

WHOLE OF GOVERNMENT DOCUMENT CLASSIFICATION CASE STUDY

Letters and email communication are routine daily business activities of all UK Government agencies. Employees spend one third of their time managing emails. As email is one of the most used customer touchpoints more than half of an organisation's knowledge was stored in inboxes.

The UK Government sought a solution that could:

- Process 100,000 letters per day from about 400 financial institutions in an automatic way.
- Perform automatic classification on content according to 12 different categories.
- Automatically extract 10 different datapoints.
- Enable forwarding to letters to responsible person via a digital workflow management system.
- Monitor the responses back to the original author.
- Be flexible enough to deal with previously unseen letters.
- Generate consistent and explainable outcomes.

Pure Machine Learning technologies had already failed to perform this task

Due to the very sensitive nature of the project, the customer chose a global partner of Expert System to develop the linguistic package to categorise the communications in 12 different categories and the extraction of all the relevant entities (with 10 relevant entities for each category) from the content.

The solution was implemented on AWS. The letters are scanned on receipt using ABBYY FineReader, an intelligent OCR solution. The text is then passed to Cogito Discover which extracts the information required.

Cogito Discover enables semantic tagging, automatic classification, and clustering of documents. standard taxonomy and metadata which includes main concepts, people, organisations, geographic locations, dates, and measures. customised taxonomy and metadata enabled.

The data generated from the text is stored in SQL Server. The letters are then digitally processed via Team Foundation Server 2018.

The customer claimed the following benefits from the use of Cogito:

- 3M+ letter processed
- 85%+ completely automated
- 97%+ accuracy
- 200K docs per day capacity
- £5M projected net benefit.