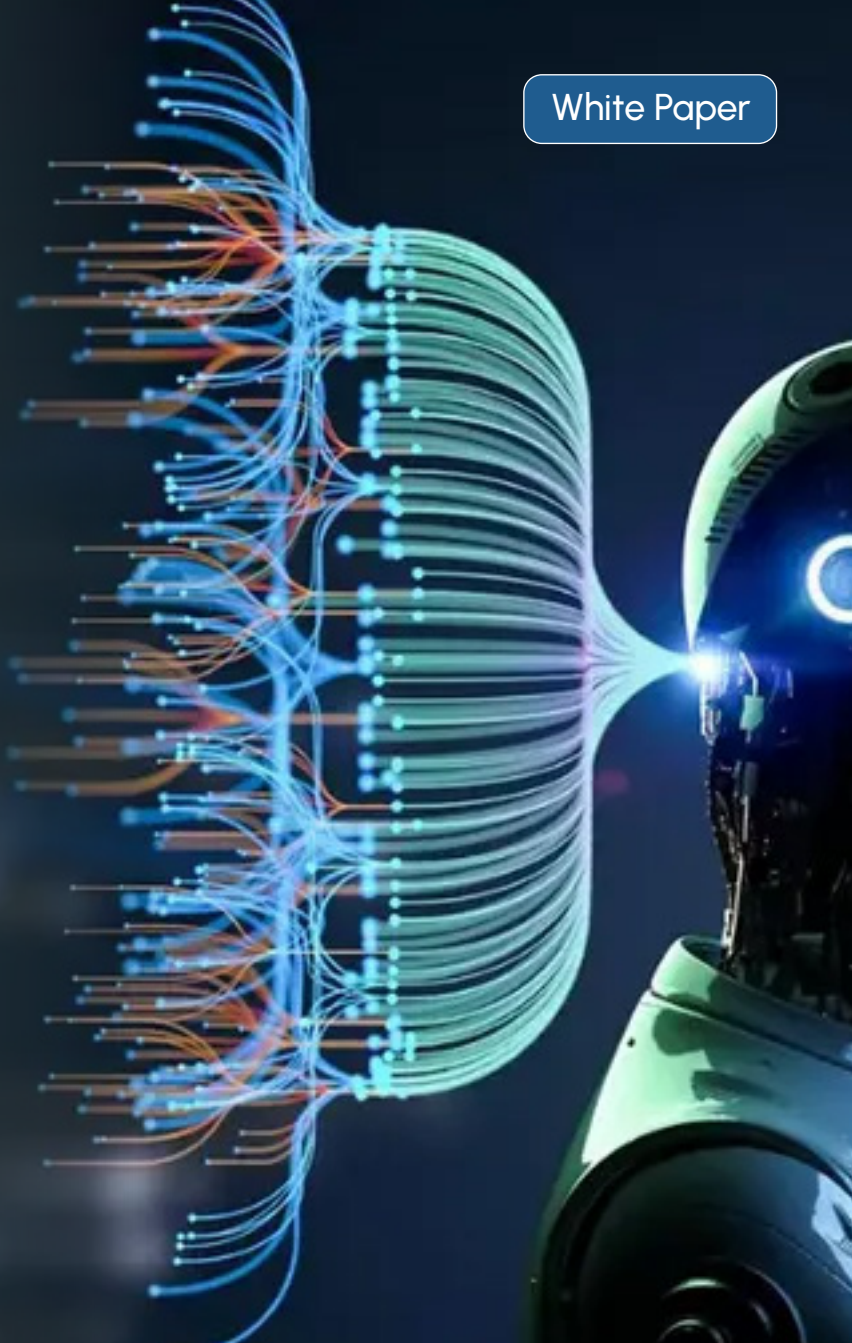


reveal

Pattern Recognition Software for Legal Compliance



Complying with legal requirements is a challenging task in the best of times. We need to dig through data whose volume and varieties continue to grow at astounding rates. As we scrutinize that data, we need to have at hand the legal requirements we are trying to meet, requirements that keep morphing and expanding.

Technology helped get us into this mess, and technology can help us get out. One prominent path out is the use of data pattern recognition software.

TABLE OF CONTENTS

1	<u>Legal Requirements</u>	3
2	<u>Compliance Data</u>	4
3	<u>Complying with Legal Requirements</u>	4
4	<u>Software to Help Recognize Patterns & Achieve Legal Compliance</u>	5
5	<u>Pattern Recognition Software: A Form of AI</u>	6
6	<u>Types of Pattern Recognition</u>	7
7	<u>Ways Pattern Recognition Software is Used</u>	11

Legal Requirements

Before we get into how data pattern recognition software is used to help organizations comply with legal requirements, let's start with the problem to solve: compliance.

