time data processing and dynamic machine learning outputs satisfied this requirement. *Id.* at *1215. The court disagreed, concluding that these features were intrinsic to the field and did not represent an inventive application. *Id.*

The *Recentive* decision highlights a consistent trend in § 101 jurisprudence: courts remain wary of patents that merely invoke AI or machine learning without providing concrete technical innovation. To be eligible for patent protection, claims must do more than describe the automation of known methods using AI—they must improve the underlying technology itself. The ruling serves as a cautionary reminder: "do it with AI" is no more patentable than "do it on a computer" unless accompanied by specific, inventive technical contributions.

The Supreme Court's two-step framework described in *Alice* is used for determining whether a patent claim falls within an excluded category of laws of nature, natural phenomena, and abstract ideas. *See Alice Corp. v. CLS Bank Int'l*, 573 U.S. 208, 217-18 (2014) (citing *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 566 U.S. 66, 75-77 (2012)). According to *Alice* step one, it must first be "determine[d] whether the claims at issue are directed to a patent-ineligible concept." *Alice*, 573 U.S. at 218. If the claim is "directed to" an abstract idea, then the elements of the claim must be examined "to determine whether it contains an 'inventive concept' sufficient to 'transform' the claimed abstract idea into a patent-eligible application." *Alice*, 573 U.S. at 221 (citation omitted). "A claim that recites an abstract idea must include 'additional features' to ensure 'that the [claim] is more than a drafting effort designed to monopolize the [abstract idea].'" *Id*.

The USPTO issued guidance for this framework in its 2019 Revised Patent Subject Matter Eligibility Guidance, 84 Fed. Reg. 50 (Jan. 7, 2019) ("Revised Guidance") found in the Manual for Patent Examining Procedure ("MPEP") §§ 2104-2106. To determine if a claim is "directed to" an abstract idea, a patent examiner considers whether it recites: (1) any judicial exceptions such as laws of nature, natural phenomena, and abstract ideas (referred to as Step 2A, Prong One); and (2) additional elements that integrate the judicial exception into a practical application (referred to as Step 2A, Prong 2). *See* Revised Guidance at 52-55. If a claim (1) recites a judicial exception and (2) does not integrate that exception into a practical application, then a patent examiner considers whether the claim (3) provides an inventive concept by adding a limitation that is significantly more than the judicial exception or (4) merely appends well-understood, routine, conventional activities known to the industry, specified at a high level of generality (referred to as Step 2B). *See* Revised Guidance at 56; MPEP § 2106.05(d).

Alice Step Two – No Inventive Concept

With respect to *Alice* step two, the Court found that no "inventive concept" transformed the abstract idea into patent-eligible subject matter. Rather, the claims merely described generic computing environments and generic use of ML without explaining how the ML was technically improved or modified. The Court found that the dynamic updating or iterative training steps were inherent in ML and not novel in themselves: "[s]uch a position plainly fails to identify anything in the claims that would 'transform' the claimed abstract idea into a patent-eligible application."

In doing so, the Court distinguished from more favorable cases that are often applied in the software domain. In particular, the Court distinguished this case from *McRO, Inc. v. Bandai Namco Games America Inc*., 837 F.3d 1299 (Fed. Cir. 2016) and *Koninklijke KPN N.V. v. Gemalto M2M GmbH*, 942 F.3d 1143 (Fed. Cir. 2019), where claims recited specific technological improvements. In the Court's view, the patents at issue in *Recentive* merely applied ML as a tool without improving it or solving a technical problem in its operation.

The key holding of *Recentive* is thus that claims applying generic machine learning techniques to a new data environment (e.g., scheduling events or generating broadcast network maps), without disclosing improvements to the machine learning model itself, are patent-ineligible under § 101.

While *Recentive* addresses aspects of machine learning for the first time at the Federal Circuit, it is useful to put it in context with other key Federal Circuit decisions involving software-related Section 101 issues, particularly where machine learning or data processing technologies were at issue. Accordingly, Table 3 compares such cases with *Recentive*:

2. *Brightex Bio-Photonics, LLC v. L'Oreal USA*, 2025 WL 722445 (N.D. Cal. Mar. 6, 2025)

In *Brightex Bio-Photonics, LLC v. L'Oreal USA, Inc.*, the U.S. District Court for the Northern District of California invalidated a set of patent claims that allegedly involved artificial intelligence (AI) technology. Brightex had asserted U.S. Patent No. 9,842,358 against L'Oreal, claiming infringement of its invention titled "Method for Providing Personalized Recommendations," which was directed to improving facial cosmetics through computerized analysis. *Brightex Bio-Photonics, 2025 722445 at *1.* The representative claim described a method involving image capture of a user's face, computerized skin analysis using population-based comparisons, and the delivery of personalized skincare recommendations. Notably, Brightex emphasized that its invention utilized AI for skin condition analysis, allegedly improving the precision of treatment suggestions. *Id.* at *2.

L'Oreal moved to dismiss the suit under Rule 12(b)(6), contending that the claims were invalid under 35 U.S.C. § 101 for being directed to an abstract idea without an inventive concept. *Id.* at *3. Specifically, L'Oreal argued that the patent merely recited the abstract idea of recommending treatments based on severity rankings of skin conditions, using only generic computing components. *Id.* at *3. Brightex responded by pointing to elements like a "photo guide" as an innovative feature intended to improve facial image capture, arguing that such elements rendered the claims sufficiently technical to survive dismissal. *Id.* at *9.

The court disagreed. While the specification of the patent discussed AI-based techniques and benefits, the court emphasized that these features were not incorporated into the claims themselves. *Id.* at *16. As such, the court found the claims lacked the necessary specificity to show a technological improvement. The claimed "photo guide," according to the court, did not constitute a meaningful technical innovation but merely ensured proper image positioning—something well-known in the art. *Id.* at *18-19. Without demonstrating how any element provided a technological improvement to computer functionality or device operation, the court found the claims to be abstract and lacking an inventive concept. *Id.* at *20.

This decision underscores a critical lesson for patent practitioners drafting software- or AI-related claims. Courts continue to demand that patent claims—not just the specification—include technical features that demonstrate a concrete improvement to underlying technology. As emphasized in the *Brightex* ruling and reflected in Federal Circuit precedent, omitting such claim-level detail not only invites § 101 rejections during prosecution but also exposes issued patents to invalidation risks in litigation. Patent drafters should take care to explicitly integrate AI-related functionalities or device improvements into the claims to strengthen their enforceability and avoid the fate encountered in *Brightex Bio-Photonics*.

3. In re Starrett, 2023 **U.S.P.Q.2d 684 (Fed. Cir., Jun. 8, 2023)**

In *Starrett*, the Federal Circuit considered an appeal from the Patent Trial and Appeal Board (PTAB) regarding U.S. Patent Application 15/299,124 ("the '124 application"). The '124 application claimed an invention for maintaining "data structures representing categories of biological signals in a body such as a 'Nervous System' and a 'Sensory System.'" *Starrett,* 2023 U.S.P.Q.2d 684.

Claim 1 of the '124 patent recited, in part, a "machine learning" element directed to a specific "configuration," specifically: "[b)] configur[ing] to receive, relay, transmit, or distribute one or more signal [*sic*] wherein at least one signal comprising data representative of information about one or more biological body [*sic*] wherein the processing of biological systems data using at least one *machine learning* task intelligibly recovering perceived, experienced, remembered, or imagined imagery, sounds, or feelings as one or more computational, visual, auditory, textual, numeric, symbolic, coordinate, or haptic representation…" *Id.* (citing '141 application, claim 1.) (emphasis added).

During prosecution, the examiner rejected all claims for lacking enablement. The rejection was appealed to the PTAB, which affirmed the examiner. *Id*. The PTAB found that claim 1 was a type of genus claim that "contain[ed] forty-seven 'or' clauses, thereby allowing it to cover over 140 trillion embodiments." *Id.* In addition, while the patent applicant had argued that claim 1 was "fully enabled" by the patent application's "laboriously detailed" specification, the PTAB disagreed, finding such assertions merely conclusory. *Id.* at *2-*3. Finally, the Board noted that the patent applicant's contentions essentially amounted to "argu[ing] that if an apparatus is well-known . . ., then any function that [the inventor] claims for that apparatus is also fully enabled." *Id.* at *3 (citations omitted).

On appeal, the Federal Circuit affirmed, citing the Supreme Court's precedent regarding enablement in the recent *Amgen Inc. v. Sanofi* decision: "If a patent claims an entire class of processes, machines, manufactures, or compositions of matter, the patent's specification must enable a person skilled in the art to make and use the entire class. In other words, the specification must enable the full scope of the invention as defined by its claims. The more one claims, the more one must enable. *Id.* (citing 143 S. Ct. 1243, 1254 (2023)).

The Federal Circuit then applied this enablement principle to the '141 application, finding that "[h]ere, much is claimed, and little is enabled." *Id*. In particular, the Federal Circuit found particularly troubling the '141 patent application's failure to explain *how* the claimed features would operate without undue experimentation: The application's disclosure of a broad and abstract organizational structure used to accomplish the maintenance of augmented telepathic data amounts to little more than a "research assignment" requiring a skilled artisan to undertake undue experimentation to discover what types of devices are encompassed by the claim limitations and *how* they would function. *Id.* (citing *Amgen*, 143 S. Ct. at 1256) (emphasis added).

While the Federal Circuit did not specifically address the "machine learning" element of claim 1, it did find, more generally, that claim 1 was "rife with broad, vague concepts." *Id.* at *5. For this reason, the Federal Circuit invalidated the claim based on a lack of sufficient enablement. *Id.* The Federal Circuit also addressed the applicant's contention that the claimed features were "well-known" and, as a consequence, allegedly "fully enabled." *See id.* (discussing this aspect as part of a *Wands* factor analysis, of which consideration of "well-known" components is a part. *See In re Wands*, 858 F.2d 731, 737 (Fed. Cir. 1988)). However, the Federal Circuit found that whether a feature is "well-known" (or not) is but one factor of the Wands analysis and is not dispositive on its own. *See id.* Again, the Federal emphasized the importance of describing *how* the claim elements function. *Id.* (stating that "the Examiner's discussion of the Wands factors properly faulted the specification for failing to describe how the claim elements function.").

Ultimately, claim 1 (containing the "machine learning" element) was found non-enabled because the applicant had failed to describe how this (and other) aspects of the invention worked, and the applicant could not rely on the knowledge of a person of skill in the art to cure this defect, no matter how "well-known" such prior art elements were. As the Federal Circuit noted, "[a]lthough the knowledge of one skilled in the art is indeed relevant, the novel aspect of an invention must be enabled in the patent." *Id.* (citing *Auto. Techs. Int'l, Inc. v. BMW of N. Am., Inc.*, 501 F.3d 1274, 1283 (Fed. Cir. 2007)).

4. ***Realtime Data v. Array Networks,*** 2023 U.S.P.Q.2D (BNA) 901 (Fed. Cir., Aug. 2, 2023).

