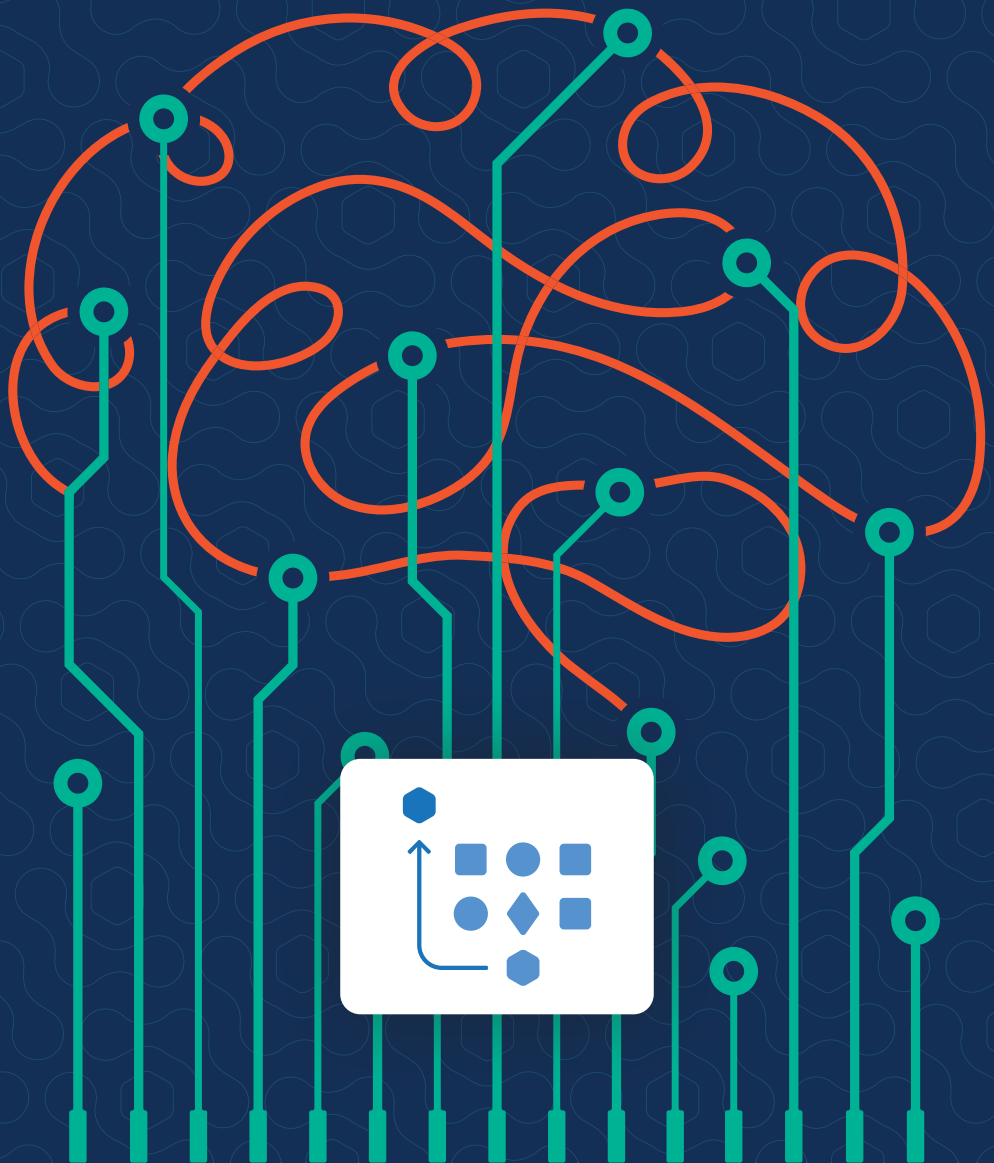
expert.ai

Core Hybrid Use Cases for the Enterprise Business



The enterprise has a language data problem. Not only are organizations sitting on troves of unstructured (often language-based) data, but they are generating more of it every day at record-setting rates. What is worse is one-third of enterprises acknowledge that more 50% of their data is unused, according to Komprise. Ready or not, it is time to put this data to good use.

Recognizing the problem is one thing, but addressing it is another entirely. While many enterprise organizations are turning to natural language processing (NLP) solutions to process and analyze their data, there remains fragmentation in terms of the artificial intelligence approach used to do so.

Traditionally, there are two primary schools of thought when it comes to AI and natural language. You have a rule-based approach in symbolic AI or a statistics-based approach in machine learning. While each offers its own unique set of benefits, neither offers a universal solution fit for every business case. Together, though, you can maximize your natural language capabilities in a hybrid approach.

