

From minutes to seconds: Transforming document management with AI

Digital Office Company and IBM



Overcoming manual document management challenges

Digital Office Company (DOC), a Finland-based provider of information management solutions and services, helps businesses find and manage large swaths of documents at speed and scale by enabling them to identify, classify and extract insights from their documents.

Most often, automated document classification systems provide superficial insight into the nature of documents through generic tags or metadata enrichments. This superficial classification doesn't fully take into account the contents of the document itself, meaning that the true value of the data cannot be leveraged unless each document is manually evaluated.

The bottom line is that such a process of manual evaluation of documents to identify document types, detect misplaced documents and extract crucial details would be very laborious and time consuming. Moreover, lack of proper metadata handling poses potential GDPR concerns and negatively affects data quality for down-stream tasks. Unresolved, such challenges can threaten the competitive edge and efficiency of DOC's customers because of these data quality and regulatory compliance issues.

2.4 seconds

Time to classify a document reduced from a few minutes to 2.4 seconds

Applying the best of both worlds

Hello! How can we help you?

To tackle these hurdles, DOC collaborated with the IBM Ecosystem Engineering Build Lab, IBM Client Engineering and IBM Technology Expert Labs to develop a pilot that

leverages a combination of traditional machine learning and generative AI—large language models (LLMs)—with [IBM® watsonx.ai™](#) and [IBM Watson® Discovery](#).

Through a 6-week co-creation pilot, DOC developed an IBM Watson Discovery platform-powered data pipeline solution that uses custom machine learning models alongside Mistral AI's Mixtral-8x7B LLM to classify documents with custom labels and metadata tags. The focus of the pilot was set on the real estate industry because of the variety of data types and regulatory requirements faced by DOC's customers in this domain. This industry focus led to the creation of an additional capability that expanded the scope of the solution—using LLMs to extract rich insights such as board decisions from meeting minutes documents.

Enhanced document classification, insights and cost savings

The pilot proved to be highly successful in terms of the increases in speed and quality of document classification and insights of large volumes of documents. In addition, the results of the pilot provide a clear indication of the benefits of a combined approach of traditional machine learning and generative AI, in which the shortcomings of each approach are covered by the strengths of the other. What used to take an individual a few minutes to manually look through a document and classify it can take just 2 seconds per document through an automated process that requires a human evaluation only for documents that are flagged as outliers by the system.

Looking ahead, DOC aims to expand on this pilot by further developing its solution to expand it to other industries and customer segments.



About Digital Office Company

Hello! How can we help you?

[Digital Office Company](#) (DOC) (link resides outside of ibm.com) is a Finland-company founded in 1996 that offers modern information management solutions for organizations

of different sizes. Offices are located in Espoo, Hämeenlinna, Lahti and Lappeenranta, Finland.

Solution component

[IBM® watsonx.ai™](#)



[IBM Watson® Discovery](#)



Watsonx.ai is helping companies build custom AI solutions to suit their specific needs.

[Explore watsonx.ai](#)



[Start your free trial](#)



Legal

© Copyright IBM Corporation 2024. IBM, the IBM logo, IBM Watson, and watsonx.ai are trademarks or registered trademarks of IBM Corp., in the U.S. and/or other countries. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on <https://www.ibm.com/legal/copytrade>.

This document is current as of the initial date of publication and may be changed by IBM at any time.

Examples presented are illustrative only. Actual results will vary based on client configurations and cannot be provided. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

Hello! How can we help you?

results

ARRANTIES

The client is responsible for ensuring compliance with all applicable laws and regulations. IBM does not provide legal advice nor represent or warrant that its services or products will ensure that the client is compliant with any law or regulation.

Hello! How can we help you?