Enablement could become a more prominent area of focus with respect to computer-implemented inventions if, for example, Section 101 is resolved via legislation. This topic was recently previewed in *Realtime Data v. Array Networks*. 2023 WL 4924814 (Fed. Cir., Aug. 2, 2023).

In *Realtime Data*, the patent at issue generally related to computer-implemented technology (and not AI), where the claims recited methods and systems for digital data compression. The majority opinion, by Judge Reyna, affirmed a district court's decision invalidating the claims of the patents-at-issue as abstract ideas pursuant to Section 101. *Id.* In particular, the Federal Circuit agreed with the district court that the claims were directed to the abstract idea "of manipulating information using compression." *Id.* The court admonished the claims and specification, stating that "[w]e have determined that "the claim itself ... must go beyond stating a functional result" and that "the claim must "identify *'how'* the functional result is achieved by limiting the claim scope to structures specified at some level of concreteness, in the case of a product claim, or to concrete action, in the case of a method claim." *Id.* (citing *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC,* 967 F.3d 1285, 1302 (Fed. Cir. 2020)) (emphasis added).

Because the claims "failed to do this," the majority opinion held the claims to be invalid pursuant to Section 101. *Id.* at *8. In particular the majority found that "none of the claims at issue specifies any particular technique to carry out the compression of data—the particular rules for producing a smaller set of data out of a larger starting set." *Id.* Rather, the claims "all take the availability of compression techniques as a given and address the threshold matter of choosing to use one or more such available techniques." *Id.* The majority further faulted the abstract nature of the claims stating that "even as to making such a selection, the claims are directed to only abstract ideas, calling for unparticularized analysis of data and achievement of general goals." *Id.*

Judge Newman dissented, arguing that the proper lens for determining was not Section 101 but Section 112, in particular, enablement. *See id.* At *12 (Newman, J., dissenting) (stating that "This is properly an enablement case."). Judge Newman did not analyze the claim under Section 112. Rather she advocated that the proper review belonged under Section 112, and not Section 101: "I write separately to note once again that § 101 was never intended to bar categories of invention in this way. This judicial exception to eligibility is an unnecessary and confusing creation of the courts. This case is an example, for the enablement requirement of § 112 is better suited to determining validity of these claims than is the distortion of § 101. I respectfully dissent, and would remand for determination of validity under § 112." *Id.* (Newman, J., dissenting)

She ended her dissent by noting that "[e]ligibility law has been called a 'morass of seemingly conflicting judicial decisions'" (citations omitted) and that "[w]e should not wade further into this morass." *Id*. "This case is another example that conforms with our flawed precedent. I respectfully dissent. I would remand for a determination of validity under § 112 and, if applicable, §§ 102 and 103." *Id.*

In view of the *Starrett* decision, practitioners should endeavor to explain sufficiently in the written description the specific aspects of ML (and other computer-implemented invention features). In particular, practitioners should endeavor to describe *how* a claimed computer-implemented invention (e.g., an AI invention) operates or otherwise works.

Finally, in view of the *Realtime Data* decision, practitioners can get a preview of what may yet come in the event of a legislative change to Section 101 (or a ruling from the Supreme Court bringing about the same), where the new invalidity battleground is not Section 101, but instead Section 112.

D. District Court Decisions

(Ryan Phelan; David Pointer**)**

The following set of AI-related cases are from district courts of different geographic locations.

1. *Health Discovery Corp. v. Intel Corp.*, 577 F. Supp. 3d 570 (W. D. Tex. 2021)

In this case, Intel filed a motion to dismiss Health Discovery Corp (HDC)'s complaint on the grounds that the claims are invalid under 35 U.S.C. § 101. HDC's complaint alleges that Intel infringed HDC's patents related to using learning machines, such as Support Vector Machines (SVM), to identify relevant patterns in datasets and to identify a selection of features within the datasets that best enable data classification. The asserted patents were directed to feature ranking, selection, and reduction using SVM to facilitate a Recursive Feature Elimination (RFE) process on a large dataset.

Under step 1 of the eligibility test, the Court reasoned to follow the guidance provided from prior cases that analyzed patents with similar subject matter. After selecting two similar cases, the District Court asserted that the specification merely describes improving a mathematical analysis used by conventional systems. The specification explained how conventional systems reduce a feature size in data sets by ranking and eliminating features according to correlation coefficients. Similarly, the asserted patents involve ranking and eliminating features using SVM-RFE, which the court characterized as a purportedly novel, but mathematical technique. As such, the District Court reasoned that the claims are directed to an abstract mathematical concept of SVM-RFE.

Under step 2, the District Court stated that HDC's complaint failed to sufficiently allege an inventive concept. The Court stated that improving data quality was an unpersuasive argument, and the Court also considered that the claims were not limited to a particular field of invention. Accordingly, the Court granted the motion to dismiss.

2. *Pavemetrics Sys. v. Tetra Tech*, 2021 U.S. Dist. LEXIS 117651, 2021 WL 2548959 (C.D. Cal. 2021)

In this case, Tetra Tech moved for a preliminary injunction to enjoin Pavemetrics from importing, using, and selling their "Laser Rail Inspection System" (LRAIL) products on the grounds of patent infringement. Tetra Tech's '293 patent generally relates to a three-dimensional railway track inspection and assessment system that collects and processes data during and/or after a high-speed assessment of a railway track. In 2018, Pavemetrics began developing AI-based deep learning algorithms using convolutional neural networks to identify defects in railway tracks in their LRAIL products.

Ultimately, the District Court concluded that Tetra Tech could not demonstrate a likelihood of success for a preliminary injunction because of substantial questions regarding infringement and invalidity. With regard infringement, the Court noted that substantial questions remain because Tetra Tech relies the "gradient neighborhood" limitation being reflected on a prior design of Pavemetrics' LRAIL product from two years prior to the issuance of the asserted patent. At that time, Pavemetrics significantly changed how LRAIL processed data when it switched to detecting missing features using deep neural networks. The Court agreed that Tetra Tech had not provided sufficient evidence to indicate Pavemetrics' "deep neural network" design within the current product meets the "moving gradient neighborhood like a sliding window over the 3D elevation data using the processor," as recited in claim 1.

With regard to invalidity, the Court determined there are substantial questions because Pavemetrics alleged that one of its prior designs anticipates claim 1 of Tetra Tech's '293 Patent. For these reasons, the Court denied the motion for preliminary injunction.

3. *IBM v. Zillow Grp., Inc.*, 2022 U.S. Dist. Lexis 41831, 2022 WL 704137 (W. D. Wash. 2022)

In this case, Zillow filed a motion to dismiss IBM's claims of patent infringement on the grounds of ineligible subject matter under 35 U.S.C. § 101. The '676 patent generally relates to a method of annotating response sets via an adaptive algorithm and the supplied annotations are used for a visualization system that presents resource response results. In the specification, the '676 patent states that the "*s*ystem discovers contexts and context attributes among users which can be used predictively," by using "a highly specialized and optimized combination of supervised & unsupervised logic along with" automated entry of learned results. Col. 19, Lines 39-44.

Although the specification of the '676 Patent appears to describe improvements relating to computer and/or search engine functionality, the Court focused on the subject matter recited in the claims for step 1. The Court found that the claimed subject matter was directed to an abstract idea because the claim language was result-oriented and recited a process that could be performed with a pen and a paper. The claims failed to recite the inner functionality of the invention.

Under step 2, the Court concluded that the claims failed to satisfy the second Prong of the test for a couple of reasons. First, the specification failed to provide a description of the alleged inventive concepts offered by IBM. Second, the Court determined that the claims in the '676 patent do not provide a specific, discrete implementation of the abstract ideas for applying an ordering and annotation function, mapping the user context vector with the resource response set, or generating an annotated response set. Accordingly, the Court concluded the '676 Patent is invalid under 101 and the related patent infringement count is dismissed for failing to state a claim upon which relief can be granted.

E. Detectability

(Michael Carey; Sumon Dasgupta; Thomas Burton**)**

In addition to issues that arise in court, AI related patents pose unique challenges related to detecting infringement. Generally, a patent is more valuable if infringement is easily detectable. However, many AI patents relate to software, and in some cases, to a configuration of software parameters that is not easily discernible or interpretable. This can create challenges in detecting infringement.

Detectable features are observable from the use, appearance, or construction of a product or process embodying an AI invention. In some cases, infringement may be detected through using a publicly available product, reviewing publicly available documents (e.g., academic publications, market materials, specifications, etc.), through reverse-engineering, or through application to technological standards.

When drafting an AI-related patent application, significant attention should be paid to identifying and claiming any features that are observable at inference (i.e., during use). For example, an AI related invention may include prompting a user for key inputs, and providing an output based on those inputs. The relation between the inputs and outputs may be observable even if the inner workings of the system are not observable. For example, AI features that are part of an autonomous device (e.g., vehicle, robot, game, etc.) or a device that involves learning (e.g., an intelligent thermostat) may be detectable in the operation of the device. Moreover, recent advances in black-box auditing techniques enable detection of AI system behaviors without direct access to the underlying models. These techniques can be leveraged to detect patent infringement by analyzing system outputs.

AI features may also be detectable based on publicly available documents. For example, because AI is a rapidly developing field that involves many academic contributions, many inventors publish their work in academic journals, conferences, and archives. In another example, companies may disclose features of an AI model in other ways such as user guides and advertising materials.

AI features may also be detectable via reverse-engineering a product. This is made more challenging by the prevalence of Software-as-a-Service (SaaS) AI products. That is, it may be challenging to reverse-engineer a product that is not publicly available. However, in some cases, AI products use publicly available data or publicly defined data structures that expose the inputs or the outputs to inspection. Moreover, methods exist for "attacking" an AI model hosted on a network. Using such techniques, an entity can steal a high-value AI model with very limited resources, which could significantly harm the interests of the AI product owner. Recent advancements in neural network watermarking techniques provide a method to detect unauthorized use of or theft of AI models or portions thereof. These watermarking techniques can be integrated into patent claims as a way to make infringement more detectable.

AI features can be detected at several different stages including data collection, training, testing, and operation. In some cases, different features may be exposed by attempts at distributing the process via federated learning, distributed data processing, distributed training, edge ML, and other emerging technologies that alter the workload balance between edge devices and centralized (e.g., datacenter, cloud) compute resources. In some cases, communication between the edge devices and the centralized cloud compute resources is much easier to capture and investigate via technical methods than the centralized compute resources themselves. In addition, with the rise of federated learning and distributed AI systems, new patterns of detectability emerge with edge devices. Therefore, when claiming ML inventions, pay particular attention to the dataflow between centralized compute resources and edge devices. For distributed AI systems, focus patent claims on the observable effects in distributed architectures rather than the internal implementation details.

Some features of AI patents may also be detected by virtue of relating to technological or communication standards. In many cases, standards are needed for the adoption and compatibility of new technologies. Standard bodies, such as American National Standards Institute (ANSI), Institute of Electrical and Electronics Engineers (IEEE), or the Third Generation Partnership Project (3GPP) have different patent policies. For example, these patent policies may request patent holders to disclose information regarding patents or patent applications that are relevant to the standard. Some may require that any license issuing from these patents be granted under fair, reasonable and non-discriminatory terms (FRAND license). Some standard bodies may even request that a license must be royalty free. Thus, it may be worth contributing patented technologies to a standards body standard.