Your organization is beset by an ever-expanding set of laws, regulations, and business needs mandating compliance on your part. Just in the privacy arena, there is a profusion of requirements. There are well-established laws, such as the Health Insurance Portability and Accountability Act of 1996 (HIPPA), with extensive and comparatively clearly delineated requirements to be met and consequences for failing to do so. There are newer regulations, such as the General Data Protection Regulation (GDPR), whose boundaries are starting to set but nonetheless can be unclear. And there is a host of new, emerging, or gestating US-based data privacy laws, from the California Consumer Privacy Act (CCPA), passed in 2018, to the Virginia Consumer Data Protection Act (VCDPA), signed into law on March 2, to the Utah Consumer Privacy Act, introduced in February.

Should you fail to comply with the various laws, regulations, and even internal policies that apply to you, the penalties can loom large. They may include fines, sometimes substantial, potentially increasing daily, that can range from hundreds of dollars to millions, even, in situations such as cartel infringements, up to 10% of global group turnover. Failure to comply can go beyond monetary penalties to damaged reputations, loss of goodwill, forced divestitures, and even imprisonment.

Compliance has many components, of course. You should know which rules, policies, and processes regulate your business practices. You should have policies and procedures in place, based on those rules and requirements. You should make sure you have the right people in place to implement and manage to those policies and procedures.

And, to support all that, you need to find data. You need to be able to locate data showing efforts you have taken to get into compliance or stay there. You also need to be able to get at data suggesting failures to comply.

Compliance Data

Your organization has data, lots of it.

Even a small organization can have terabytes of data, large and even medium ones easily have petabytes. Specific numbers are hard to come by, but as far back as 2015 researchers estimated that **1/5 of Australian businesses** were managing **more than one petabyte of information**.

That data has been created in many ways. Everyone has communications data. This includes email, of course; various forms of short messages such as chat, text, and Slack; and data from collaboration platforms such as Zoom and Team. Almost everyone uses some sort of productivity apps. These might be Word, Excel, and PowerPoint from Microsoft, or comparable tools from Google or Apple, or even open-source versions.

The data exists in different forms. Much of big data is text. It can be images, with or without searchable text, or audio, or video. A lot of data resides in databases of various types - SQL, MySQL, and so on.

Complying with Legal Requirements

To comply with legal requirements, we need to go through a process that boils down to this:

- ✓ **Identify** the application law, rule, regulation, etc.
- ✓ **Determine** what that law, etc., requires.
- ✓ Amass potentially applicable data.
- ✓ **Look for patterns** in the available data that tend to indicate that the compliance requirements were, or were not, met.
- ✓ **Assess** any patterns identified.
- ✓ **Repeat** as needed, if, for example, you need to determine compliance with more than one rule.

Software to Help Recognize Patterns and Achieve Legal Compliance Requirements

Sometimes we can perform this entire process manually. However, with the volumes and variety of today's data and the multiplicity of legal requirements to be complied with, a manual approach often is not enough.

When that happens, we can turn to technology for help with data mining. Maybe not the same technology that got us into this mess, but technology nonetheless.

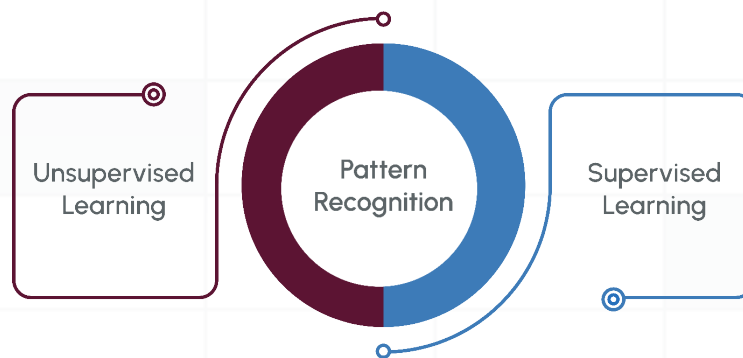
One particularly powerful form of technology we can turn to is pattern recognition software.

