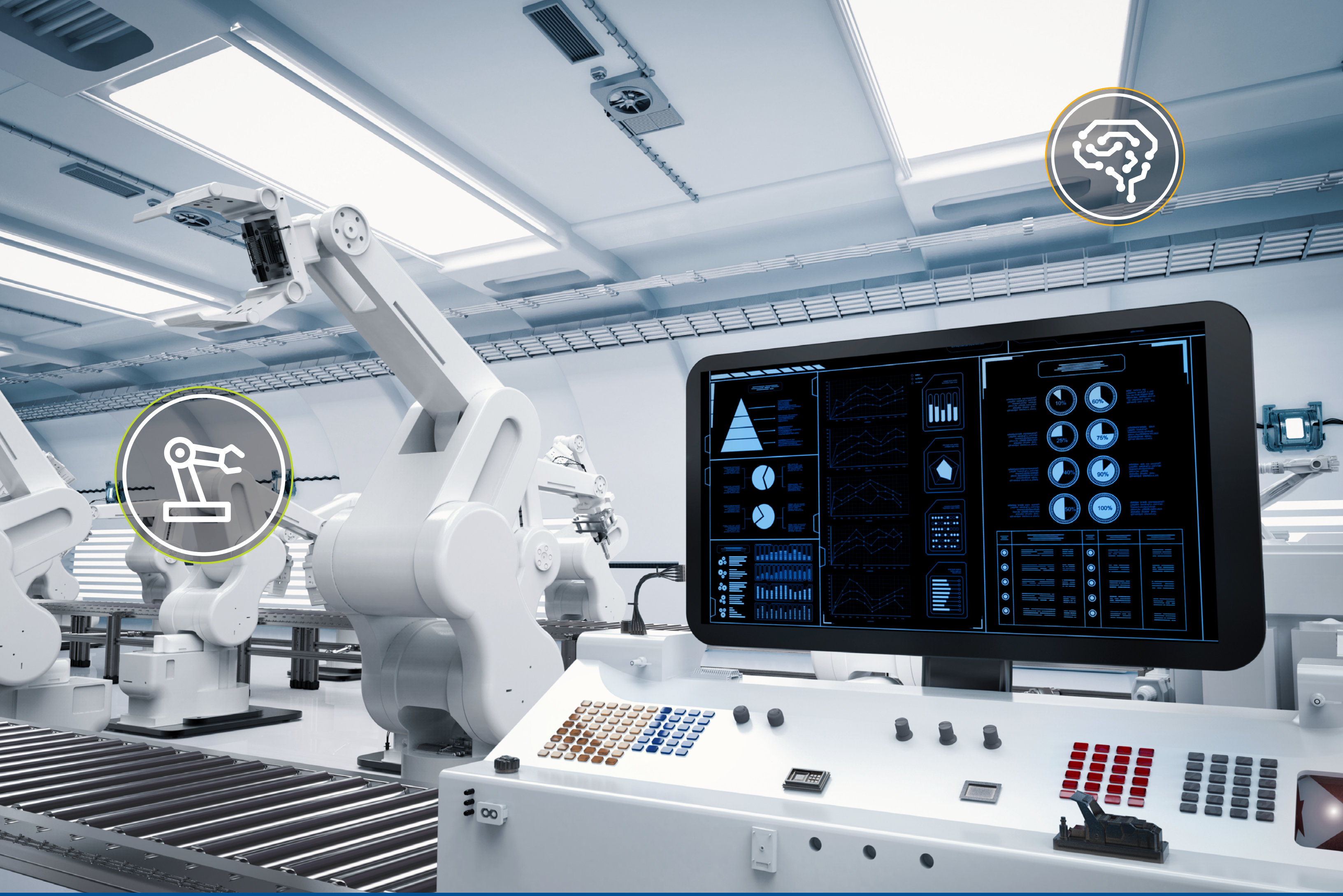


INTELLECTUAL PROPERTY ASPECTS OF MACHINE LEARNING

www.nxp.com

INTELLECTUAL PROPERTY ASPECTS OF MACHINE LEARNING

It is common for manufacturers and suppliers to offer maintenance contracts to their customers for operation-critical equipment. To help avoid failures that could impact the customer's business, a preventative maintenance application based on a Machine Learning (ML) Model is used. The supplier spent time, money and effort to build this model.