Finally, the growing field of "Explainable AI (XAI)" offers new possibilities for making AI systems more transparent and their operations more detectable. In particular, AI developed to meet XAI guidelines by design may reveal the inner workings of such AI models, making their decision-making processes more visible and understandable.[12] Since such AI has model features that are detectable, patent claims that include explainability features can improve both the value of the patent while also addressing regulatory concerns related to AI transparency.

IV. Other Considerations

A. Ethics for AI Inventions

(Thomas Burton; Sumon Dasgupta**)**

With the emergence of AI-enabled products and tools such as ChatGPT and Bard, the US government and other countries are introducing guidelines and legislation to address other ethical concerns related to AI. For example, in October 2022, the US government published the "Blueprint For An AI Bill of Rights: Making Automated Systems Work for the American People "(hereafter, the "Blueprint") to provide companies guidance to address potentially inherent ethical risks of AI enabled systems. The Blueprint is not law. However, it can be a signal of potential government action (e.g., laws, rules, regulations).

[12] See, Defense Advanced Research Projects Agency. (n.d.). *XAI: Explainable Artificial Intelligence*. DARPA. Retrieved July 28, 2025, from https://www.darpa.mil/program/explainable-artificial-intelligence.

Likewise, the European Union Artificial Intelligence Act ("EU AI Act") that was enacted 1 August 2024 (and expected to fully effective by 2027) also addresses ethical risks characterized as "unacceptable" or "high risk" under the Act.

As a corollary, patent practitioners now have other ethical issues to consider when preparing a patent application directed to an AI related invention. This section provides patent drafting practice tips for avoiding such ethical issues. For example, patents should not employ discriminatory language to avoid ethical issues, such as using "binary" pronouns like he or she that may not be viewed as being inclusive of all. Such language can have negative public impact.

AI principles under the Blueprint, EU AI Act and other models, such as the recent G7 code of Conduct include:

1. Safe and Effective

An "automated system" that uses AI to "determine outcomes, make or aid decisions, inform policy implementation, collect data or observations, or otherwise interact with individuals and/or communities" (including "people connected by affinity, identity, or shared traits") should be safe and effective based on their intended use. In particular, such automated systems should be "developed with consultation from diverse communities" as well as "domain experts" and undergo pre-deployment testing, risk identification and mitigation, and ongoing monitoring" to support that they are safe and effective.

To illustrate why this principle is important, the Blueprint highlights real examples of AI enabled automated systems that violated this principle. For example, a proprietary model system was developed to predict the likelihood of sepsis in hospitalized patients and was implemented at hundreds of hospitals around the country. An independent study showed that the model predictions underperformed relative to the designer's claims while also causing "alert fatigue" in certain patients by falsely alerting likelihood of sepsis. Thus, this AI enabled prediction model was ineffective and caused harm to certain patients that were falsely identified.

Practice Tip: In preparing a patent application, patent practitioners should always avoid making any definitive statements in the patent application that the underlying invention "is safe", "error free" or any other statement that could be used to support a product liability claim against your client (the patent owner) or lead to public scrutiny of the effectiveness of your client's invention. This is especially true of an AI enabled invention where the boundaries of the AI invention are not yet fully tested, such as the case of the above AI enabled sepsis prediction model system. To avoid such product liability issues, the patent practitioner should counsel his client not to make any definitive claims of the AI system's effectiveness.

Instead, provide facts related to the data sets used to train the AI system and use safety terms such as "inhibits safety issues" (such as with outbreaks of sepsis in a hospital in which the AI system is deployed) and/or "enhances the likelihood of an accurate prediction." Such patent preparation practices are aligned with this Blueprint's principle for Safe and Effective Systems.

2. Algorithmic Discrimination Protections

Users of "automated systems" should be protected against algorithmic discrimination. Such systems should be designed to include "proactive equity assessments" and "use of representative data" to protect against bias and ensure equitable treatment.

To illustrate the necessity of this principle, the Blueprint provided the following real example. The U.S. Transportation Security Administration (TSA) deployed certain body scanners at airport checkpoints that required the operator to select a "male" or "female" scanning setting based on the passenger's sex. But this manual setting selection introduces bias into the body scanner system based on the operator's perception of the passenger's gender identity. These scanners are more likely to flag transgender travelers as requiring extra screening done by a person. TSA has recently announced plans to implement a gender-neutral algorithm while simultaneously enhancing the security effectiveness capabilities of the existing technology.

Practice Tip: In preparing a patent application directed to an AI enabled system (such as a body scanner that purports to implement a "gender-neutral algorithm", the patent practitioner should counsel the inventor(s) on the need to ensure the underlying algorithm is not discriminatory. The practitioner should ask if the system assesses data associated with people (as opposed to data associated with a substance, an object, an animal, a car, a building, etc.) and, if so, ask how the system takes "proactive" measures to "enhance the effectiveness" of the system to operate equitably so that the patent application can disclose those measures (which may aid in identifying potential elements to claim) and inhibit third party claims of algorithmic discrimination. Moreover, the practitioner should question the inventor(s) on the data sets used to train the applicable AI enabled system so that such data sets could be disclosed to further aid in preventing third party claims of bias.

3. Data Privacy

People should be able to control how their data is used, and they should not be subjected to abusive data practices. For example, the Blueprint describes ensuring that data collection conforms to reasonable expectations and that only data strictly necessary for the specific context is collected. Designers, developers, and deployers of automated systems should notify and seek permission with regard to data usage, and respect decisions regarding collection, use, access, transfer, and deletion of private data in appropriate ways and to the greatest extent possible. Systems should provide clarity about user choices, and not obfuscate user choice or burden users with defaults that are privacy invasive. For example, consent requests should be brief and understandable. Enhanced protections and restrictions for data and inferences related to sensitive domains (e.g., health, work, education, criminal justice, and finance), and for data pertaining to youth should only be used for necessary functions.

Practice Tip: Be cautious about describing the types of data that are used in some types of AI applications. It is almost second nature to the patent practitioner to broaden the scope of an application, but caution should be exercised with respect to data, and particularly in sensitive domains. Unwanted implications of unauthorized and/or unexpected data usage should be avoided in patent applications. Furthermore, bear in mind that while the Blueprint is non-binding, other information-related restrictions are in place. For example, the Health Insurance Portability and Accountability Act of 1996 (HIPAA) is a federal law that describes standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge.

4. Notice and Explanation

Users should know why and how an AI system made its determination. People should have an understanding that an automated system is being used and understand how and why the automated system contributes to outcomes that impact them. For example, it may be necessary to provide generally accessible plain language explanations (including clear descriptions of the overall system functioning and the role automation plays), notice that such systems are in use, who is responsible for the system, and explanations of the outcomes. In other words, people should timely receive notice as to how and why an outcome that impacts them was determined by an automated system.

An example provided in the Blueprint describes that "a lawyer representing an older client with disabilities had been cut off from Medicaid-funded home health-care assistance couldn't determine why, especially since the decision went against historical access practices. In a court hearing, the lawyer learned from a witness that the state in which the older client lived had recently adopted a new algorithm to determine eligibility. The lack of a timely explanation made it harder to understand and contest the decision."

Practice Tip: One of the many applications of AI can include the determination of a decision that impacts a user. In such cases, it may be prudent to describe in the patent application, that the user is notified in some fashion of the decision of the AI, that the AI is being used, and so forth. It may also be prudent to inform your client of the above guidance to mitigate potential issues in the future.

5. Human Alternatives, Consideration, and Fallback

Finally, the Blueprint sets forth the principle that an "automated system" should provide people a choice to opt out of AI decision-making and have "fall back" access to a human if the system has an error, fails, or they want to challenge a decision made by the system. The Blueprint highlights that "automated systems with an intended use within sensitive domains, including, but not limited to, criminal justice, employment, education, and health, should additionally be tailored to the purpose" in addition to incorporating "human consideration for adverse or high-risk decisions."

To illustrate the problems this principle seeks to address and protect against in AI systems, the Blueprint identified several examples, including the following. A fraud detection system for unemployment insurance distribution incorrectly flagged entries as fraudulent, leading to people with slight discrepancies or complexities in their files having their wages withheld and tax returns seized without any chance to explain themselves or receive a review by a person. A patient was wrongly denied access to pain medication when the hospital's software confused her medication history with that of her dog's. Even after she tracked down an explanation for the problem, doctors were afraid to override the system, and she was forced to go without pain relief due to the system's error. A large corporation automated performance evaluation and other HR functions, leading to workers being fired by an automated system without the possibility of human review, appeal or other form of recourse.

Practice Tip: Preparing a patent application directed to an AI system trained for a particular use case (e.g., unemployment insurance fraud detection) typically would not require the applicant to disclose or include an "opt out" feature for an individual to contact a human to question an output of the AI system. But where the output of an AI system is based on accessing data that is personal to an individual, a patent practitioner should be alert to the harm that the AI system may cause the individual when the output is based on data that is not error-free but includes "slight discrepancies or complexities" that could lead the AI system to output a false positive output.

Accordingly, where an invention is directed to an AI system that processes personal data (e.g., AI system for insurance fraud detection, determining patient medication prescription, or automated employee performance evaluation), a patent practitioner should advise his client that such an AI system may be viewed by the public and government agencies as providing "high-risk" decisions for persons. To mitigate any negative sentiment from the public based on the patent application being directed to such a "high-risk" AI system, the patent practitioner may recommend that the applicant/inventor add potential further features or embodiments to the AI

System for an individual to address any “high-risk” decision relating to the individual. Although likely not patentable subject matter, the patent practitioner may recommend that the applicant also include an embodiment where an individual may alternatively contact the owner or operator of the AI System to dispute any output or decision of AI system. Such an embodiment may be positively viewed by the public as in line with the objective of the Blueprint’s “Human Alternative” or “Fallback” principle.

6. G7 Code of Conduct

In addition to the Blueprint, AI ethics are being considered by a number of organizations around the world. For example, the Group of 7 (G7) countries recently adopted a voluntary AI Code of Conduct. The G7 Code is largely directed at companies that operate in G7 countries.

This 11-point ethical code seeks to ensure that AI is both safe and dependable on a global scale. It is intended to furnish voluntary guidance for organizations engaged in the development of cutting-edge AI systems, encompassing advanced foundation models and generative AI systems. The primary objective of this code is to harness the advantages of AI while effectively addressing the associated risks and challenges. It places a notable emphasis on urging companies to implement measures that identify, assess, and mitigate risks throughout the entire lifecycle of AI. Furthermore, it compels companies to confront and rectify incidents and patterns of misuse that may arise after the deployment of AI products in the market. Additionally, the code encourages companies to disseminate public reports detailing the capabilities and limitations of AI systems, as well as their usage and potential misuse. Furthermore, it underscores the importance of investing in robust security controls to ensure the responsible development and deployment of AI technologies.

