



NTT DATA automates document classification and information extraction for Iberdrola, with Dolffia, an AI-based processing platform



With a history of over 170 years, today Iberdrola is a Spanish global energy leader, the number one producer of wind power, and one of the world's biggest electricity companies in terms of market capitalization. The group supplies energy to almost 100 million people in dozens of countries, has more than 600,000 shareholders, a workforce comprising almost 39,000 employees and assets worth more than €123 billion.

The challenge

As an energy corporation that's performing at a global level and services millions of clients, Iberdrola processes countless documents and contracts specific to subcontractors, health and safety check-ups, energy controls, b2c service contracts and more.

From subcontractors alone, the client is processing and validating more than 300.000 documents per year from 49 different types, which in the past were done manually. For these elaborate tasks, they had developed an internal

system called Siroco, where subcontractors had to submit the necessary documents. The documents were then processed by a team of people who checked the documents one by one, verified whether the information met the required internal standards and approved or rejected the request.

Given the significant time and resources needed to process these documents, Iberdrola decided to implement a solution with advanced capabilities that would automate this process, but also help boost efficiency and lower costs.

In order to find the best solution that satisfied all their requirements, Iberdrola organized a tender and invited NTT DATA, among other providers, to develop a proof of concept to automate the extraction of information from IDs and other documents. NTT DATA used Dolffia, an AI-based document processing platform, for this particular project and proved its advanced capabilities by having the best results. As a result, Iberdrola chose NTT DATA as a technological partner and Dolffia as the best solution for their needs.

The client's goal was to achieve the following objectives:



Automate the classification and extraction of certain information from documents using a solution with advanced technologies such as AI, machine learning or natural language processing (NLP) .



Increase efficiency, eliminate the risk of errors in information extraction, and lower costs by leveraging the advanced capabilities of an automated solution.

The solution

As a future-thinking organization that views innovation as a way to guarantee sustainability, efficiency and competitiveness in the long term, Iberdrola understood the benefits of a solution like Dolffia. Among its many benefits, it is fast and scalable, highly adaptable due to its capacity to extract information from a wide range of structured, unstructured and semi-structured documents, it classifies documents with an accuracy of 95%, it ensures compliance and offers rapid set-up time reducing indirect IT overheads to almost zero when installed in the public cloud. Using its machine learning and Natural Language Processing capabilities, Dolffia is able to learn on its own, refine itself and suggest actions that unblock information from any kind of unstructured data source.

Also, implementation required only a few easy steps:

- 1** Gather all the documents and classify them depending on their type so that Dolffia can recognize them automatically and file them accordingly.
- 2** Define and setup the necessary flows.
- 3** Activate Dolffia and start the process of auto-learning using AI and machine learning so it fine tunes itself.

Once the process of automatic learning starts, Dolffia starts recognizing a document type, sends it through the corresponding flow and extracts the necessary information. In addition to this, Dolffia also validates the information and sends confirmation that it's correct.

Given the positive results Dolffia has had over the first years since the collaboration with Iberdrola began, within the group, 2 more divisions were acquired: Sic Contratación and Customer Web.

In the case of Sic Contratación, Dolffia was brought in to automate the processing of electrical installation certificates created by the technicians that record the changes in a client contract.

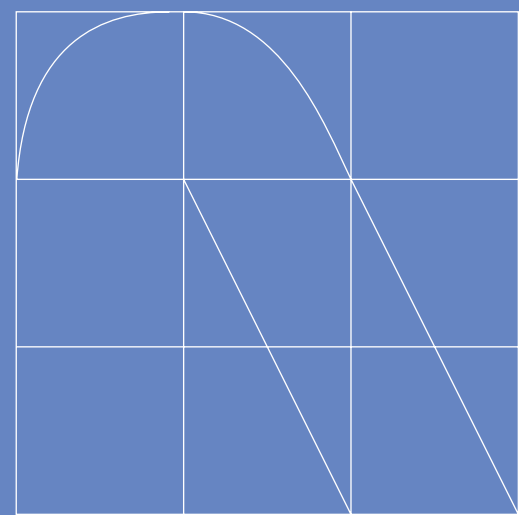
For Customer Web, Dolffia had a more complex objective: validate customer identification documents in real-time. Compared to the previous situations where the requests were sent, validated and the results were passed to the client, for this particular situation, the team had to adapt Dolffia so that it delivers a volume of requests that exceeds that of Siroco, and in real-time.