The customer could, however, copy this Intellectual Property (IP) and manage the maintenance without the supplier's assistance, eliminating the costs of a maintenance contract. Furthermore, a third-party, including a competitor, may profit from copying the model instead of spending resources to create its own. This white paper explores which parts of an ML model could be protected by which IP laws.

To build an ML model for maintenance, an appropriate training set must be collected and labelled, the architecture and training parameters must be chosen for optimal accuracy-speed trade-offs for the algorithm, and computing time is required to train the model. If the IP of the ML model for maintenance is not properly protected, a competitor can copy and steal the ML model with very little time and effort, lightly tweak the model to avoid detection, and deploy the model in their own products.
This is just one example among many in which a company will want to protect its investments and IP. The question is: How will IP be protected in the context of ML – now and in the future?

An ML model can be a considerable investment and a valuable asset for any company. Although ML-driven business is gaining more and more traction, some companies are reluctant to make the required investment for data collection and model building simply because they have concerns about other individuals and companies, including competitors, potentially exploiting and profiting from their work.

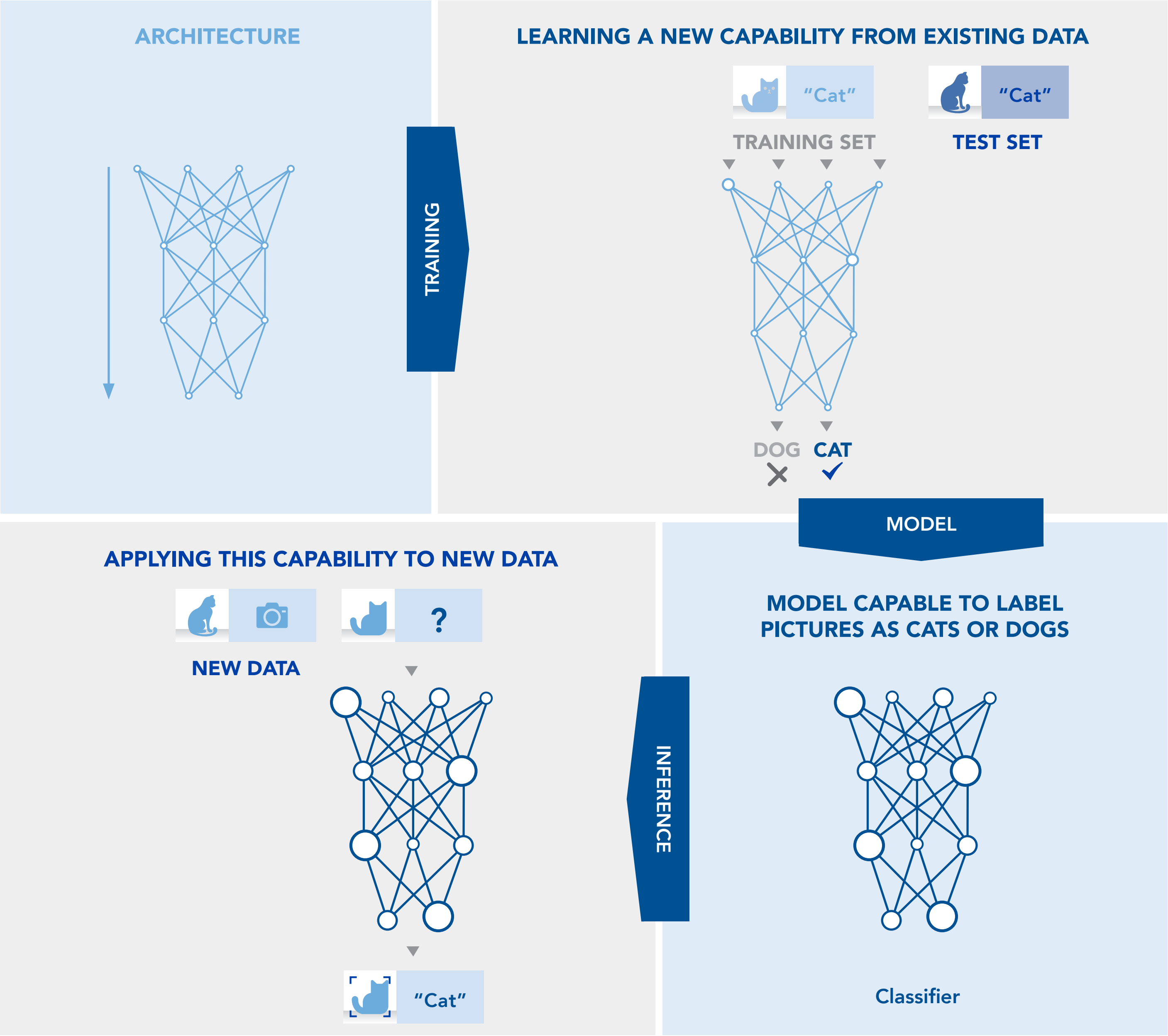
Traditional intellectual property rights such as patents or copyrights are available to protect investments in the creation of non-physical assets. However, across the legal community, an open question remains about the protection of ML through IP rights, and how much – and what – is covered by IP rights. This white paper sets out to add clarity to the legal background and challenges for the IP aspects of ML.