7. European AI Act

The European Union's AI Act represents the most comprehensive regulatory framework for AI to date and should also be considered when drafting AI patents. The EU AI Act categorizes AI systems based on risk levels and imposes varying obligations accordingly, which may impact patent strategy.[13] For example, risks characterized as "unacceptable" under the EU AI Act include AI systems that (i) "use subliminal or deceptive techniques to distort decision-making and impair informed choices", (ii) "take advantage of individuals' vulnerabilities related to age, disability, or socio-economic status", (iii) "evaluate or classify individuals based on their social behavior or personal traits, leading to detrimental treatment", (iv) "assess the likelihood of an individual committing a crime based solely on profiling or personality traits", or "infer sensitive attributes (e.g., race, political opinions) from biometric data, except in specific lawful contexts". Risks characterized as "high risk" under the EU AI Act include AI systems that are (1) important for product safety under the Union's harmonized legislation on product safety; and (2) can significantly affect people's health, safety, or fundamental rights in specific use cases listed in the AI Act.

In preparing a patent application directed to an AI enabled system, the patent practitioner should counsel the inventor(s) on the need to ensure that the specification and claims do not disclose any features or embodiments of the AI enabled system that may be characterized as "unacceptable" or "high risk" to avoid potential liability under the EU AI Act.

8. Summary

Although the "ethical risk mitigation" principles set forth in the Blueprint are not mandatory for companies or individual inventors to follow when implementing an AI system, the EU AI Act presents similar ethical concerns with AI systems classified as "unacceptable" and "high risk". Accordingly, patent practitioners would be wise to address these principles with their company or individual clients while preparing a patent application directed to AI system that may be viewed as having ethical concerns. By following the above practice tips, a patent

[13] See, *AI Act: The first legal framework on AI worldwide, with a risk-based approach*. European Digital Innovation Hubs Network. Retrieved July 28, 2025, from https://european-digital-innovation-hubs.ec.europa.eu/knowledge-hub/european-ai-innovation-ecosystem/ai-act-first-legal-framework-ai-worldwide-risk-based-approach

practitioner can prepare a patent application for an AI system that aid the client to address any potentially inherent ethical risks in the AI system that the five principles seek to protect the public against.

B. Considerations for the EPO

(Alexander Korenberg; Edoardo Mirabella; Henry Aldridge; Mike Jennings)

The EPO assesses inventions by using the so-called two-hurdle approach. The first hurdle aims to establish whether the claimed subject matter is excluded from patentability or not. The second hurdle constitutes substantive examination and, among other goals, aims to establish whether the claimed subject-matter is inventive and whether the requirements of sufficiency are met. When considering AI-related inventions, the first hurdle is usually trivial to overcome. The second hurdle is where difficulties can lie.

1. Eligibility

The EPO considers AI and ML algorithms to be mathematical methods *per se*. Mathematical methods are not considered technical by the EPO, and according to the European Patent Convention (EPC), mathematical methods claimed *per se* (i.e. without any technical means such as a computer) are excluded from patent protection (i.e. they are included in the non-exhaustive exclusion list of excluded subject matter in Article 52 of the EPC).

While this may appear to suggest that AI-focused inventions are not patentable at the EPO, this exclusion is usually trivial to avoid. The presence of any technical means (e.g., a computer) is sufficient to overcome this first hurdle.

Recommendation

Include at least one technical feature (e.g. a computer) in the claims. This can be done implicitly, such as by claiming a *computer-implemented* method.

2. Substantive Examination

Inventive Step

For AI/ML-related inventions, claims that overcome the first hurdle are checked for any features that contribute to the technical character of the invention. The EPO is very comfortable with mixed-type inventions that include technical features (e.g. a computer) and features that would in isolation be considered non-technical (e.g. an AI/ML algorithm), and all features that contribute to the technical character of the invention are considered. Claims should not be rejected merely because one feature is based on an algorithm or is implemented in software. Each invention is considered as a whole with an attempt to identify all features that contribute to solving a technical problem and therefore contribute to the technical character of the invention. All features that contribute to technical character must be considered when assessing inventive step.

The assessment of inventive step for such mixed-type inventions is carried out via the so-called COMVIK approach, which prescribes how to apply the problem-solution approach in the presence of technical and non-technical features. While all claim features are taken into account for the assessment of novelty, the COMVIK approach requires that only those differences from the closest prior art that contribute to technical character are considered for inventive step. In practice this means that non-technical features (e.g. the steps of a business methods) are often entirely ignored when considering inventive step.

Importantly, however, non-technical features are considered by the EPO if, in the context of the invention, they contribute to the technical character of the invention by contributing to producing a technical effect. The steps of a business method always fail this test, but AI/ML algorithms may not.

Under EPO practice, there are two ways in which an AI/ML algorithm (or, in fact, any mathematical method) can be considered to contribute to the technical character of the invention by being:

- 1) Applied in a field of technology; or

- 2) Adapted to a specific technical implementation.

Where an AI or ML algorithm serves a technical purpose, the steps of generating the training set and training the algorithm may also contribute to the technical character of the invention, if they support achieving that technical purpose.

Application to a field of technology

The application to a field of technology must be specific, e.g. a generic purpose such as "controlling a technical system" is not sufficient. Moreover, the claim must be functionally limited to the technical application, i.e. not encompassing further non-technical applications.

Examples of technical contributions of a mathematical method are (cf. Guidelines for Examination in the European Patent Office, G-II 3.3 and G-II 3.3.1): controlling a specific technical system or process, e.g. an X-ray apparatus or a steel cooling process; digital audio, image or video enhancement or analysis, e.g., de-noising, feature detection in a digital image, or estimating the quality of a transmitted digital audio signal; classifying digital images, videos, audio or speech signals based on low-level features, e.g. edges or pixel attributes of images; and using a neural network in a heart monitoring apparatus for the purpose of identifying irregular heartbeats.

Examples of non-technical purposes include classifying text documents by using their textual content; and classifying abstract data records (or even telecommunication network data records) without any indication of a technical use of the resulting classification.

Specific Technical Implementation

AI and ML algorithms may contribute to the technical character of the invention when they are adapted for a specific implementation, i.e., their design is motivated by technical considerations of the internal functioning of the computer system or network, e.g. where a software-implemented algorithm is designed to exploit specific capabilities, or to take account of specific constraints, of the hardware on which it will run. For example, the implementation of ML techniques in a computing platform comprising a GPU and a CPU, where complex training steps are executed by the GPU and preparatory steps by the CPU, is an example where the allocation of tasks of the ML method contribute to the technical character of the invention.

Recommendations

- Where possible, focus the claim on technical features and consider omitting non-technical features if not essential for the definition of the invention. When including AI/ML algorithms (or parts thereof), aim for the specification to include specific applications to fields of technology, e.g. in the description and/or dependent claims (if not already defined in the independent claims). Moreover, if the claim features are adapted to a specific technical implementation, this should be specifically explained in the specification to give the best possible chance of convincing the EPO of this point.
- For this reason, implementation details and a detailed description of technical applications can be very important in Europe for satisfying the requirement for 'technicality' – those are the features that are considered when assessing inventive step, as well as the implementation details potentially being required for sufficiency (see below).
- In the claims themselves, avoidance of non-technical terms (where possible) can help to avoid some features being disregarded out of hand as not being technical, and this includes terminology which may be too easily associated with mathematical methods or business methods in the eyes of an EPO examiner. Furthermore, the claim features, including those relating to the AI or ML, should be readily viewed as part of a causal chain leading to the technical effect being asserted. This aim can be supported by ensuring that sufficient interconnections or interactions between claim features are present, such that the causal chain is easily highlighted.

Sufficiency

Article 83 EPC requires that the European patent application shall disclose the invention in a manner sufficiently clear and complete for it to be carried out by a person skilled in the art.

EPO Guidelines F-III, 3 provide a discussion of how issues can arise in the field of AI, stating *"[a]nother example can be found in the field of artificial intelligence if the mathematical methods and training datasets are disclosed in insufficient detail for the skilled person to be able to reproduce the technical effect without undue burden using common general knowledge over the whole scope of the claim (see also G-II, 3.3.1)."*

While this could be understood to mean that detailed discussions of the training dataset are required for AI inventions, Guidelines G-II, 3.3.1 make clear that this is not always the case: *"The technical effect that a machine learning algorithm achieves may be readily apparent or established by explanations, mathematical proof, experimental data or the like. Mere assertions are not enough, but comprehensive proof is not required either. If the technical effect depends on particular characteristics of the training dataset used, the characteristics required to reproduce the technical effect must be disclosed unless the skilled person can determine them without undue burden using common general knowledge. However, <u>in general, there is no need to disclose the specific training dataset itself</u> (see also F-III, 3 and G-VII, 5.2)."*

Recommendations

- Care should be taken to ensure that the specification includes sufficient detail for the skilled person to be able to reproduce the technical effect without undue burden. This should include consideration of whether the specific training dataset needs to be explained, as well as ensuring that sufficient specific technical detail of, e.g., training data, training parameters, input data for inference, and inference parameters is provided in the description, to sufficiently enable the claimed invention(s).
- It is also useful if disclosed technical features that are not claimed but could potentially be used as fallback positions to help patentability are disclosed in claim ready language in a suitably generalized context, for example in the summary section, given the restrictive EPO rules on added matter.

Claim Types

For AI-related inventions, a notable aspect for patenting is that machine-learning technologies invariably consist of two phases: a training phase (in which a system "learns") and an inference phase (in which that "learning" is put into effect). Separate claims can therefore be crafted to cover each phase. There have been some instances of EPO examiners objecting to the presence of both types of claim in a single application, but it is still considered good practice to include them and it is usually possible to overcome such objections without the need for divisional applications.

The technical features (or features which contribute to technical character) in each phase should form the core structure of the claim in each case, so that the technical purpose is evident in both, and the technical effect of the claimed features can be demonstrated.

3. EPO Case Law

Case law on ML inventions at the EPO is limited, but a few principles can be gleaned. ML *per se* is part of the trend of technology and hence obvious (T 2246/18, T 0161/18). An invention therefore needs to specify more than just using ML for solving a problem. And the problem must be technical; solving a non-technical problem with ML *per se* does not contribute to an inventive step (T0872/19) unless the ML has been designed with the functioning of the computer in mind (T1358/09), an example of which would be enabling or improving parallel processing (T2330/13, T2910/19). Any improvement to be used as the basis of an inventive step must be present for all embodiments covered by the claim scope (T0702/20), must be credible from the content of the application (T0702/20) and must be more than merely encoding / automating human expert knowledge (T1635/19). If neither the output of a machine-learning computer program nor the output's accuracy contributed to a technical effect, an improvement of the machine achieved automatically through supervised learning to generate a more accurate output is not in itself a technical effect (T755/18). While neural networks can provide technical tools useful for automating human tasks or solving technical problems, they must be sufficiently specified, in particular as regards the training data and the technical task addressed (T0702/20).

While ML case law is limited, the EPO treats ML in the same way as any other mathematical method and so the case law concerned with mathematical methods more widely is also relevant.

