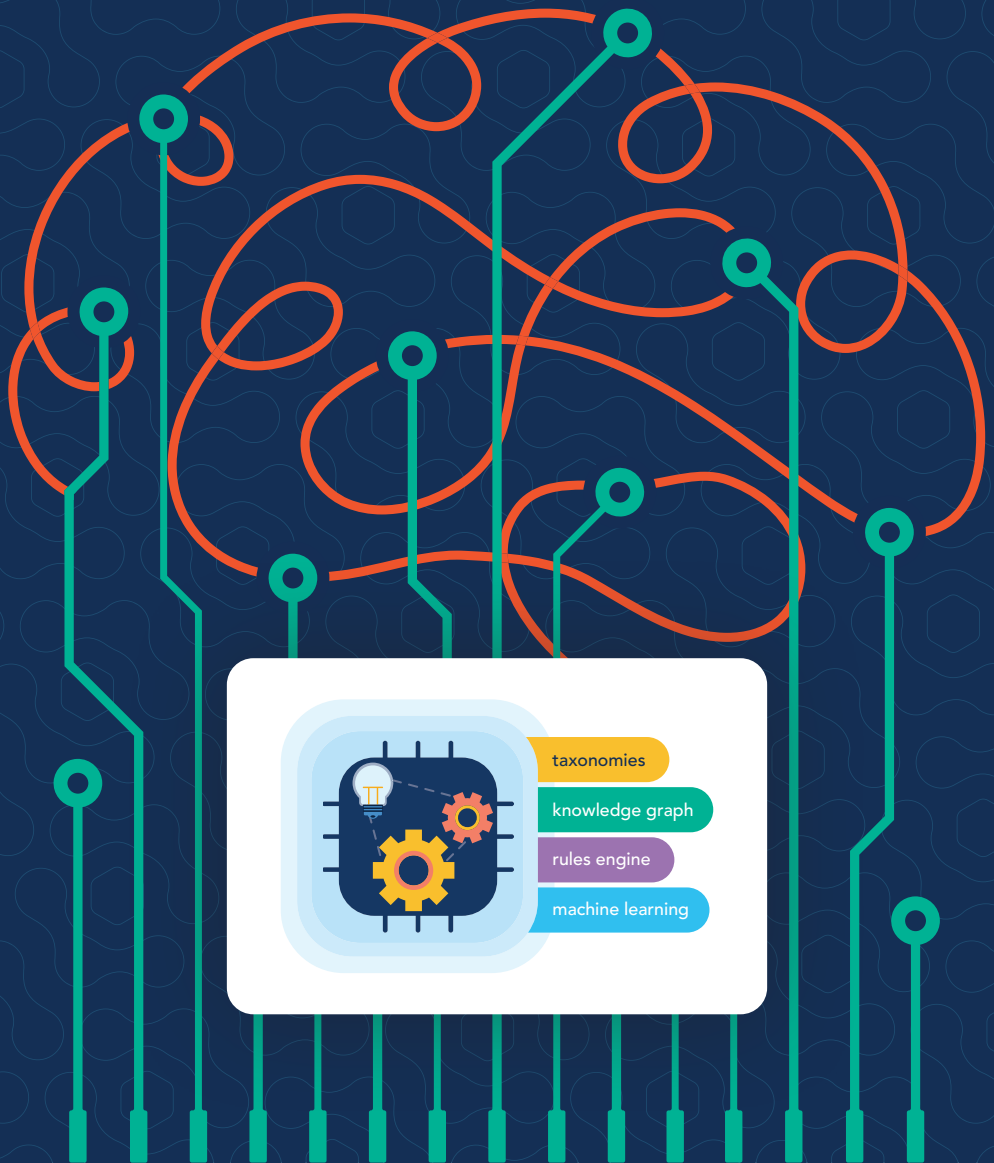
expert.ai

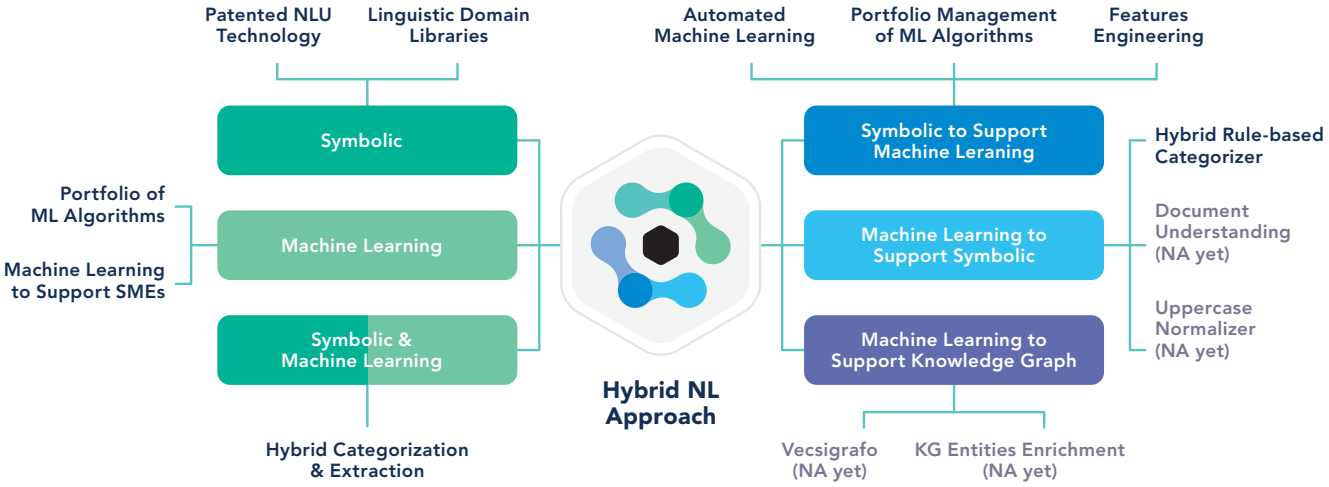
How the Expert.ai Platform Impacts Your Approach to NL



Machine learning and symbolic approaches are staples when it comes to natural language processing solutions. They each offer unique capabilities that enable organizations to process and analyze their most underutilized source of value: unstructured language data.

Though machine learning and symbolic approaches have long been viewed in opposition to one another, it is their pairing in a hybrid approach that will soon become commonplace. In fact, the hybrid approach is the driving force behind the expert.ai Platform, a development platform where multiple approaches can be used to power natural language solutions.

So what exactly does a hybrid natural language approach do that an individual approach does not? We break it down on three separate levels.



Hybrid Approach at a Platform/Pipeline level

One of the core functionalities of the expert.ai Platform is that it enables subject matters experts and other users to leverage different techniques for different tasks within a single use case. The subject matter experts have access to a robust knowledge graph and a development environment in which they can write rules. They also have access to a curated list of machine learning algorithms. They have the freedom to apply an individual algorithm or a combination of machine learning and symbolic at various steps in the pipeline.

Semantic Understanding to Improve/Feed ML

Consider a customer using the Platform to categorize press releases or news related to specific events (e.g., corporate financial transactions). The Platform enables the customer's subject matter experts to create a single symbolic rule that identifies "buy-sell" verbs with high precision. It does so independently of verb tense through generalization or disambiguation enabled by the knowledge graph.

For example, the knowledge graph can correctly identify the following as a single event and time designation with one line of code — a task ML cannot realistically accomplish.

- "Will acquire", "will purchase" and "to be acquired"
- "Has acquired", "has purchased" and "has completed the acquisition"

At the same time, the subject matter experts can use a standard categorization machine learning model to recognize "buyer" and "target" companies. By pairing the symbolic approach with the machine learning model in the pipeline, they can produce a composite relationship in which a "buyer" has bought or plans to buy a "target" company.

