

# Automate Application Processing with Machine Learning and AI



## Introduction

A large part of the work of UK based Universities is to offer courses to foreign students, and to manage their application submissions (Application Processing).

Each submission is received on one of many different mediums, including paper and PDF, requiring conversion from unstructured to structured data to allow automated processing.

This Blueprint shows how machine learning and AI, combined with business rules can deliver significant improvements to quality of service and accuracy of application processing, while reducing cost.

IBM Business Automation Content Analyzer allows us to ingest and process thousands of higher education applications quickly and consistently. Leaving only the exceptions for human attention.

With larger Universities running with upwards of 40,000 students it is easy to imagine that processing 100,000 graduate applications each year.

Each graduate application is accompanied by an "academic transcript" that must be examined to assess whether the qualifications presented are appropriate for the position.

Transcript layouts and emphasis, content, and quality of English vary by geography and even by University in some countries. Many documents supplied are "scans of photocopies", which compromises quality and readability.

## Reduce Cost and Improve Quality

An Application Processing solution was needed that allowed interpretation and pre-processing of applications to be done quickly and efficiently. The problem can be broken into to two parts.

- (1) Extracting the data into a usable format, and
- (2) Making a decision by comparing the applicant data with the position requirements.

These are distinct problems that require very different functionality.

## Currently Application Processing is All Manual

The current approach for many Universities for processing these applications is entirely manual, very labour intensive and expensive. The staff undertaking the work are highly skilled and often have to work long hours at peak periods to deliver results, particularly for the start of term or academic year.

Further complications include an expected 20% annual increase in the volume of applications and that most applications arrive within a two-month window.

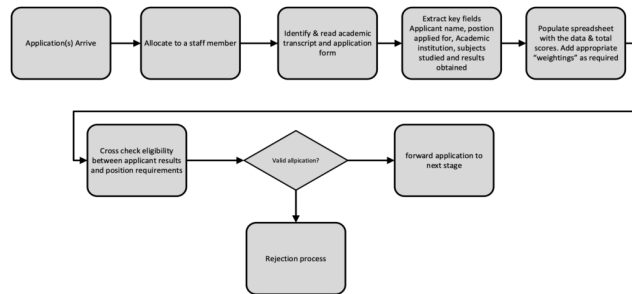


Figure 1; Example process

## Solving the Application Processing Problem

### Ingest, interpret, and first triage

A product is needed to allow highly variable documents from various sources and quality to be read, interpreted, and converted to a consistent format for onward processing.

We use IBM Business Automation Content Analyzer to ingest documents and make sense of them.

IBM Business Automation Content Analyzer is configured using a web based graphical interface. No coding is required.

This is a SaaS offering that uses Optical Character Recognition (OCR) and IBM Watson artificial intelligence (AI) to examine documents and extract the required data.

The solution generates a JSON name/value structure containing fields extracted from the original document. Each field has an associated "certainty" attribute indicating how accurate the value returned is expected to be.

The Application Processing solution is driven by a template that defines a "class of document", which in this case is an "academic transcript", and an ontology based on the names of the fields that are required from the document.

Aliases are allocated to the name/value pairs so "Surname" and "last name" both map to the same variable. Fields can also be "marked as "required" to flag errors.

IBM Business Automation Content Analyzer is configured using a web based graphical interface that does not require code to be written. The product utilises machine learning based on documents uploaded for testing and verification.

### Exception Approval and Acceptance or Rejection

The standardised document is scored against measures defined for the role being considered, and the result triaged. For example, a score of less than 70% is automatically rejected, under 10% is reviewed for possible failure of the system, over 70% goes forward for processing

IBM Operational Decision Manager (ODM) can be purchased as a cloud service or self-hosted solution. It is also a key part of the Responsiv Unity platform.

IBM Operational Decision Manager (ODM) allows non-technical users to specify rules in a natural language format rather than a coding language.

Rules can be as simple as "if score is less than 75% then reject" to a complexity where hundreds of rules are triggered to make the "rejection" decision.

Initial development requires a small amount of coding to set up rule parameters that then allow a business analyst to define rules that make the scoring decision. In this case the rules compare skills and scores the applicant has claimed, with the requirements of the position.

## Overall Application Processing Solution (how the magic works)

The whole Application Processing process is an "event driven, real-time architecture" that is triggered by the arrival of an application. ODM is used to manage the process, which calls to IBM Content Analyser to parse each document and extract data.

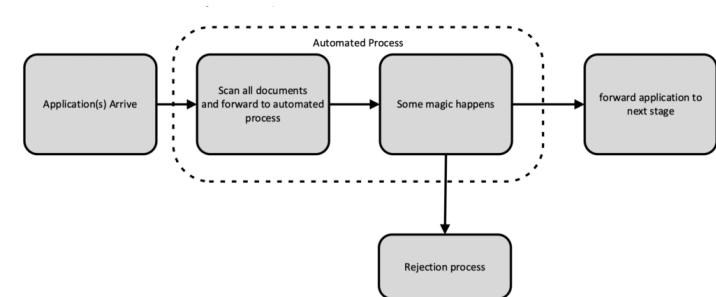


Figure 2; Example automated process workflow

The machine learning element of this solution means that the quality of its work is expected to improve over time, and the number of exceptions that require human intervention to reduce.

Further learning: submitting rejected documents to enhance the machine learning of the OCR system, and by altering the rules to better filter applicants based on feedback from the interviewers can also reduce the number of human interactions required.

### Future Enhancements

IBM Business Automation and Workflow (BAW), and other IBM and Responsiv technologies can be added to the solution to extend the automation, and improve outcomes.

## Conclusion and Next Steps

Adoption of IBM Content Analyser and IBM Operational Decision Manager starts with a conversation with Responsiv to flesh out ideas, consider the options, and decide whether it is right for you at this particular time.

We know that higher education establishments have a great deal going on, and this type of solution can be tactical or strategic, and can move from being tactical to strategic. Either way the initial implementation needs to deliver value and demonstrate the opportunity – otherwise what's the point?

If you decide to give it a go, then we will start with a discovery phase. This involves reviewing the processes that may be candidates for automation, and identifying two or three that can be delivered quickly, and that will demonstrate value.

Following discovery, a Responsiv subject matter expert writes a set of instructions that implement a simple process that can be continually refined.

Any exceptions or escalations are reviewed and resolved by an experienced business user to be added to the automation or treated as exceptions.