Requirements for an enabling disclosure

Based on the guidance in T1669/21, we believe a good summary of what is required for an enabling disclosure according to EPO practice is as follows:

- **Clear definition of the machine learning model**: The type of model (e.g., neural network, support vector machine), its architecture, and the specific algorithms used must be explicitly stated. Simply referring to a generic "computational model" is insufficient.
- **Detailed parameter mapping**: The application must provide clear guidance on how to select, pre-process, and represent input parameters within the model. This includes specifying how to handle time-varying or multi-dimensional parameters. Examples are crucial for illustrating these steps.
- **Transparent training procedures**: The description should cover the training data used, the training process, and the criteria for evaluating model performance. It should also address potential challenges such as data scarcity and the prevention of artefacts from random correlations. In instances where data augmentation is considered important, this should also be included.
- **Working examples**: Where possible, include concrete, workable examples demonstrating the implementation of the invention. This could involve providing sample data, model configurations, and training scripts. With a view to supporting a technical effect, also for inventive step arguments, showing result and performance is also advisable.

This is a comprehensive approach, which is recommended in light of the importance of an enabling disclosure and the inability to repair any issues after filing. Of course, practitioners will need to adapt, reduce, or add to this list on a case by case by basis. For example, the Guidelines for Examination, Part G, Chapter II, Section 3.3.1 makes it clear that the fundamental test is that sufficient information to reproduce the invention is disclosed: *"If the technical effect depends on particular characteristics of the training dataset used, the characteristics required to reproduce the technical effect must be disclosed unless the skilled person can determine them without undue burden using common general knowledge. However, in general, there is no need to disclose the specific training dataset itself"*.

C. Considerations for Japan

(Jennifer Che; Sumon Dasgupta**;** Christina Huang**)**

While Japan Patent Office (JPO) did not revise the Patent Examination Guideline specifically for AI, JPO has published several AI invention examples that raised awareness on the enablement and disclosure requirements in 2019. According to Patent Act Article 36(4)(i), "[t]he statement of the detailed explanation of the invention shall be clear and sufficient as to enable any person ordinarily skilled in the art to which the invention pertains to work the invention." Patent Act Article 36(6)(i) provides that a claimed invention shall be disclosed in the description.

The JPO provided Comments to Patenting Artificial Intelligence Inventions in September 2019, "[i]n order for the AI-applied invention to satisfy its enablement requirement, the description that the invention can achieve a certain degree of accuracy in estimation processing should be in Specification, that is, the capacity to create a learned model with a certain degree of accuracy in estimation processing is required for the description in Specification. . . Therefore, if there is any relationship between input and output data in the training data used to create the learned model, we consider that the AI algorithm can create a learned model that performs accurate estimation processing based on the above-mentioned input and output data relationship."

As an example, in Case Example 49, the description discloses that (i) a feature value representing a face shape of a person is a face-outline angle, which is defined between a tangent line to a jaw and a tangent line to a cheek, and (ii) there is a statistically significant correlation between a cosine of a face-outline angle and BMI (defined as a body weight divided by the square of a body height) of a person. However, the description only discloses that any feature value other than a face-outline angle representing a face shape may be obtained from a face image and used. It does not disclose a correlation or the like between (i) a feature value other than a face-outline angle representing a face shape and (ii) a body height, weight, and the like of a person and BMI based on these. As such, the application fails to meet the support requirement or the enablement requirement in Example 49.

The Japan Patent Office (JPO) has published 25 case examples[14] that further clarify standards on inventive step, eligibility, and description requirements for AI related inventions. This included 5 examples in 2017, 10 examples in 2019, and an additional 10 examples on March 13, 2024.

1. Eligibility

Inventions must constitute a "creation of a technical idea utilizing a law of nature" as a whole. Examples of inventions that are eligible include “[t]hose concretely performing control of an apparatus, or processing with respect to the control (e.g. engine control)” and “[t] hose concretely performing information processing based on the technical properties of an object (e.g. image processing).”

For example, claims focusing solely on the content of data, such as sugar content data of apples measured by a sensor, without technical features in the presentation means or methods, are considered mere presentations of information and are not eligible. Conversely, a method for predicting sugar content data of apples using computer software is considered concretely performing information processing based on the specific chemical or biological properties, and is therefore deemed eligible.

[14] https://www.jpo.go.jp/e/system/laws/rule/guideline/patent/ai_jirei_e.html

2. Enablement

Based on the disclosure in the description or common knowledge, a skilled person in the art should be able to recognize that there are certain relationships between the different types of data in a training data set used for machine learning. For products having certain functions due to AI, enablement often requires evaluation with actually made products.

3. DABUS Case

On January 30, 2025 the Japan High Court affirmed a lower Tokyo Court Decision that AI cannot be an inventor, and inventors must be natural persons. This follows similar results in other jurisdictions.

D. Considerations for China

(Jennifer Che)

1. Standards of Evaluating AI Inventions

Subject Matter Eligibility

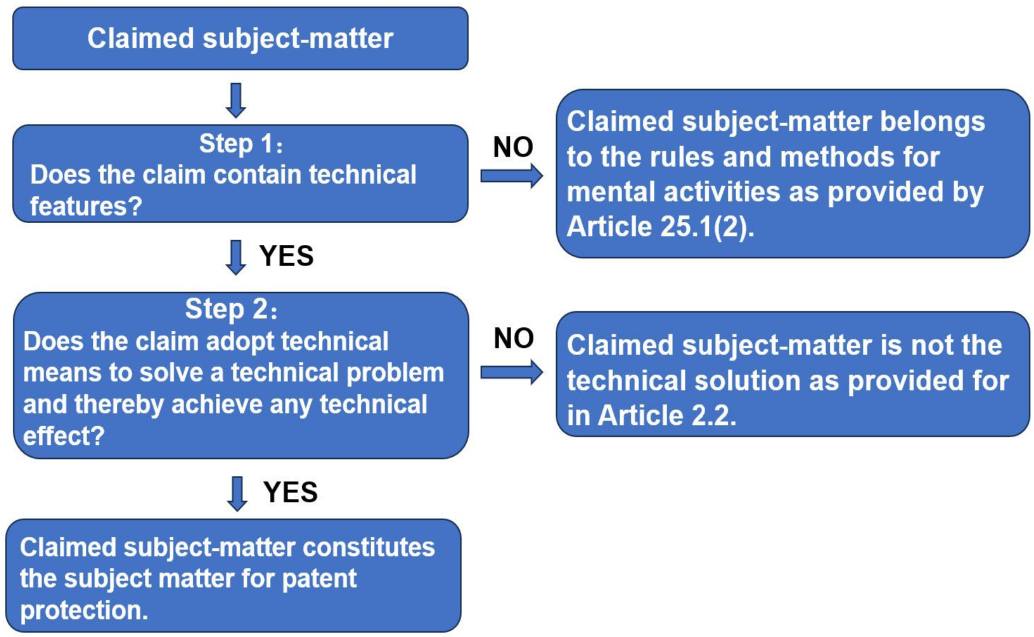
Under Article 2.2[15] of the Chinese Patent Law, an invention is any new technical solution relating to a product, a process, or an improvement thereof. Exceptions are laid out in Article 25[16], and those that are relevant to AI include scientific discoveries and rules and methods for mental activities. Additionally, more unique to China, Article 5.1[17] stipulates no patents may be

[15] Article 2.2: an invention is any new technical solution relating to a product, a process, or an improvement thereof.

[16] Article 25: no patent rights may be granted for scientific discoveries, rules and methods for mental activities, methods for the diagnosis or treatment of diseases, animal or plant varieties, methods of nuclear transformation and the substances obtained by the nuclear transformation method; and designs that are mainly used for marking the pattern, color or the combination of the two of prints.

[17] Article 5.1: no patents may be granted for inventions that (i) violate the laws, social ethics, or harm public interest, or (ii) are accomplished by relying on improperly obtained genetic resources.

granted for inventions that (i) violate the laws, social ethics, or harm public interest, or (ii) are accomplished by relying on improperly obtained genetic resources.



<u>Claim Types</u>

The 2023 Examination Guidelines added "computer program product" as an allowable claim type in China. This is after "computer readable storage medium" and "device comprising modules" were added in 2021. Below are the claim types that are allowable in China and that are often used to protect software related inventions, including AI inventions:

Method

A method for solving [technical solution] comprising the following steps:

Device Comprising a Processor + Software

A device comprising a processor and software executed by the processor

Computer Readable Storage Medium[18] **(New 2021)**

[18] or readable medium storing software, readable storage medium, etc.

A computer readable storage medium storing software that realizes a method when executed by a computer

Device Comprising Modules (New 2021)

A device comprising modules A, B, and C, wherein module A executes the steps of 1, 2, 3

Computer Program Product[19] **(2023)**

A computer program product comprising instructions which, when executed by a computer, causes the computer to carry out [the steps of] the method

Analysis Approach

The 2020 Examination Guidelines emphasized the need to analyze all features in the claim as a whole. To determine inventiveness, one should determine whether the invention as a whole **possesses technical features** and solves a **technical problem** using **technical means** that achieve a **technical effect (using natural laws).** The 2023 Examination Guidelines provided further examples, clarifying certain aspects of eligibility and inventive step as they relate to algorithms in AI, big data, and user experience.

AI or Big Data Algorithms: Improving Internal Performance of a Computer

Algorithms features are patent eligible if the algorithm features and the technical features functionally support and interact with each other. AI or Big Data algorithms (e.g., deep learning, classification, clustering) that improve the internal performance of a computer are patent eligible if the algorithm has a specific technical relationship with the internal structure of the computer system (**technical feature)** and the algorithm improves efficiency or performance of hardware computing (**technical effect)**. Examples of technical effects include reducing the amount of data stored, reducing the amount of data transmitted, or increasing the processing speed of the hardware.

[19] or software product, computer product, computer-readable product, etc.

Specific Application

If a big data algorithm does not specifically impact the internal performance of a computer, a claim directed towards the big data can still be patent eligible under Article 2.2 if the claimed solution processes big data in a specific application field by using big data methods (examples: classification, clustering, regression analysis, neural networks, etc.) to mine the intrinsic correlations in the data that conform to natural laws leading to the technical solution of improved reliability or accuracy of big data analysis.

Inventive Step

In China, an invention possesses inventiveness if "as compared with the prior art, the invention has prominent substantive features and represents an obvious progress." (Article 22). If algorithm features and technical features functionally support and interact with each other and result in improvements to

a) technical means (of a specific application field);
b) internal performance of a computer (new); or
c) user experience (new);

this should be considered when evaluating inventive step.

Even if only the algorithm is novel, but it is specifically tied to improved technical effects or improved user experience of a specific application when interacting with the technical features (for example, if the improvement in user experience is brought about by the adjustment of data architecture and/or data communication method that interact with the ordering and/or pickup notification procedures), **this can still be sufficient to provide inventiveness to the entire claimed invention**.

Sufficient Disclosure

Article 26(3) of the Chinese Patent Law stipulates that "[t]he description shall contain a clear and comprehensive description of the invention or utility model so as to enable a person skilled in the relevant field of technology to carry it out."

