Artificial Intelligence (AI)

Feith uses Artificial Intelligence as a tool to analyze, predict, and categorize FOIA content without explicit instructions or programming. AI allows you to reduce costs, increase reliability, gain speed, increase volume, and understand what your goals are.

Natural Language Processing

Feith has been using Natural Language Processing (NLP) to handle massive amounts of documents that contain massive amounts of information for many years now. Implementing Natural Language Processing allows your computer to understand spoken or written text.

In addition to storing documents with the appropriate metadata sets, Feith also utilizes entity extraction. Entity extraction captures what standard metadata cannot, such as names, places, organizations, dates, states, crimes, or any number of subjects. Feith can also identify key information from text data. Once identified, the data can be categorized into predefined categories. If this process were to be done manually, a knowledge worker would have to go into each individual document and mark down all the proper names and places.

Feith can also utilize relation extraction to relate words, phrases, or sentences to other words, phrases, or sentences. This AI feature extracts the relationship between two entities in unstructured sources. For example, if a specific pronoun is related to a certain part of a sentence.

Feith uses causation detection to identify the underlying web of causes or behavior. It can then derive insight from that cause or behavior. For example, if someone were to write an email containing the term stomach, the system would be able to relate it to the term “anatomy”. The system analyzes the sentence structure and detects the fact that there are relationships within the text, scores its confidence level based on those relationships, and also extracts out other concepts from other datasets.

Natural language processing gives users a visual representation of the standard things that all software should be able to do to reduce costs, increase reliability, gain speed, increase volume, and understand what your goal is. The result is a payoff where users are getting valuable information out of the system, and FOIA workers can relate data that would previously be unrelatable.

Automatic Categorization

AI for categorization and classification teaches computers to predict outcomes based on data, and how a document should be classified or categorized based on that data.

Feith’s module for automatic categorization uses a feature called regular expression and computing, which means that it will find the exact word being searched or find a word or phrase that follows a given pattern. For example, if every 1099 MISC document contains the term “1099 MISC” at the bottom of the page. The system will be confident that the document is a 1099 MISC document because of its format.
Feith’s AI module for automatic categorization classifies documents that can’t be categorized by normal means. The solution multi-tags the document by using the entire dataset behind it to decide what words and phrases stand out from these documents. The Feith interface allows users to give the system an uncategorized document to look at, and the system comes back to them with all other uncategorized documents that it believes are similar. If the documents that the system found are similar, the solution will recommend categories based on its level of certainty that the term applies. The system can then be trained how to categorize documents more accurately based on the examples that it has been given. Once it has been trained, it can be trusted to categorize content with 90% or higher accuracy.

Recognize PII and CUI
To recognize PII and CUI, a computer needs to be trained to understand both text content and images. Feith’s AI module uses the text, the categorization, and also the images within the FOIA documents that the organization works with.

Feith uses a technology called computer vision to recognize images as certain “types” based on the sample data given and all the background information in the system. When documents enter Feith, they get broken down into text and images, where the text goes to the text engine and the images go to the image engine. The image engine can find things like driver’s licenses, signatures, passports, photos, and anything necessary to complete a FOIA request.

Feith also uses computer vision for signature identification. For instance, the solution can determine if a 50-page contract is signed. Feith’s computer vision can also detect if the contract has been annotated in a way that there are strikes through certain content. Feith’s image engine also does not need many training examples to function on its own.