

INTENDA

START-UP
CASE STUDY

INNOVATIVE
REAL-WORLD
SOLUTIONS / FOR
TODAY / FOR THE
FUTURE

FRAXSES



Firstly, we're eliminating the client's cumbersome copy-and-paste approach to lifting text from PDFs.

THE CLIENT

Our client compiles data that is sourced from the public domain and published by government municipalities for public consumption.

The data is lifted from public PDFs, stored in a central database, and served to users in the form of dashboards.

The client needed an entirely different approach to data processing and management, and a solution that would provide best in class processes for PDF processing, data storage and business intelligence.



THE SOLUTION

Our solution for this client entailed two parallel work streams, the first being the conceptualising of the future state end-to-end process, and the second being the configuration of the corresponding technical components this would require, using the Fraxes platform.

We deployed Fraxes on-premises to host and virtualize the final data set, so that content is delivered via the Fraxes visualisation tool.

The client's original time estimate for completion of their project was two years. We delivered the entire solution in under four months, effectively streamlining two of their major production bottlenecks by implementing two microservices:

- Firstly, we're eliminating the client's cumbersome copy-and-paste approach to lifting text from PDFs. Originally, workers would sift through electronic documents and transcribe data, page by page.

Now, business analysts begin by building a metadata library for each PDF.

They identify which pages contain relevant data, record them, and then capture the coordinates of the data in a zonal capture programme. An OCR service can then digest those coordinates and auto-matically lift the text data from the PDFs.

Fraxes can then re-assemble useful data and filter out the noise from the result sets provided.


The client's original time estimate for completion of their project was two years. We delivered the entire solution in under four months.

(CONTINUED)

- Secondly, we have developed a classification microservice that automates enrichment and classification - this would otherwise be a manual process. This enrichment service is built using human-defined keyword matching.

This process adds over 100 unique dimensions to the dataset. These dimensions are later used to customise the user experience in the Fraxses visualisation tool. This solution is offered as a Software as a Service (SAAS) model by Intenda. The input to the data platform is an Excel database that a business analyst has reviewed for accuracy.

The Integrate module runs a batch process to dimension the data into its components while virtual Data Objects rebuild the original structure to create a hosting friendly data model.



This process adds over 100 unique dimensions to the dataset.

BENEFITS

While the current end-to-end processes are loosely coupled, the future state of these processes entails tight integration with document management software.

Enrichment and classification will move to machine learning.

Essentially, business analysts will be able to run these procedures from a data workbench in an automated fashion, which would not be possible without the metadata catalogue approach and automation we have provided for this client.



Essentially, business analysts will be able to run these procedures from a data workbench in an automated fashion.

THANK YOU
GET IN TOUCH

____WE
SEE
DATA
INSIDE
OUT____

www.intenda.tech