In a 2023 Alipay patent invalidation case[20], the CNIPA clarified the current standards of sufficient disclosure for AI inventions, indicating that the specification should 1) clearly define the data involved; 2) specify the AI models; and 3) describe the training/optimizing methods, such that a person skilled in the art can implement the claimed solution based solely on the information disclosed in the specification. If results are not predictable, the specification should ideally include experimental results, simulations, or practical examples to support the feasibility and functionality of the invention.

2. Drafting Tips

Improving the Computer Itself

If possible, demonstrate how the algorithm directly improves the functioning of a computer. Ideally, show at the hardware level how hardware resources are being specifically scheduled, reduced, or improved

Improvements for Specific Applications

Draft the technical effects of the algorithm ("practical application") into the claims as filed. Try to draft many different types of claims covering various angles of the practical applications. Provide as much information as possible (examples, detailed description) about how the algorithm interacts with technical features /specific data to result in specific technical effects for each type of specific application. Create clear steps that are specific to each application and link these steps to actual tangible inputs (e.g., natural language text, data, etc.). The more specific, the more different examples, the better. Explain the technical effect in the specification.

[20] patent invalidation case for Alipay Chinese Invention Patent, application no.: 202010440505.5; titled "A Control Method and System for Payment Operation." published in the CNIPA's "2023 Compilation of Key Decisions on Patent Reexamination and Invalidation Cases".

<u>User Experience</u>

If the algorithm improves user experience, demonstrate how the algorithm works/interacts with the technical features to improve the user experience.

<u>Inventive Step</u>

Show how the algorithm is <u>specifically tied</u> to improved technical effects or improved user experience.

3. Additional Considerations

Consistent with most jurisdictions around the world based on the DABUS case, China does not permit AI to be listed as an inventor, since inventors must be "natural persons." If an invention is co-created with AI, the natural person co-creator needs to "makes creative contributions to the substantive features of an invention"[21] to be listed as an inventor. Content generated by AI must be clearly labelled starting September 2025[22].

E. Considerations for Korea

(Jennifer Che; Sumon Dasgupta; Christina Huang**)**

Korea appears to be less stringent with respect to patent eligibility. For example, Korea identifies patentable subject matter based on novelty and inventiveness (if technical ideas are embodied in a computer). It is worthwhile to note that a technical idea embodied within a "general purpose computer" may be sufficient to satisfy patent eligibility if software and hardware operate together. Some inventions may fail to meet patent eligibility (e.g., economic laws, mathematical formula, mental activity, etc.) that do not satisfy the above (e.g., lack of software and hardware together, lack of technical idea, etc.).

---

[21] Rule 14 of the Implementation Regulations

[22] Measures for the Labeling of Content Generated by AI released by the Cyberspace Administration of China and 3 other agencies

Moreover, the enablement requirement in Korea may require that the application includes a description of a relationship (e.g., a correlation) between input data and output data from a trained model to implement AI-related inventions. For example, a correlation may be met when learning data is described, correlations between learning data and solution to a technical problem, a description of a learning model/method based on input data, how a trained model for solving a technical problem is generated based on input data and methods.

The Korean Intellectual Property Office (KIPO)'s published **Examination Guidelines for Computer-Related Inventions** in 1984, and has revised it multiple times, with the most recent update in July 2014. KIPO published the **Examination Guide in the Artificial Intelligence Field**[23] on January 18, 2021.

1. Patent Eligibility

To qualify as patent eligible subject matter, an AI related invention shall be directed to "a creation of technical ideas utilizing a law of nature" as a whole. Inventions that rely only on mental activities or artificial determination are not considered "using the law of nature", while inventions specifically controlling devices or implements information processing based on the technical nature of a subject are considered invention that utilize the law of nature.

The Korean approach is stricter than China in that it does not allow pure software product claims. Instead, software inventions (including AI-related inventions), can only be claimed as a method or a computer readable medium having a program recorded thereon.

2. Examples of Allowable Claim Types

- A method for solving [technical solution] comprising the following steps:
- A computer (program) readable medium[24] having a program recorded thereon, wherein the program makes the computer execute procedure A, procedure B, procedure C, . . .

---

[23] https://www.kipo.go.kr/upload/en/download/Examination%20Guide.pdf

[24] Or a computer (program) recorded in a medium, a recording medium having data structure, etc.

- A computer (program) readable medium having a program recorded thereon, wherein the program makes the computer operate as means A, means B, means C, . .
- A computer (program) readable medium having a program recorded thereon, wherein the program makes the computer implement function A, function B, function C . . .
- A computer program recorded in a medium to accomplish a specific task when combined with hardware (e.g., to make a computer execute procedure A, procedure B, procedure C)
- A computer-readable recording medium having recorded data structure presenting structure A, structure B, structure C, … operated on a computer

3. Outputs of Computer Programs

New drug candidates discovered using computer programs or algorithms and the methods of discovering the new drug candidates can be patented if they are supported by specific experimental results that can prove the efficacy of the substance in the description of the invention. Claims based solely on *in silico* methods without supporting experimental data are not patentable for failing to meet utility and description requirements.[25]

DABUS Case

On June 30, 2023, a South Korean court decided that AI cannot be an inventor. In 2024, the Seoul High Court dismissed the appeal, and currently the case is pending in the Korean Supreme Court.

F. Considerations for the UK

(Mike Jennings)

Since many applicants from outside Europe are more familiar with EPO practice (as set out in section B above) than with UK law and practice, a brief comparison with EPO practice

[25] “Patent Protection of AI outputs at KIPO” presentation by Korean Intellectual Property Office (6 November 2024)

seems a good starting point. For most inventions that make a technical contribution to the art by virtue of a ***technical effect outside the computer***, the UK courts and UKIPO practice are now very similar to the EPO (see "***specific technical application***" in section B above), applying a relatively narrow interpretation of the UK's exclusions from patentability for computer programs, mathematical methods and mental acts and allowing claims that recite a specific technical purpose such as control of a physical apparatus or industrial process, or enhanced image processing.

However, the UK test for patentability remains somewhat different from the EPO test, so outcomes can differ for some types of AI invention. Notably, UK courts and the UKIPO have **not yet** followed the EPO's lead in approving patents for some of the ***specific technical implementations*** of mathematical methods such as machine learning algorithms, where the design of an algorithm is specifically adapted to the system or network on which it will run. (At the EPO, if the design of an algorithm is motivated by the requirements of the hardware – either constraints or special features – our claims fall in a patenting "safe zone"), but the UKIPO has test has been interpreted more narrowly by the UKIPO. This may change soon. We are awaiting a judgment from the UK Supreme Court relating to patentability of a particular ANN and its training. The Appellant (and an intervention from CIPA and IP Federation) recommended that the Supreme Court increases consistency with the EPO's approach to the assessment of computer-implemented inventions including AI, whereas the UKIPO's recommendation to the SC favoured continuity of current UKIPO practice. It is too soon to predict the final outcome, but we hope to have a UK Supreme Court judgement by the end of 2025.

We recommend seeking an update from a registered UK patent attorney in early 2026, to guide your decision on which patent office to choose for your AI/ML inventions. The current situation is summarized below.

Section 1(2) of the UK Patents Act corresponds to Article 52(2) of the European Patent Convention and lists certain subject matter that is not to be considered an invention:

"*(a) a discovery, scientific theory or mathematical method;*

(b) a literary, dramatic, musical or artistic work or any other aesthetic creation whatsoever;

(c) a scheme, rule or method for performing a mental act, playing a game or doing business, or a program for a computer;

(d) the presentation of information;"

However, these exclusions only apply "*to the extent that a patent or application for a patent relates to that thing as such*". This wording has been interpreted differently for the different categories of excluded subject matter, with the exclusion of mental acts interpreted narrowly (implicit computer implementation has been accepted as avoiding the exclusion), a strict scrutiny of computer programs and mathematical methods to determine whether there is a technical contribution, and quick rejection of business methods.

The UK approach to the assessment of excluded subject matter involves a structured test (the so called "Aerotel test" from 2007, confirmed in *Comptroller-General of Patents, Designs and Trade Marks v Emotional Perception AI Ltd* [2024] EWCA Civ 825) involving these steps:

(1) Properly construe the claim.

(2) Identify the actual contribution (although at the application stage this might have to be the alleged contribution).

(3) Ask whether it falls solely within the excluded matter.

(4) If the third step has not covered it, check whether the actual or alleged contribution is actually technical.

When considering the second step of identifying the technical contribution, the UKIPO and UK courts often consider the following signposts for inventions involving computer programs and mathematical methods - asking:

i) whether the claimed technical effect has a technical effect on a process which is carried on outside the computer;

ii) whether the claimed technical effect operates at the level of the architecture of the computer; that is to say whether the effect is produced irrespective of the data being processed or the applications being run;

iii) whether the claimed technical effect results in the computer being made to operate in a new way;

iv) whether a program makes a computer a better computer in the sense of running more efficiently and effectively as a computer;*

v) whether the perceived problem is overcome by the claimed invention as opposed to being merely circumvented.

*These signposts were initially set out in *AT &T Knowledge Ventures LP's Patent Application* [2009] EWHC 343 (Pat), and then signpost (iv) was revised in *HTC v Apple* [2013] EWCA Civ 451. It should be noted that these signposts were never intended to be a prescriptive list and in *HTC v Apple* it was noted that signpost (iv) is only an illustration of the "*broader question whether the invention solves a technical problem within the computer*". If the UK Supreme Court chooses to endorse this more general question in 2025 (see below), we will be able to recommend changes to UKIPO practice for greater consistency with the EPO's allowance of a slightly wider range of technical contributions (the EPO's "*specific technical implementations*").

As at the EPO, claims that are functionally limited to a specific technical application of AI/ML are treated positively, but claims to abstract mathematical methods and applications of AI/ML technology for non-technical business methods will be refused. Therefore, it continues to be important to include within an AI-related patent application a description of all of the inputs, training, outputs, and technical applications – including all interactions with physical systems outside the computer – as these may be important for patentability. It is also necessary to include a detailed description of the system/software architecture for sufficiency reasons, as in all jurisdictions.

In its September 2022 guidance on AI patenting, the UKIPO endorsed the EPO Board of Appeal's insufficiency decision T161/18, which criticized a lack of detailed disclosure related to the training of an ANN in an application that relied on that training as its point of novelty.

In 2024, the UK Court of Appeal overturned a 2023 UK High Court decision relating to patentability of claims to a system and method for providing media/text file recommendations using an Artificial Neural Network (ANN) that was trained to perceive semantic similarity between pairs of files. The claims were rejected by the Court of Appeal with the ANN considered to be "a computer" and the weights and biases treated as an excluded "program for a computer" (*Comptroller-General of Patents, Designs and Trade Marks [2024] EWCA Civ 825* overturned *Emotional Perception AI v Comptroller-General of Patents [2023] EWHC 2948 (Ch)*).In the High Court, the Judge found that the claimed system made a technical contribution to the art, at least because a selected file having certain attributes such as similarity characteristics was output to an end user by a system that had established the identification system and implemented it. The recommended media/text file was identified as being semantically similar by the application of technical criteria, which the system had worked out for itself. This was considered to be a technical effect, which contributed to patentability. The Court of Appeal largely disregarded the implementation details and took the view that the purpose of recommending semantically-related media files did not qualify for patent protection, and considered the claimed subject matter to be a mere computer program as such. The Applicant (and others) considered the Court's technical reasoning to be flawed and the Applicant appealed to the Supreme Court.