TERMINOLOGY

Before we further dive into the IP aspects of ML, it is important to understand the terminology. Broadly, ML is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using manually programmed instructions, relying instead on patterns and inference. Typically in ML, a set of training data is used to derive weights for the statistical models. These weights are then applied to new situations to obtain an answer from the model that is applicable to the new situation. One popular type of ML models are neural networks.

To clarify the process of employing neural networks, we refer to the picture below:

Machine Learning model to label pictures as cats or dogs


ARCHITECTURE
TRAINING
LEARNING A NEW CAPABILITY FROM EXISTING DATA
"Cat"
"Cat"
TRAINING SET
TEST SET
DOG
CAT
MODEL
MODEL CAPABLE TO LABEL PICTURES AS CATS OR DOGS
Classifier
INFERENCE
APPLYING THIS CAPABILITY TO NEW DATA
NEW DATA
?
"Cat"

In the neural networks type of ML, there are two steps. First, in the training phase, parameters for the architecture are derived to give the model a specific functionality. We call this training the model. The quality of the trained model is measured by using test data. Second, in the inference phase, the trained model is used to make predictions, for example, to perform classification on new data.

Although terminology for all these concepts varies in the literature, in this white paper, we use the following terminology:

NEURAL NETWORK ARCHITECTURE

The collection of neurons in the neural network, the connections between them, and the activation functions used. This architecture can be visualized as a directed graph.

TRAINING SET

A set of data which is used to train the architecture, allowing it to determine the right weights.

TEST SET

A second set of data, used to test and validate that the model is providing the expected result.

ML SYSTEM

The software and hardware that implements machine learning (training and/or inference).

MODEL

For neural networks, the model is the collection of weights associated to the connections of the neural network architecture. These weights are collected during training.

TRAINING PARAMETERS

Parameters are used to steer the training algorithm. Eg., how many times should we repeat the training set? How many data items do we process before we update the weights? How large are the changes we apply to the weights per update? What cost function do we use for optimization?

ML is used today for a wide variety of tasks. A popular application is classification, e.g. recognizing certain objects in images or videos, classifying texts as particular categories and detecting fraud or anomalous measurements, which includes our previous example for predictive maintenance.

Other applications include forecasting and object detection as used in autonomous vehicles. For many companies employing ML, the training set and the model used for an ML application are valuable pieces of information that competitors should not access. This has led to the question on how to protect these and other ML elements through legal means, IP rights.

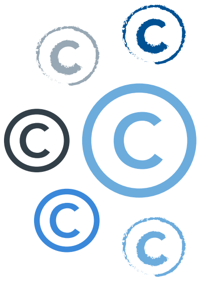
INTELLECTUAL PROPERTY RIGHTS

IP rights (IPR) are legal rights that protect non-tangible business assets against various types of misuse by third parties. Such misuse can be stopped by a legal injunction issued by a court, often combined with claims of financial damages and/or seizure of infringing products. However, each type of IPR has its own particular requirements and limitations. In this whitepaper, we discuss copyrights, patents, database rights and trade secrets.

COPYRIGHT

Copyright is the most well-known type of IPR. A copyright is the right to forbid copying and dissemination of a protected work. Traditionally this right has been used much in the creative arts, e.g. for music, books and photographs. However, copyright applies just as much to business works such as software, manuals, whitepapers (even this one!), company videos and so on.