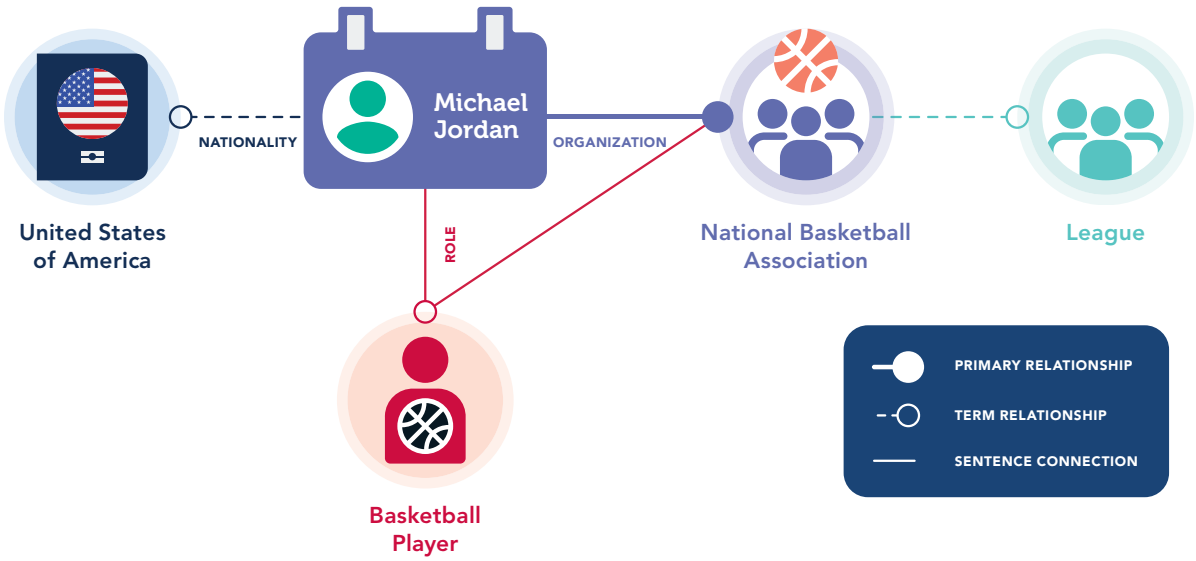
Knowledge Graph Entity Enrichment

Different industries and enterprise environments make use of extensive technical and industry terminology including large numbers of specific entities (products, components, companies). Unfortunately, most organizations do not have structured dictionaries, glossaries or lists that house these entities, despite their value towards knowledge discovery and management.

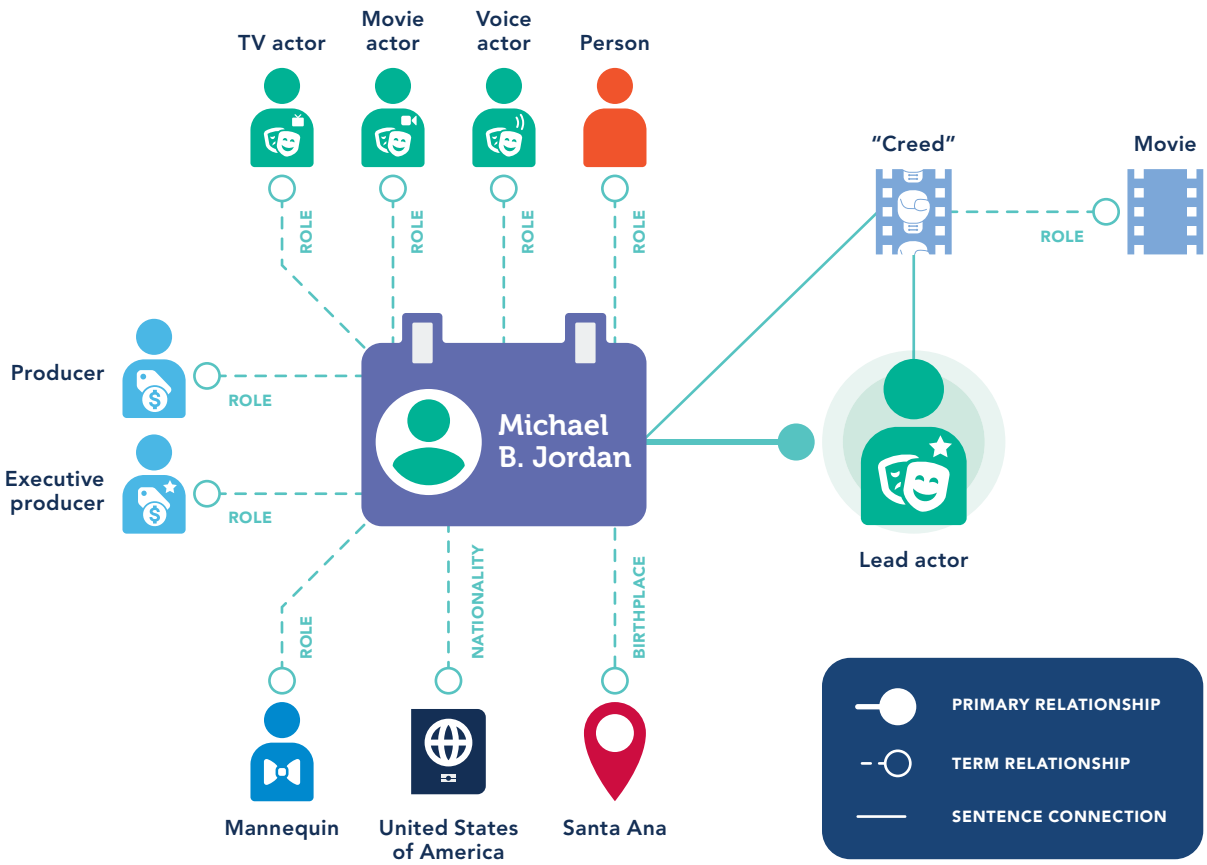
Manual creation and curation of related terminology requires reading or scanning documents and assembling a knowledge base. By combining expert.ai's rich, embedded language knowledge graph with ML algorithms, you can **automate the processing of relevant documents and create your own industry- or enterprise-specific knowledge graphs**. These extensions to the knowledge graph leverage the position, attributes, and relationships of these specific entities to ensure an accurate and effective understanding of domain- or company-specific content.