How to Leverage a Hybrid Approach

A hybrid approach offers you the best of both machine learning and symbolic methodologies. Where one approach is weak, the other can make up for it. This is evident across several core natural language use cases.



Categorization of Highly "Complex" Documents

Pure machine learning has proven extremely effective when it comes to high-volume, low-complexity scenarios (e.g., extracting information from invoices). However, when it comes to processing complex documents, data scientists have struggled to achieve the same results. One of the most common issues they encounter is the same textual passage meaning two different things in two separate parts of a document.

Consider the following scenario: the monetary amount of $50,000 is identified twice within an insurance policy. If that number refers to a deductible or premium, it could imply a reduction in risk for the insurer. If it refers to coverage, it could increase the risk of the insurer.

In both cases, a hybrid workflow could significantly improve your understanding of each document and, thus, your ability to categorize it accurately. In this workflow, a symbolic approach assigns specific roles and characteristics to each document segment (e.g., coverage vs. deductible). This creates a rule-based structure that your machine learning model can leverage to process large volumes of documents.

Extraction of Complex Information to Identify Events from Financial (and Other) News



Pure machine learning is sufficient when it comes to extracting key words or phrases from text. However, external knowledge is critical to extracting more complex information. This is particularly applicable to those who depend on news coverage to carry out their work.

Consider an occurrence like the change in control of an acquired company. Machine learning can detect the more evident changes in a shareholder majority based solely off training data. These changes could refer to a "full merge" where 100% of company control changes hands, or a "small participation" where a small percentage of equity is involved and there is no change in control.

However, training data alone makes it extremely difficult for a model to detect more subtle changes in company control. These scenarios are rarer (not enough for a training set) as they include a combination of transactions or only involve select participating companies. In these cases, symbolic rules can be injected by a knowledge engineer or subject matter expert to address these specific cases and help the machine learning model increase its overall accuracy.

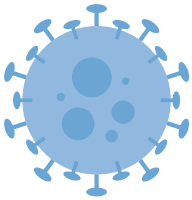
Extraction of "Relational Information"



The effectiveness of machine learning-based information extraction is dependent on a couple factors. First, you must have a sufficient volume of annotated data to train your model. Second, you should focus your extraction on independent data points, meaning those not influenced by other occurrences.

If you account for both factors, your model can produce highly accurate results. However, if your extracted information relies on a sequence of related data points (e.g., company X is buying asset Y from company Z"), your model will often struggle to contextualize the text and extract the correct information.

A hybrid approach that joins machine learning predictions with symbolic rules is a more effective approach to extracting relational information. By establishing rules specific to specific sequences of data, you add valuable context to your model that enables it to identify and extract more complex information that your business processes depend on.



Controlling Data Shift Over Time

Purely data-driven machine learning is not robust to changes in data through time. That means that any unfamiliar events, terms or concepts (or variations to existing terms) introduced to a model cause the accuracy of outputs to suffer.

For instance, **we saw what happened to chatbots when the COVID-19 pandemic began**. To resolve the issue, you would need to inject large volumes of newly annotated data into the original training set and then retrain the system from scratch. Between data collection and annotation, you add a significant amount of time and effort into the equation.

On the other hand, humans can more easily comprehend such an event and identify its context within the model. So rather than retrain your model with new data, you can simply make a rule change to it. This quick adjustment can save you valuable time when it comes to maintaining and updating your model.



Anti-Money Laundering Automated Support in Banking

International law requires banks to **implement an anti-money laundering process over money transfers with a "four eyes" approach**. A first-level analysis detects possible noncompliant transfers while a second-level analysis investigates and eventually flags the transfers as noncompliant.

First-level filtering can be automated, but the information that the first-level agent relays to the second level requires enrichment with an "explanation." This ensures the second level knows which information was detected by the agent (or which inference the agent applied) to determine the noncompliance of the transfer.

This is a very strong example of where "explainability" is required by law rather than an abstract goal. **A hybrid approach to anti-money laundering processes ensures explainability and interpretability of your model.** This enables you to adopt automatic learning processes for non-machine learning-friendly tasks.

All language-intensive processes deserve a hybrid natural language solution. Every enterprise organization deserves a hybrid solution as simple and powerful as the expert.ai Platform.

The Platform makes it easier, faster and less expensive to design, deploy and operate impactful natural language solutions. Purpose-built with end-to-end tools that turn your language into usable and actionable data, you have everything you need to power the processes that power your business.

Now is the time to act on your language data. Make it count with a hybrid natural language approach.

Explore the Platform
OR Set Up Your Platform Demo Today

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