The law on this type of right is greatly standardized around the world. A work is protected automatically upon creation, with no application or registration needed. Not even a copyright notice is required, although this is often done in an effort to scare off would-be copyists. The only real requirement is that some form of creativity is present in the work. A mere list of dates, for example, is not copyrighted, but a cleverly formulated sentence could be.



A limitation of copyright is that it only protects against actual copying. An independent recreation of the same work is not an infringement of copyright. The independence of the recreation can be demonstrated through documentation or logs of the process of the creation.

PATENTS

Patents are the heavy lifters of the IPR world. When an innovation is protected by a patent, no one may make, use or sell any device incorporating that innovation. Unlike a copyright, a patent protects any independent re-creation. The patent holder can demand royalties or simply put an end to someone's commercial use of his innovation.The major drawback is the application process, which involves costly fees and a multi-year examination process with uncertain outcome. A complication with software is the strict case law on "software patents," which are perceived negatively around the world. It is hard to enforce a patent on an innovation that heavily draws on software or automation.

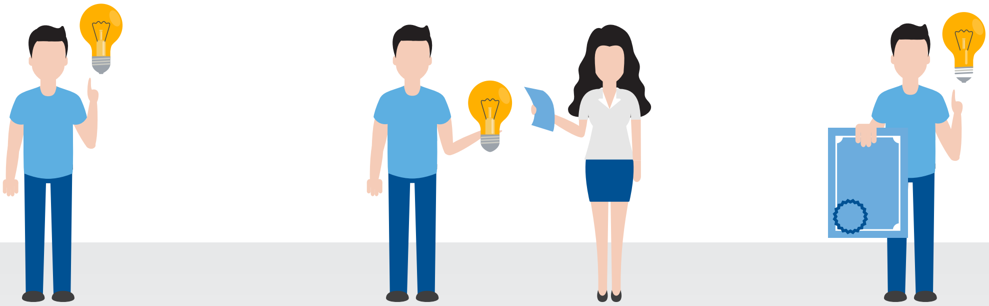
In general, to be awarded a software patent, the invention must provide a real-world improvement – not just better working software. For example, today a compression algorithm is typically considered patentable, as would a more memory-efficient matrix multiplication technique. An algorithm to accurately predict the next soccer world cup winner would not be patentable.

DATABASE RIGHTS

A relative newcomer in the IPR world is the database right. Introduced in Europe in the late 1990s, the database right protects a collection of information against copying and reuse. The main requirement to qualify for a database right is that substantial investment was made in the creation or maintenance of the data in the database. As with copyright, no formal registration or application is required.

Examples of protected databases include online dictionaries, labelled image collections and source data for cartographical maps. In all cases, the data must be organized for search and browsing.

Outside of the European Union, however, the database right is not recognized, which further complicates IPR. The U.S. has a long-standing legal tradition that collections of data are not protectable by IPR; only creative works can be protected under copyright.



TRADE SECRETS

The status of trade secrets in the IPR world differs around the world, but in general, misappropriation of well-protected information is actionable by law. In this instance, the owner of the information would be required to show how it applied adequate security measures against unauthorized access. A would-be trade secret thief could then counter by proving that the information was already available in the public domain.

Typically, companies guard their trade secrets by signing non-disclosure agreements (NDAs) with customers or other third parties. Strict contractual obligations then prohibit copying or reuse, in some jurisdictions strengthened by contractual fines or other legal measures. NDA provisions may also be present in other agreements. However, someone who learns the confidential data from a legitimate purchase of a product is not bound by such provisions, even when using special techniques such as reverse engineering. This limits the strength of trade secret law.

IP PROTECTION FOR MACHINE LEARNING

A competitor or other entity with less than noble intentions has various options to profit from the work or investments made by the creator of an ML system. Given the unique nature of ML, the question then arises: How can IP law be applied to protect the various aspects of this novel technology?

PROTECTION OF THE TRAINING SET

Creating a good training set for a particular ML application can be a time-consuming and expensive effort and, although in a typical setting an infringer has no direct access to this training set, the ability to copy the set is easy if the person has access. This is where IP law comes in.