The UKIPO changed its examination practice after each of the High Court and Court of Appeal judgements. It will now continue to apply the Court of Appeal's guidance for ANNs and related inventions, until we have a judgement from the Supreme Court. The prosecution of other applications for ANNs and other AI/ML inventions will not be stayed to await the Supreme Court, so applicants may wish to consider applying for extensions of time for AI/ML UK national patent applications that have received objections and have critical deadlines within 2025, if it is not appropriate to limit the claims to a specific technical purpose outside the computer.

The UKIPO will review its examination guidelines when the UK Supreme Court judgement is available, and so updates are expected in early 2026.

It is worth noting that the UKIPO and EPO worked closely on a recommendation for member states of the European Patent Organisation in 2023 with the aim to achieve greater harmonization of patent practice for computer-implemented inventions (within the limits of what is possible while taking account of EPO and UK national case law). Despite our desire to further increase harmonization, the different approaches of the UKIPO and EPO often do achieve similar outcomes. In brief, the majority of patent applications for technical applications of AI/ML will be accepted as patentable in the UK just as easily as in the EPO (and other national patent offices of EPC contracting states), with patents available if the claimed invention is new, inventive, and sufficiently described; but we await the Supreme Court's judgement to confirm how the UK courts and UKIPO should assess inventions based on computer-implemented algorithms and software that are designed to solve/mitigate technical problems within the computer.

G. AI Inventorship

(Sumon Dasgupta; David Kincaid; John Pienkos; Jennifer Lacroix)

According to United States patent law, the threshold question in determining inventorship is who conceived the invention (MPEP 2109). Conception can mean "the complete performance of the mental part of the inventive act" and it is "the formation in the mind of the inventor of a definite and permanent idea of the complete and operative invention as it is thereafter to be applied in practice." (MPEP 2138.04).

AI systems pose unique challenges when determining inventorship. A fundamental inquiry is whether the AI systems are used as a tool to help natural persons conceive of an invention, or whether the AI systems conceive of the invention. If the AI system is simply used as a tool, then the inventor would be the natural person. That is, actions by a natural person(s) that qualify as a contribution to the conception of an invention are unaffected by use of an AI system. Such actions have been broadly interpreted, and in most instances (nearly all to date), a natural person can be fairly called an inventor. As an example, actions including designing the architecture of the AI system, choosing input data to provide to the AI system, or developing an algorithm to enable the AI system to process data, may be adequate to qualify a natural person as a contributor to conception.

Still, there are circumstances (and possibly an increasing number of circumstances) in which a human has only minimal interactions with an AI system and questions can arise as to whether any human at all should properly be considered an inventor and, if not, whether in some sense the AI should be an inventor.

There are different hypothetical scenarios in which AI might be considered to constitute an inventor with respect to an invention when human involvement in the inventive process is or seems minimal. Consider generative AI models. Generative AI models can generate new and original content (e.g., computer code, designs, architecture, art, drugs, etc.). In one particular example, it is possible that an AI-enabled drug development process results in a new and novel drug being identified. Since AI is extensively used in the drug development process, there is a possibility that human intervention decreases to a point where the AI could be the only meaningful contributor to the origination of the invention. As a different example, consider that some generative AI models can generate code based on a request received from a natural person.

It is potentially possible that the request can frame a problem (e.g., a code to solve an existing problem), but does not provide any possible clues or suggestions as to how to solve the problem. In such an instance, could the natural person be considered an inventor if the resulting code is novel and patentable? Thus, at what point does the AI model cease to be merely a tool, and possibly rise to the level of an inventor?

Notwithstanding such considerations, as will be explained below, under most current patent laws, AI cannot be listed as an inventor for a patent application. It remains to be seen whether, as AI continues to evolve and expand, these laws remain the same or change over time in regard to the fundamental question of who or what can be considered an inventor.

1. AI Inventorship in the United States

There have been numerous developments in the past several years in which various actors in the United States government have considered and confronted many issues regarding the use of AI in innovation.

For example, the USPTO has published two notices in the Federal Register, in 2019 and 2023, seeking comments regarding the use of AI and inventorship, and in February 2024 the USPTO additionally issued "Inventorship Guidance for AI-Assisted Inventions" (88 FR 10043 (February 13, 2024). The USPTO also held its inaugural AI/ET Partnership meeting in June of 2022, which discussed issues such as whether AI could actually "conceive" of inventions.

The first notice was in August of 2019, and requested comments on patenting AI inventions ("Request for Comments on Patenting Artificial Intelligence Inventions," 84 FR 44889 (August 27, 2019)). The result was a report titled "Public Views on Artificial Intelligence and Intellectual Property Policy," published by the USPTO in October of 2020. The report indicated that the comments received included very mixed views regarding whether AI is merely a tool that cannot invent without human intervention, or whether AI could contribute to the creation of inventions, jointly with humans or even on its own.

The second notice was published in February of 2023, and the time period for comments closed in May of 2023 ("Request for Comments Regarding Artificial Intelligence and Inventorship, 88 FR 9492 (February 14, 2023)). In the 2023 notice, the USPTO recognized that it "plays an important role in incentivizing and protecting innovation, including innovation enabled by artificial intelligence (AI), to ensure continued U.S. leadership in AI and other emerging technologies (ET)" and sought "stakeholder input on the current state of AI technologies and inventorship issues that may arise in view of the advancement of such technologies, especially as AI plays a greater role in the innovation process" *Id.*

Despite the USPTO's curiosity and willingness to consider the issues regarding AI, the law to date in the United States remains clear that AI cannot be named as an inventor. In July 2019, Thaler filed U.S. Patent Application Nos. 16/524,350 and 16/524,532, naming an AI system (Device for Autonomous Bootstrapping of Unified Sentience (DABUS)) as the sole inventor. In April of 2020, the USPTO issued decisions denying the applications, and concluding that the Patent Act defines “inventor” as being limited to natural persons. Thaler appealed to the U.S. District Court for the Eastern District of Virginia. However, that court granted summary judgment in favor of the USPTO, agreeing that the Patent Act requires an “inventor” to be a natural person. Thaler then appealed to the Federal Circuit, which affirmed (*Thaler v. Vidal*, 43 F.4th 1207 (Fed. Cir. 2022)). The Federal Circuit explained that the Patent Act expressly provides that inventors are “individuals,” and that the term “individuals” means a human being. *Id.* at 1211 (relying on *Mohamad v. Palestinian Auth*., 566 U.S. 449, 454 (2012)). Accordingly, the Federal Circuit concluded that, “Here, Congress has determined that only a natural person can be an inventor, so AI cannot be” (43 F.4th at 1214). Thaler filed an appeal to the Supreme Court of the United States in March of 2023; the Court declined to hear the case.

Subsequently, the “Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence” (Executive Order 14110) was issued on October 30, 2023. The Executive Order set forth a policy “to advance and govern the development and use of AI in accordance with eight guiding principles and priorities” and directed that “[w]hen undertaking the actions set forth in this order, executive departments and agencies (agencies) shall, as appropriate and consistent with applicable law, adhere to these principles.” *Id.* The principles generally are: (1) Artificial Intelligence must be safe and secure; (2) Promoting responsible innovation, competition, and collaboration; (3) The responsible development and use of AI require a commitment to supporting American workers; (4) Artificial Intelligence policies must be consistent with my Administration's dedication to advancing equity and civil rights; (5) The interests of Americans who increasingly use, interact with, or purchase AI and AI-enabled products in their daily lives must be protected; (6) Americans' privacy and civil liberties must be protected as AI continues advancing; (7) It is important to manage the risks from the Federal Government's own use of AI and increase its internal capacity to regulate, govern, and support responsible use of AI to deliver better results for Americans; (8) The Federal Government should lead the way to global societal, economic, and technological progress, as the United States has in previous eras of disruptive innovation and change. *Id.*

Pursuant to the Executive Order, the USPTO issued a notice entitled "Inventorship Guidance for AI-Assisted Inventions" (88 FR 10043 (February 13, 2024)). The guidance explains that *Thaler* is an "acknowledgment that the statutory language clearly limits inventorship on U.S. patents and patent applications to natural persons" and that "while AI-assisted inventions are not categorically unpatentable, the inventorship analysis should focus on human contributions, as patents function to incentivize and reward human ingenuity." *Id.* Specifically, the guidance states that AI-assisted inventions "are not categorically unpatentable due to improper inventorship if one or more natural persons significantly contributed to the invention." *Id.* For determining whether a human can properly be named on an invention developed using AI, the USPTO's guidance walks through standard rules for determining inventive contribution, reiterating the factors set forth in *Pannu v. Iolab Corp*., 155 F.3d 1344, 1351 (Fed. Cir. 1998).

The USPTO guidance sets forth five Guiding Principles:

1. A natural person's use of an AI system in creating an AI-assisted invention does not negate the person's contributions as an inventor.
2. Merely recognizing a problem or having a general goal or research plan to pursue does not rise to the level of conception.
3. Reducing an invention to practice alone is not a significant contribution that rises to the level of inventorship.
4. A natural person who develops an essential building block from which the claimed invention is derived may be considered to have provided a significant contribution to the conception of the claimed invention even though the person was not present for or a participant in each activity that led to the conception of the claimed invention.
5. Maintaining "intellectual domination" over an AI system does not, on its own, make a person an inventor of any inventions created through the use of the AI system.

The USPTO guidance also includes an assessment of the impact of AI-assisted inventions on various aspects of patent practice.

As to the duty of disclosure and the naming of inventors, the USPTO guidance states that the duty of disclosure is not being modified or changed, and that "applicants rarely need to submit information regarding inventorship" pursuant to that duty. However, "special care should be taken by those individuals subject to this duty to ensure all material information is submitted to the USPTO to avoid any potential negative consequences." *Id*. The USPTO guidance also states that "[i]n situations in which it is determined that contributions by a named inventor to the claimed subject matter do not rise to the level of inventorship" a correction of inventorship should be filed, and that "[i]n situations in which inventorship with respect to a particular claim cannot be corrected (i.e., no natural person significantly contributed to the claimed invention), the claim must be canceled or amended." *Id*.

Regarding the duty of reasonable inquiry, the USPTO guidance states that there is no modification or change, but reminds applicants and practitioners that the duty includes questioning proper inventorship, including "questions about whether and how AI is being used in the invention creation process." *Id*.

With respect to requirements for information, the USPTO guidance states that requirements for information may be used "in the context of applications or patents for AI-assisted inventions such that if an examiner or other USPTO employee has a reasonable basis to conclude that one or more named inventors may not have contributed significantly to the claimed subject matter, the examiner or other USPTO employee may request information from the applicant regarding inventorship even if the information is not material to patentability." *Id*.