Fast and scalable

Highly adaptable due to its capacity to extract information from wide range of structured and unstructured documents.

It classifies documents with an **accuracy of 95%**.



Required capacities:

- Real-time document validation
- Identify and anonymize certain fields of information



Results

- Over 500.000 documents have been processed
- Accuracy increased to 95% for document classification and to 85% for information extraction
- Drastic reduction in document rejection rates
- Improved response rates, reducing validation time to less than 3 seconds per document

Reference to offering and capabilities

Dolffia was created with one clear goal: help businesses improve efficiency and lower costs by automatically processing the growing quantity of unstructured documents they deal with every day. By leveraging its powerful advanced capabilities, Dolffia boosts throughput, reduces the risk of errors and allows employees to focus on more demanding tasks.

Dolffia uses NTT DATA's extensive MLOps expertise in AWS and it has a microservice architecture deployed in AWS.

With Dolffia,
Iberdrola has
achieved:

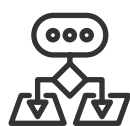
Over 500.000
documents
processed

Accuracy
increased:

to 95% for
document
classification

to 85% for
information
extraction

Among its main features are:



Document classification using the text, machine vision or a combination of both methods. Parallelization techniques are used to reduce the time required to train the ML algorithms on a document collection.



Information extraction by tokenizing and segmenting the document. It then applies natural language processing to correctly identify the text.



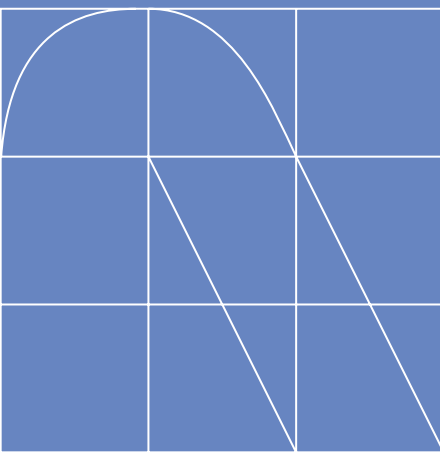
Natural language understanding including knowledge graphs, language models and semantic OCR for more advanced use cases. For such situations, it uses these techniques to improve the understanding of the documents and extract the knowledge of the content.

Customer quote

The implementation of Dolffia has led to an increase in the processing of documents automatically by 90%, a reduction in registration times for new online users and decrease of delays, which has also helped to reduce the claims.

On the technical side, we have learned to leverage Dolffia's AI and machine learning capabilities and teach the system, configure it properly and make the most of this advanced solution for different use cases.

Arantxa Ríos, Head of Document management Systems and Engineering , Iberdrola





NTT DATA QUOTE

“Thanks to Dolffia we have helped our client Iberdrola to automate the End to End of several business processes. We are currently obtaining a success rate in document validation and extraction above 90%. We started processing 3,000 documents per month in the first year and in the second we already processed more than 20,000 documents per month”.

Jacinto Estrecha Cádiz. Head of Artificial Intelligence at NTT DATA

What's next

As one of the top IT and consulting companies in the world, NTT DATA is committed to investing in technology and continuously improving solutions such as Dolffia, which was designed and developed at the NTT DATA Center of Excellence for AI.

A potential next step in this project is implementing the upcoming advanced features such as intelligent or semantic search. Our engineers are also working on identifying new and more complex applications for Dolffia, aimed at optimizing the processes of information search and extraction from documentation that goes online from intelligent search.

For more information

www.es.nttdata.com