A training set would be protected with a database right, if the owner of the training set has its principal place of business in the European Union. However, such right would only be enforceable against an infringer in the same jurisdiction.

Whether copyright can be claimed on an ML training set is a more difficult question. A training set is not created to be a piece of art. The typical intention is to ensure the data fits the use case. Creating a well-fitting set of data on a topic is not a creative activity under copyright law. One potential copyright claim are the data classification descriptors.

If categories are chosen through a creative process – "beautiful/ugly", "strong/weak", "big/small" – then the training set could be said to be protected by copyright through this creative labelling. A classification based on factual elements – "cat/dog", "traffic light/streetlight/parking sign" – does not impart creativity and therefore does not allow for copyright protection.

In some applications, training sets are generated by simulation or other artificial means. Arguably, these training sets could be copyright protected as the choice of how to simulate or generate could be seen as a creative choice. However, to date, this has not been challenged in court.

Companies will often consider their training sets to be carefully guarded secrets. Since a training set is not required to be shared for the ML model to be used, it seems straight forward. The best approaches are to both guard the training set from illicit copying and apply strong contractual restrictions to parties that must have the training set.

PROTECTION OF THE TRAINING PARAMETERS

The training set and model are only a part of the value of a good ML system. The parameters that steer the training algorithm may also have value: choosing the right training parameters takes time and effort from highly trained engineers.

For the set of training parameters that create the ML system, copyright protection is a sensible approach. If a data scientist determining these parameters uses creative efforts to select the right training parameters, the resulting set of parameters would likely be protected by copyright. But if the training parameters were found through exhaustive search (e.g. evaluating a number of options proposed in the literature) or algorithmic process, no copyright would be available. The same would apply to the model that is produced using those training parameters and a given training set.

A database right is least likely on the parameter set because one criterion for database rights is that it must concern a collection of individual elements that are systematically or methodically arranged. A parameter set is unlikely to fit that criterion.

PROTECTION OF THE ARCHITECTURE

The architecture of the system is the underlying foundation for the ML system. Its design is a key aspect of the proper functioning of the system. After training, the architecture can be put in practice.

A system like this has two aspects: the graph defining the architecture and the software implementing it. The graph is protected under the same conditions as given for the protection of the model parameters. Patents would theoretically be available for innovative hardware aspects of it, but this is unlikely given most innovation in this area is purely software. The software that implements training and/or inference would typically be protected with copyright, as it is principally software designed using creative efforts.

PROTECTION OF THE ML SYSTEM

In theory, a computer system programmed with a well-chosen parameter set and trained on a specific training set could fall within the realm of patentable subject matter. However, current case law in Europe and the United States would require the system to be designed to perform real-world tasks such as steering a car or recognizing images from the real world. To date, it would be speculative to conclude a patent is obtainable on an ML system that operates in a more abstract manner, e.g. recognition and/or classification without a specific use case in the real world.

The software of the ML system could be protected by copyright just like any other software.

A database right for the ML system is theoretically arguable: in a way the dataset is made searchable through the model and the software executing that model. However, this has never been decided in court or outlined in legal literature.