The USPTO guidance states that "[t]here is no change in oath or declaration practice for the named inventors in a patent application." Id. As to ownership and assignment of inventions, the USPTO guidance states that "there is no change in practice for AI-assisted inventions with regard to the applicant or assignment of ownership rights." *Id*.

With respect to claiming the benefit of priority to a prior-filed application, the USPTO guidance states that "[f]or all applications and patents, including those that cover AI-assisted inventions, the prior-filed application and the United States application or patent claiming the benefit of, or priority to, the prior-filed application must name the same natural person as the inventor, or have at least one joint inventor who is a natural person in common." *Id*.

Finally, the USPTO guidance stated that it "views the inventorship guidance on AI-assisted inventions as an iterative process and may continue with periodic supplements as AI technology continues to advance and/or as judicial precedent evolves." *Id*.

In the United States, the law remains clear: only natural persons who made a significant contribution to an invention—not AI systems—can be named as inventors.

2. AI Inventorship in Other Jurisdictions

From a global perspective, AI-generated inventions have sparked significant debate and legal scrutiny. Thaler's patent applications for DABUS have become the primary case study globally, as they were filed across numerous jurisdictions worldwide. These jurisdictions include Australia, South Africa, the European Patent Office (EPO), the United Kingdom (UK), Germany, Japan, and Israel. Among these, South Africa currently stands alone in recognizing AI as an inventor, while the rest rejected Thaler's applications on improper inventorship grounds.

South Africa's Companies and Intellectual Property Commission granted Thaler's application naming AI (DABUS) as an inventor. However, South Africa's patent system does not conduct a substantive examination of patent applications and merely checks to see if formal requirements are satisfied. For this reason, some have questioned the validity of this outcome. (*See*, e.g., "South Africa was wrong to patent an AI's 'invention'" by Mhangwane et al.). However, others have championed the outcome as a progressive and pro-science stance. (See, e.g., "AI inventorship: The right decision?" by Thaldar et al.).

Other jurisdictions have held that AI cannot be an inventor. For example, the Australian Patent Office initially refused Thaler's application for naming AI (DABUS) as an inventor, and Thaler successfully appealed. The Federal Court found that an AI system or device can be recognized under the Australian Patent Act 1990 as an inventor, but the court held that "a nonhuman inventor can neither be an applicant for a patent nor a grantee of a patent." *Thaler v. Commissioner of Patents* [2021] FCA 879. The Australian Patent Office then appealed the decision, and the Full Court of the Federal Circuit ruled against Thaler, reasoning that "the origin of entitlement to the grant of a patent lies in human endeavor." A subsequent request for special leave to appeal to the High Court of Australia was denied.

The EPO has also rejected AI as an inventor. In the EPO, Thaler's applications were rejected because the European Patent Convention (EPC) requires an inventor to be a natural person. The EPO stated, "AI systems or machines have at present no rights because they have no legal personality comparable to natural or legal persons." See Grounds for Decision 18275147.3 dated January 27, 2020. Thaler appealed, and the EPO Legal Board of Appeal agreed with the EPOs decision.

The respective authorities in Germany, the UK, Israel, and Japan, have also refused to grant Thaler's applications naming AI (DABUS) as an inventor. In Germany, the Federal Patent Court found that AI generated inventions are not excluded from patentability, but a human inventor must be named. An applicant can identify the AI being involved in the description of the invention. In the UK, Thaler's applications were initially denied, and Thaler appealed. The UK Supreme Court affirmed that only humans can be inventors and indicated that legislative changes would need to be made in order to recognize AI as an inventor. Similarly, in Japan, after the Japanese Intellectual Property Office initially denied Thaler's application on the grounds of improper inventorship, the Intellectual Property High Court ruled that AI systems like DABUS cannot be named as inventors on Japanese patent applications because inventors are limited to natural persons. The Israel Patent Office also denied Thaler's applications.

The overwhelming consensus across global patent systems is that inventors must be natural persons, meaning that AI systems cannot be named as inventors.

3. Tips for AI-Assisted Inventions

Even if AI cannot be named as an inventor in most global jurisdictions, there can arise additional questions about who should be listed as an inventor regarding a patent application when the claimed subject matter of the patent application involves AI. The extent to which such other questions arise can depend upon the manner in which any given invention may involve or relate to AI. Indeed, as articulated by others, an invention may involve or relate to AI in any of several ways, including whether the invention relates to a technical improvement of an AI system or method of implementing AI (e.g., an improved method of training an AI system, or selection of a particular training data set), whether the invention performs a process that entirely or in part relies upon AI to make a determination or take an action that is part of or influences the process, or whether the invention relates to a data output or result provided by an AI system (e.g., a chemical formula) that has utility in other contexts.

As currently articulated, the law generally presumes that AI constitutes a tool during the invention process. There has been movement towards the notion that AI may meet the threshold for inventorship in some cases. For example, while the wording of the "Duty of Disclosure" requirement has not changed as of the drafting date of this White Paper, the USPTO did clarify that improper inventorship in AI-assisted inventions is a ground of rejection under 35 U.S.C. §§ 101 and 115. Further, parties identified in 37 C.F.R. 1.56(c), 1.555(a), and 42.11(a) have a duty to disclose to the USPTO information that raises a *prima facie* case of unpatentability due to improper inventorship or that refutes, or is inconsistent with, a position an applicant takes in opposing an inventorship rejection or asserting inventorship (e.g., evidence that demonstrates a named inventor did not significantly contribute to the invention because the person's purported contribution was made by an AI system).

When AI is viewed as a tool, human beings can be considered inventors in any of several manners, and it is appropriate to list human beings who fulfill any of these types of roles. First, when a patent application concerns an invention relating to an improvement to an AI system or training methodology conceived of by a human being, it seems generally to be appropriate that the human being should be listed as an inventor, just as if the human being developed an improved motor or gear system for an electric drill. Second, when a patent application concerns an invention relating to an improved process envisioned by a human being in which AI is used to perform or implement the process so as to make a determination or take an action—but in which the AI being used is a conventional AI system (that is, one not involving any particular innovative feature contributed by that human being or any other human collaborator)—then it would appear that the inventor for that patent application would be solely that human being who envisioned the improved process (but not any human being involved with the development or implementation of the AI system used to perform the process).

Third, when a patent application concerns a useful data output or result generated by the operation of an AI system, where the usefulness of the data output or result is something that is recognized by a human evaluator, the human evaluator should be the inventor in that context. Of course, if there are multiple human beings who collaborated and jointly contributed to the claimed subject matter of a given patent application in any of the above manners, then it may be appropriate for all of those individuals to be listed as joint inventors. Indeed, for example, if a patent application concerns the useful data output or result generated by the operation of an AI system in which one human collaborator recognized the usefulness of the data output/result and another human collaborator provided an improvement to the AI system allowing for that data output or result to be generated, then in that circumstance it may be appropriate to list both of those individuals as joint inventors.

Issues surrounding who should be considered inventors in relation to patent applications that involve or relate to AI will undoubtedly continue to be the subject of debate in the years ahead as AI technology continues to advance. Additional nuanced approaches may be developed in terms of determining whether any given human being should be included as an inventor in regard to any given patent application involving or relating to AI. As to whether laws in the U.S. and around the world will more widely over time come to view AI itself, a non-human entity, as potentially being an inventor in regard to patent applications, this remains to be seen. Changes in laws to that effect would necessarily reflect profound changes in our understanding of what constitutes an acceptable basis for attributing inventorship, in terms of concepts such as whether conception is key to inventorship and what inventorship truly entails in terms of consciousness or sentience. Further, changes in laws on this subject may also be the subject of vigorous debate given that such changes could have enormous ramifications upon what persons or entities control or own the fruits of innovation efforts as AI plays a greater and greater role in these efforts.

H. Proposed Legislation

The *Patent Eligibility Restoration Act of 2023* introduced by Senators Thom Tillis and Chris Coons, reintroduced in May 2025, will no doubt impact the assessment of subject matter eligibility under the current *Alice* and *Mayo* framework, particularly with respect to the implementation of the two-part test to identify claims that are directed to a judicial exception (Step 2A) and to then evaluate if additional elements of the claim provide an inventive concept (Step 2B) (also called "significantly more" than the recited judicial exception).

Importantly, the proposed bill in its current form addresses 101 by addressing Step 2A via the elimination of judicial exceptions, and Step 2B by eliminating Alice's "well-understood, routine and conventional" test. However, two things are notable with respect the proposed bill's approach:

1) Despite the elimination of all court-created judicial exceptions under Section 2 Part A, mathematical formulas and mental processes are still designated as being explicitly excluded from categories of inventions eligible for patent protection under Section 2 Part D.
2) The proposed bill negates the well-understood, routine and conventional test to determine what is "significantly more" than an abstract idea. This negates an oft-used test that is primarily implemented by the Federal Circuit to the detriment of patentees. Yet the proposal remains silent on the USPTO's practical application analysis implemented in the 2019 Guidance (PEG) that is followed by the PTAB.

The combination of these two factors will likely significantly impact the Federal Circuit's approach towards subject matter eligibility. However, by failing to address the practical application analysis of the USPTO's Step 2A Part 2, the proposed bill will still leave the PTAB's current approach of subject matter eligibility mainly intact. That is, AI and ML method claims are still equally subject to a determination of being directed to a mathematical formula or a mental process. More importantly, when confronted with a combination of apparatus, method and computer program claims, the PTAB's predominant practice is to adopt the method claim as exemplary, thereby subjecting non-method and apparatus claims as being subject to being found as being directed to a mental process along with the method claim.

Namely, the PTAB's practice of utilizing a method claim as exemplary rests in Rule 41.37(c)(1)(iv) which specifies that, when an applicant does not provide separate arguments for different patent claims, the Board may select a single claim from a group and decide the appeal on the basis of the selected claim alone. This practice was implemented as recently as the PTAB's decision in *Ex parte Philip E. Vasey* (Appeal 2022-001109) dated July 5, 2023.

As such, method claims presented with non-method claims can render the non-method claims to be vulnerable unless the applicant explicitly presents different 101 arguments with respect to each different type of patent claim. The current rule places the burden on the applicant to preemptively address representative claim treatment during prosecution or in the brief.

Therefore, the need to present problem/solution language in claim drafting may remain a dominant practice in AI and ML claim drafting despite the "technological improvement" analysis being a progeny of the well-understood, routine and conventional test set up by the Federal Circuit in *Berkheimer v. HP*. While the bill's elimination of this test should theoretically eliminate the technological improvement test, the USPTO's adoption of the technological improvement analysis under the 2019 PEG moved the assessment of improvements to a question of practical application and out of the realm of determining what is well-understood, routine and conventional. As the bill in its current form is silent on the question of practical application, the USPTO framework for Prong 2A Part 2 may still be largely implemented by examiners and the PTAB alike, thereby rendering the need for applicants to maintain this practice of presenting problem/solution claims.