This software typically is used in two main ways. Pattern recognition algorithms are pointed at datasets with the goal of using classifiers to identify meaningful patterns in that data, and it is used to categorize the content in real-time. It might be used to find documents with similar content, such as email messages discussing a particular issue. It might be deployed to find examples where someone's behavior deviated from the norm. For example, someone who almost always communicated with co-workers primarily through business channels such as corporate email addresses and only during standard business hours, then abruptly began exchanging SMS messages with a small cadre of officemates during evenings and weekends. It could be used to find patterns suggesting inappropriate sexual comments or offensive language.

Pattern Recognition Software: A Form of AI

Pattern recognition software is a form of artificial intelligence (AI) software.

AI comes in various flavors, including machine learning. Machine learning, in turn, comes in different subsets including supervised learning and unsupervised learning. Both supervised and unsupervised learning have subcategories. One subcategory, which falls under both these forms of machine learning, is pattern recognition.



Pattern recognition techniques are not new. In a 2001 book review published in IEEE Transactions on Neural Networks , reviewer Ke Chen described pattern recognition as “a process of description, grouping, and classification of patterns”. When used in supervised learning, Chen continued, pattern recognition “identifies an unknown pattern as a member of a predefined class”. I might, for example, predefine a class by training an AI model to look for documents containing warnings not to drive a utility side-by-side vehicle at higher speeds. I then might ask the supervised learning tool to look for and return more documents like those.

When used in unsupervised learning, Chen noted, pattern recognition “groups input patterns into a number of clusters defined as classes hereafter. For automatic pattern recognition, the primary task consists of feature extraction and classification”. The Brainspace cluster wheel is an example of a data visualization tool using pattern recognition with unsupervised learning.

Types of Pattern Recognition Software

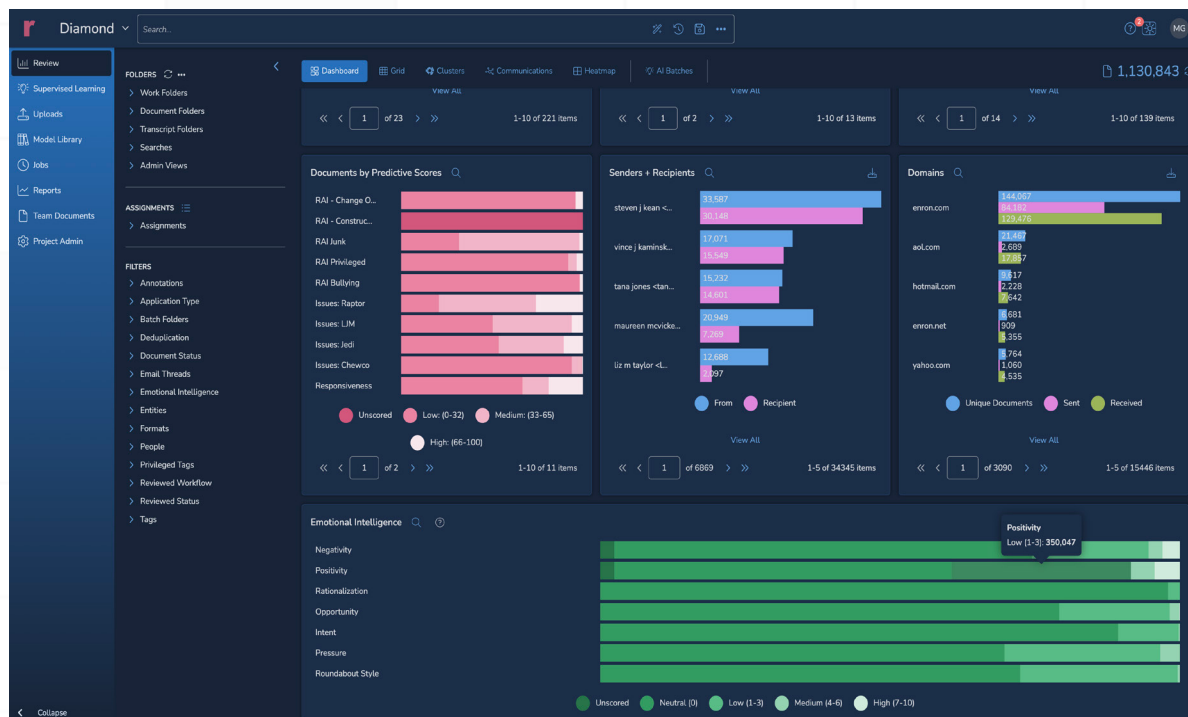
Data Clustering

A type of data analytics that inspects textual documents and groups conceptually similar documents. Clustered documents may relate to a subject or a type of communication. For example, documents in an Earnings Call cluster might be given a higher cluster score for being closer to the center of the cluster (for example, a report or transcript of such a call as opposed to preparatory discussions).