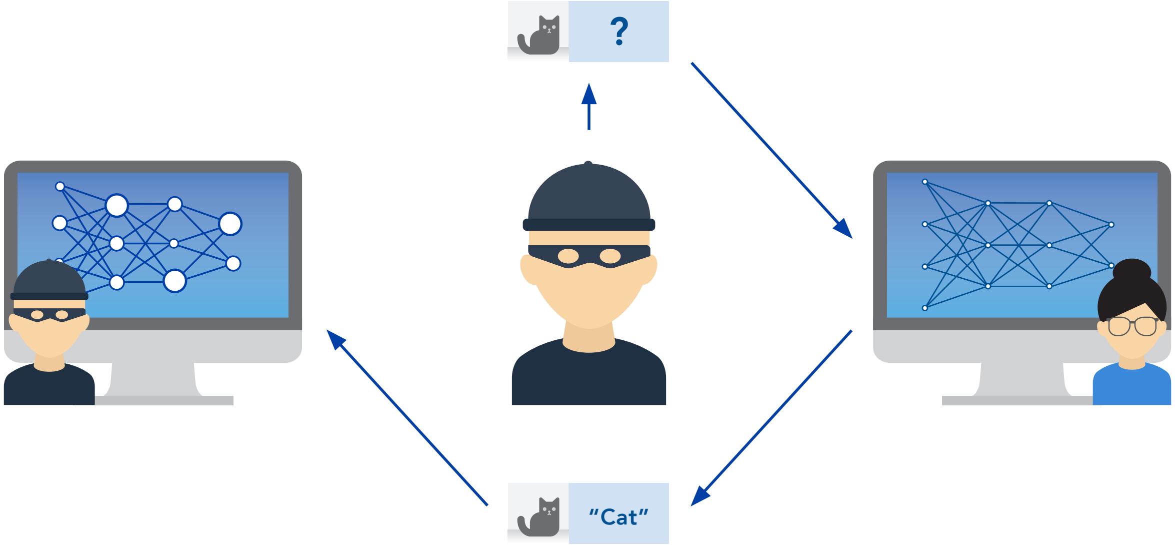
STEERING A CAR

RECOGNIZING IMAGES FROM THE REAL WORLD.

BURDEN OF PROOF

Spotting an infringer and proving infringement in court are two very different things. The burden of proof in court cases on IP can be high to meet. As a general rule, the courts need to be convinced that it is very likely that something was infringed. The alleged infringer has no obligation to cooperate in delivering this proof. Therefore, if some evidence under his control is needed, the IP rightsholder may have a problem. Some jurisdictions allow for seizure of evidence or require parties to engage in a discovery process, but it is far from certain that it provides the rightsholder the evidence needed.

Under copyright law, if two items are very similar, then a court may reverse the burden of proof: the infringer then must show his work was independently created. This is a very fact-specific analysis upon which a rightsholder should not rely. Under trade secret law, a rightsholder sometimes has the option to request the court to keep evidence secret, or to get an independent party such as a notary public to compare evidence against the secret information without having the secret become part of public court records.



PROTECTION OF THE MODEL AGAINST COPYING

When an ML system is available without contractual or usage restrictions to the public, a unique way to copy its functionality becomes available. Essentially, the copyist has a dataset of unclassified items and submits each item to the ML system. Each answer is carefully recorded as the classification of the copyist’s dataset. The obtained labelled dataset can then be used to train a model of similar quality. It has been shown that this works effectively, even if the dataset contains non-problem domain data and if the architecture and model parameters of the target and clone do not match. Under copyright or database law, it is unclear if this act is legal or not. The dataset from the original ML system is not copied; only its output is used, and then only to label a different dataset.

If the dataset classification is creative in itself, the copyist may infringe that copyright by reusing the labels. This could even apply if only the labels are copied and reused to classify a completely independent dataset. However, this has never been tested in court.

WATERMARKING IN MACHINE LEARNING

One practical aspect of IP law is that a rights holder has to prove that their rights have been infringed. Proving ML models or training sets are copied may be exceptionally hard. Especially when the data concerns real-world elements. The copyist can then easily argue they merely collected the same or highly similar data from the original source location. Without a way to counter that argument, the rights holder would be left with no recourse.

> Watermarking is the process of embedding information in the content, the embedded information may not be apparent upon normal observation. The term "digital watermarking" was coined in 1992 and has been used by rightsholders since the late 1990s for detecting and possibly tracing leaks of movies and songs. The embedded information of a digital watermark can reveal the source of the leak, or the network that originally broadcasted that content.

With NXP Semiconductors eIQ® Model Watermarking tool addition in its eIQ Toolkit for machine learning development, watermarking has also found its way into ML. The tool provides a workflow for the developer to extend the original training data with so-called trigger images that are generated by combining images from a given class with a secret drawing provided by the developer. These trigger images get as label a "watermark class", which is a selected class different from the actual class of the underlying image. Training with this extended training set results in a model with a unique functionality on trigger images. This functionality is the watermark of the ML model. When trigger images are presented to an independently trained model the resulting classification is of the actual class underlying the trigger images, but both the originally trained ML model as well as a system that copied the watermarked ML model would return the "watermark class" as classification. This would show that the model was copied from its original.

An additional benefit of the NXP eIQ Model Watermarking tool is that the watermark is based on a creative element --- the secret drawing --- thus adding a piece of copyright-protected information to the ML model. This helps strengthen a copyright claim towards any copyist.

