GOSECURE BRINGS AI CAPABILITIES TO WORLD'S LARGEST DATABASE OF LIBRARY COLLECTIONS

GoSecure

CASE STUDY

BACKGROUND

OCLC is a nonprofit global library cooperative that provides shared technology services, original research and community programs so that libraries can better fuel learning, research, and innovation. Through OCLC, member libraries cooperatively produce and maintain WorldCat™, the most comprehensive global network of data about library collections and services. Libraries gain efficiencies through OCLC's WorldShare™, a complete set of library management applications and services built on an open, cloud-based platform.

The quality of the WorldCat™ database is vital to a variety of OCLC services, so maintaining a robust and accurate catalog, free of duplicates, is key. All new records added to WorldCat™ are subject to a quality check to validate against cataloging standards and a matching process to identify if the bibliographic record already exists. OCLC wanted to enhance their processes to improve upon the accuracy and flexibility of this key step using AI and machine learning (ML).

SOLUTION

OCLC partnered with GoSecure to create an initial model focused on audiovisual materials that could probabilistically score similarity between new records and existing records in WorldCat™. Further, OCLC required a model that was easily incorporated into their matching software in order to rapidly deliver value to the community, with a development process flexible enough to expand to multiple material types in the future.

This was a complicated business process and required strong collaboration and trust between OCLC and GoSecure. Weekly working sessions were key in fostering a shared understanding of the entire business process function and aided the ML skill uplift the OCLC team desired.

GoSecure created a custom solution and associated test environment to provide OCLC with a complete baseline of their current system. This enabled continual monitoring of improvement during modeling and the ability to directly quantify improvement.

GoSecure found that a custom-built sequence-to-sequence (seq2seq) probabilistic model yielded optimal performance. This solution is similar to algorithms being produced by research labs at Google, Facebook and Hugging Face and, additionally, is more accessible and custom-tailored to OCLC's immediate needs. The model takes new incoming records and scores their similarity to existing records. The scored records are then passed back into OCLC's existing process.

"At OCLC, we have been delivering high-quality cataloging services to the global library community for decades, continuously investing in our technology platforms. Our mission is to constantly innovate our ability to match, discover, and enrich the nearly half billion bibliographic records in WorldCat™. We partnered with GoSecure to develop an Artificial Intelligence (AI) solution for this critical service."

—Ryan Shondell,
Executive Director
Data Services

BENEFITS

1 **Custom, AI model reduced errors in foundational business process**

The new methodology addressed many scenarios in which new records returned "no match" even though a match existed in the database (false negatives), increasing the match rate of the AV materials 5–10%. This will help reduce duplicate records and staff time spent investigating and merging duplicates into a single record, which will increase library satisfaction and provide a cleaner database.

2 **Scalable AI solution easily integrated into OCLC's current business process**

GoSecure made it an objective to deliver this model in a form that would be easily consumed and integrated directly into OCLC's internal environment. The solution was containerized and packaged into a stable, reproducible environment and made available as a simple API call. The outputs were in a format that could directly be reintroduced into OCLC's existing process. GoSecure optimized the code to ensure API calls were fast for scalability; as evidence, a single call took less than 200 milliseconds.

3 **Maintenance reduced and flexibility increased over rule-based algorithms**

Currently, OCLC spends a great deal of effort maintaining and developing business rules that improve their match quality. These rules get very specific and require specialized expertise in library cataloging. The new AI solution affords lower maintenance scalability and improved performance.

4 **Software engineering team upskilled on machine learning process**

GoSecure collaborated very closely with OCLC's software engineering team. OCLC's expertise allowed GoSecure to build a test environment with custom API calls. GoSecure shared how to test performance on their system, as well as details of the model and code with OCLC software engineering. This provided their team the knowledge to make future changes to the model and API web service. Collectively, these efforts resulted in an initial understanding of how to leverage AI solutions to advance future objectives.

> *"As an added benefit, GoSecure helped to expand the capabilities of our software engineers on machine learning. The GoSecure AI team was able to deliver a fast, scalable, and custom natural language processing (NLP) model that will raise our always increasing quality standards and create a better library experience."*
>
> —Ryan Shondell,
> Executive Director
> Data Services

☐ GOSECURE

LEARN MORE
www.gosecure.net

CONTACT US
www.gosecure.net/sales-contact

1-855-893-5428