Sentiment Analysis

This is a type of data analysis that reads the tone of a segment - all or part of a document or message. Sentiments can be negative, associated with negative connotations. Sentiments can be positive. A single segment could contain only negative sentiments. It could contain only positive ones. It also could contain both negative and positive sentiments, as with the segment, “Everything is a mess and we need to shut this down right away. However, the staff is nice”.



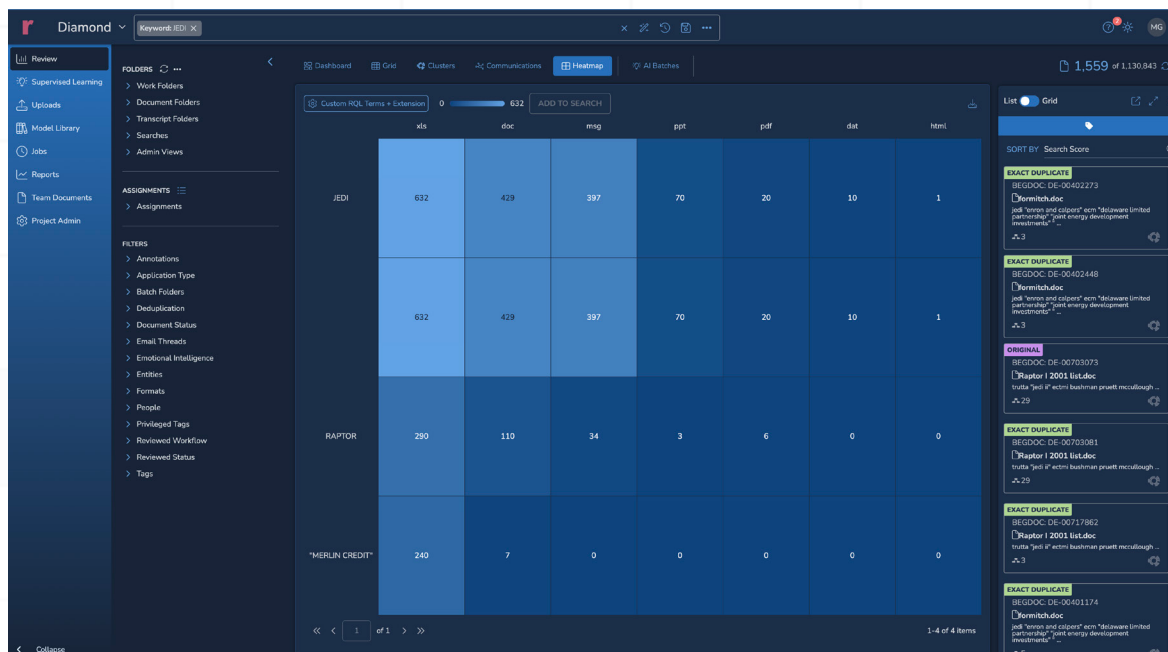
Entity Extraction

Input data and find named entities in text and classify them into pre-defined categories. A named entity is an extracted piece of data identified by a proper name. A named entity could be a person (“George Socha”), a geo-political unit, such as a city, state, or country (“Minnesota”), or an organization (“Reveal”). Other entity types include money (current discussed), temporal (dates discussed), law (legal jargon), quantity (measurements), location, technological jargon, and so on.



Term Heat Map

Enter search terms or select Custodians and identify hidden connections or patterns between these values and selected metadata. Identify anomalies in the data where higher document volumes may exist between intersecting metadata values. Make connections between people and topics or timeframes of interest.



Ways Pattern Recognition Software is Used

Five ways pattern recognition can be deployed **separately or in groups.**

- Combine email and document text analytics with entities and concepts, using AI pattern recognition software to help you better understand the meaning in the words of the files you are searching.
- Use unsupervised machine learning algorithms to classify content, helping you find important content you may not have anticipated existed.
- Use supervised machine learning algorithms to classify content, helping you find more like this - more content similar to the content you already have and perhaps content that fills in the picture more completely.
- Combining anomaly detection and social network analysis, helping you better identify and understand the behaviors of key individuals or organizations.
- Complying with legal requirements is a challenging task, but by thoughtfully using data pattern recognition software, you can go more quickly and more reliably find the data that matters most.

If your organization is interested in leveraging pattern recognition and machine learning software, contact Reveal to learn more. We'll be happy to show you how our authentic artificial intelligence takes review to the next level, with our AI-powered, end-to-end document review platform.