The copyist could counter-argue that they employed the same watermark independently, or actually created the watermark themselves to reverse the allegation of copying. To address such arguments, copyright owners must keep clear records of dates and times when the watermarks were chosen and inserted. Without good proof, a copyright holder would not be able to establish a claim infringement. With NXP's eIQ Model Watermarking tools the necessary records associated with the inserted watermark are captured and further instructions on creating the necessary date and time records are provided to the developer. Furthermore, the NXP eIQ Model Watermarking tool is optimized to incur no performance penalty on the model.

FUTURE OF ML AND IP

ML-driven business is gaining more and more traction. Interest in IP rights is also increasing to protect investments, from copyrights on training sets to patents on classification systems. Current IP law and practice is evolving, and case law is sparse. It's uncertain how legal protection for ML-systems and ML-driven products will mature.

However, some general indications are available:

	Intellectual Property Right (IPR)			
	Patent	**Copyright**	**Database right**	**Trade secret**
Protects	Technological innovation	Creative expression (i.e. not just hard work or investment)	Substantial investment in creation of the collection	Information is kept secret (e.g. by NDA)
Jurisdiction	Worldwide	Worldwide	Owner and infringer must be in the EU	Worldwide
Protected item				
Architecture	No, but see software below	Unlikely for underlying graph, unless creativity in choices	No	Yes
Training set and Test set	No, but see software below	No, except creative labels or creatively handpicked dataset	Yes	Yes
Training parameters	No, but see software below	Unlikely, unless creativity in choices	No	Yes

	Intellectual Property Right (IPR)			
	Patent	Copyright	Database right	Trade secret
Model	Unlikely	Unlikely unless creativity in watermark, labels, parameter or architecture choices	Unlikely	Yes
Software implementing the ML functionality	Yes, as part of trained system with the model and only when aimed at real world tasks	Yes, but the functionality that it implements is excluded from protection	No	Yes

In this white paper, we have outlined which ML IP could be protected by which IP law in the future. What does this mean for our introductory example of the maintenance of capital equipment? Although the ML model maintenance is not patentable, the implementation of the model might be, since it is aimed at performing real-world tasks. Furthermore, a copyright claim can be made on the software implementing the ML algorithm. However, if a copyist only copies the model (weights) and uses it in her own implementation or if they create a clone of the model by labelling their own training set, copyright protection is very uncertain. The developer must prove that creative choices were made in either the architecture design, training parameters, composition of the training set, or labelling of the data – and that these choices were not just made out of technical considerations. Even if they can prove this, it is uncertain if this creativity is sufficiently present in a clone or copy of the model such that a claim would hold up in court. This makes the development of countermeasures that either protect against cloning or copying (e.g. platform security) or include creativity (e.g. watermarking) critical in the protection of ML IP.

How infringement cases will be judged and whether the law will change in these matters is speculative until there is precedent set in court. Despite this, we at NXP offer the eIQ Model Watermarking tools within our eIQ ML development environment to help establish copyright protection for your ML IP by providing copyright ownership and proof of unauthorized copying.

NXP enables advanced solutions for a smarter world that makes lives easier, better, and safer. As the world leader in secure connectivity solutions for embedded applications, NXP is driving innovation in the automotive, industrial & consumer IoT, mobile and communication infrastructure markets. NXP also offers to protect your ML model, on top of the embedded platform security.

After training an ML model, it will be deployed in a system where it can be used for its intended purpose. The NXP® eIQ™ machine learning software development environment enables the use of ML algorithms on NXP i.MX RT crossover processors, and i.MX and i.MX RT families of SoCs. eIQ™ includes inference engines, neural network compilers and optimized libraries. It also includes methods that make the ML network more secure by addressing issues such as cloning, as discussed in this paper, and adversarial attacks. Other ML security measures are on the roadmap. The eIQ software is fully integrated into our MCUXpresso SDK and Yocto development environments, allowing you to develop complete system-level applications with ease.

For more info see: **nxp.com/design/software/development-software/eiq-ml-development-environment:EIQ**

TAKE THE NEXT STEP

To learn more about NXP's innovative solutions for Security and Machine Learning, visit nxp.com/ai

www.nxp.com

Document Number: IPMLWP REV 1