Consider the following sentences:

> Michael Jeffrey Jordan is an American former professional basketball player in the National Basketball Association (NBA) during the 1990s.



> Michael Jordan starred as the lead actor in the movie Creed.



With extensions to the knowledge graph, you can correctly identify the context of specific entities, allowing you to identify similar documents and reduce ambiguity throughout your document corpus.

Hybrid Approach at a Work Process Level

Natural language models are notoriously difficult to scale. Machine learning alone does not scale if the data is not already annotated. Symbolic alone does not scale in specific circumstances such as high-volume tasks (e.g., large categorization taxonomies or large numbers of different entity types to be extracted). With the Platform, a subject matter expert can exploit the knowledge graph and rules to accelerate the process of data exploration and annotation.

Utilizing a hybrid approach, we can alleviate some of the scaling issues presented by pure machine learning and pure symbolic approaches. With pure machine learning, the number of required annotations grows exponentially with higher model performance. Likewise, the number of required rules also grows exponentially with a pure symbolic approach.

Our Platform allows for symbolic and ML to work in unison with each other, requiring fewer rules and fewer annotations to achieve the same result. Additionally, you can take it a step further by allowing our smart algorithm to generate the symbolic rules for you automatically, while still only requiring a fraction of the annotations for similar model performance.

Accelerated Annotation

The Platform enables subject matter experts to use simple rules — perhaps focusing on recall without expecting too much from precision — to predict the presence of the following in a document:

- standard and custom entities
- domain-specific terminology
- sentiment
- intentions
- emotional traits
- relationships
- other concepts

After running a first analysis of a corpus using the simple rule set, your experts can validate the results or correct the "predictions" by turning off false positives or even turning on false negatives in a further reiteration to optimize the annotation. They can then use a specific feature of the Platform to propagate the symbolic logic starting from the completed annotations — regardless of whether the annotation was made by a person or the system via rules. This results in faster annotation which increases the scalability of an ML-driven model.

By leveraging the symbolic approach and the related, immediately available rule set, the Platform enables subject matter experts to:

- Annotate faster than you would starting from scratch (see above)
- Address a specific task directly and ONLY with the symbolic rules (when the rules are already ensuring good results or just need to be incrementally improved)
- Determine whether a certain task needs a "boost" from the machine learning technique

Support in Taxonomy Building

An effective classification project requires a disciplined and thorough set of tasks before designing and building the model. These tasks include analysis of a representative set of documents as well as the creation of a taxonomy for classification of the documents. Our platform can assist in both tasks.



Once a representative corpus is selected, the results of an extensive linguistic analysis are graphically presented to you in an intuitive way. This includes relevant data such as concept and lemma frequency, priority topics and the frequency of more than 20 types of entities (e.g., people, organizations, monetary values, etc.).

As the linguistic analysis is taking place, you can leverage the ML-based 'Magic Taxonomy' functionality to automatically create N clusters of concepts from the contents of the documents. The number of initial clusters is defined by you or automatically by the system and each group is assigned a label, including the most relevant concepts in each cluster.

From there, subject matter experts can edit the label to make it meaningful for the taxonomy they are trying to build. The output is a set of automatically system-predetermined classes that give you a solid starting foundation for a classification taxonomy.

Hybrid Approach at the Technology Level

Lastly, we consider how a hybrid approach can influence both semantic and ML technologies. While both symbolic and ML approaches can be powerful on their own, a hybrid approach enables you to harness the capabilities of both to simultaneously complement each other. The following are prime examples of how a hybrid approach can impact integral processes in the building of language models.

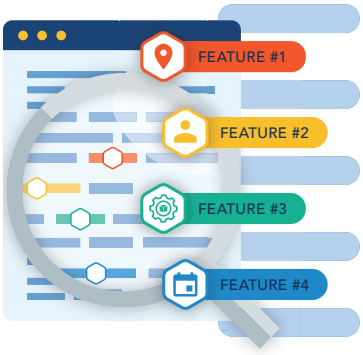
Automatic Rule Generation

A symbolic rule-based approach can efficiently produce fully explainable (in terms of computer resources) and highly accurate results for automatic classification and extraction. However, this technique is perceived to be less scalable due to the complexity of writing rules from scratch.

The hybrid technology within the Platform can **use ML to automatically curate a set of symbolic rules by taking annotations as input**. This directly addresses the scalability issue. With a symbolic rule set quickly generated via ML, the Platform transforms the black box ML approach into a fully explainable and extensible model. These rules can then be easily extended and fine-tuned via the symbolic approach for unrivaled quality control.

Optimizing Feature Engineering (and ML Algorithm Selection)

When we consider the typical machine learning model development workflow, there are three critical and time-consuming tasks: annotation, features engineering and parameter optimization. Our hybrid technology helps to accelerate each step of the process.



Feature engineering is one of the most important steps in machine learning. It is the process of transforming raw data into features (i.e., input variables for predictive machine learning algorithms) via domain knowledge of your training data. Creating and selecting these predictive variables helps your ML model better understand your desired outputs and, ultimately, deliver more accurate results.

The feature engineering process can be long, tedious, and often frustrating, so it is easy to take a shortcut or two along the way. Typically, standard open-source or legacy NLP libraries or APIs only offer basic capabilities to extract information from a sentence. This forces you to feed algorithms similar sets of features (e.g., same number of words, capital words, punctuations, stop words, etc.) that are not highly representative of the content. Once these basic features are identified, you test and progressively refine them to optimize your model.

By leveraging symbolic analysis on top of the basic NLP capabilities, you can automatically leverage a much richer set of features (e.g., the most relevant concepts, word relationships, context-specific word disambiguation, topics, etc.) that are more representative of the content. This also enables symbolic capabilities, such as navigating relationships of terms and concepts within the knowledge graph, that help you identify use case-specific elements in the text and give you more flexibility and power when developing an ML model.

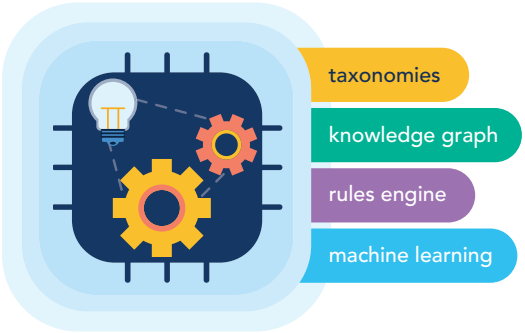
In applying symbolic techniques to enrich the features of your input data, you enable your organization to:

- speed up training with fewer examples and less training time
- improve performance while expending the same amount of manual labor and computing energy

The True Value of the Platform

You need every advantage you can get when it comes to turning language into data. By adopting a hybrid approach to NLP, you have all of them at your fingertips. And it has never been easier to do so than with the expert.ai Platform.

What makes the Platform unique?



It is powerful. The Platform offers deep language understanding driven by our natural language expertise along with a combination of symbolic and machine learning approaches.

It is comprehensive. The Platform was built to address all aspects of natural language design, deployment and operation.

It is simple. The Platform provides embedded expertise and explainable outcomes that are intuitive and easy to operate.

The Platform has everything you need to create powerful and purposeful language models. Experience it for yourself and realize the possibilities enabled by a hybrid approach.

Get Started

Interested in learning how NLU AI will transform your insurance company? Get started here.

See what expert.ai can do for you!

expert.ai

About us

Expert.ai is the premier artificial intelligence platform for language understanding that augments business operations, scales data science capabilities, simplifies AI adoption and provides the insight required to improve decision making throughout organizations. The expert.ai brand is owned by Expert System (EXSY:MIL), that has cemented itself at the forefront of AI-based natural language solutions across insurance, banking, publishing, defense and intelligence, life science and pharma, oil, gas and energy, and more.

www.expert.ai **info@expert.ai**

©expert.ai 2021