

# Adatis

## Knowledge Mining using Azure Cognitive Search and Databricks

[www.adatis.co.uk](http://www.adatis.co.uk)



Microsoft  
Partner



Gold Cloud Platform  
Gold Data Analytics  
Gold Data Platform  
Gold DevOps  
Gold Datacenter



# Adatis



Adatis, a leading Microsoft- specialist Data Analytics and Business Intelligence consultancy, work in flexible, agile and collaborative ways to help organisations unlock the power and value of their data to enable them to improve business performance, better serve and engage customers, empower employees, find efficiencies and remain compliant.



## WHAT WE OFFER

### BUSINESS CONSULTING

- Data Strategy & Advisory
- Data Governance
- Master Data Management
- PowerBI Review
- Power BI Governance & Rollout
- PowerBI Training
- Data Envisioning
- AI Proof of Concept

### DELIVERY

- Enterprise Intelligence Solutions
- Data Architecture
- Data Platform
- Modern Data Warehouse
- Data Science Platform
- Data Migration
- Application Integration
- Advanced Analytics
- Azure Synapse
- Azure Databricks
- Artificial Intelligence & IoT
- PowerBI Migration

### MANAGED SERVICES

- Experts in Data Analytics and AI
- Reactive Support
- Infrastructure & Application Monitoring
- Proactive monitoring and optimization
- Continues improvement
- Evolves with you to maximise value
- Business Continuity



Gold Cloud Platform  
Gold Data Analytics  
Gold Data Platform  
Gold DevOps  
Gold Datacenter

# Adatis Knowledge Mining Solution, using Azure Machine Learning technologies

*With Adatis's Knowledge Mining solution, ingest data from many different sources into a single unified data estate, and unlock insights across structured and unstructured content – such as documents, images, and media - by applying AI at scale across your information. With AI-driven content understanding, discover hidden patterns and relationships in your content, understand sentiment, extract key phrases, and more.*



The Adatis Knowledge Mining solution has functionality that covers a wide set of requirements including:

- **Data Collection:** Obtaining documents and storing them in a location that is accessible to the tool
- **Pre-processing:** Cleaning and transforming data so that it can be trained and scored as part of the model development or classification pipeline
- **Modelling:** Training and evaluating different classifiers to determine the best algorithm and best set of hyper parameters to give the greatest performance
- **Deployment & user interface:** Deploying the trained model into a running environment where unseen documents can be scored and presented back to users through an easy to use and aesthetically pleasing interface.

Technologies used include:

## Azure Cognitive Search

Azure Cognitive Search is the engine that underpins the processing and indexation capabilities of the solution that uses AI to enrich documents with metadata. Enriched documents are stored in an intelligent index that provides a variety of search options

## Azure Cognitive Services

A suite of pre-built, pre-trained models exposed through simple REST API's for performing common AI tasks

## Azure Databricks

Azure Databricks is the processing engine that can evaluate large and/ or complex data sets and train AI models at scale, and can be turned off when not needed to reduce costs.

# Customer success:

National Archives use Microsoft Cognitive Services and Databricks to identify and select digital assets to be stored



## Customer Challenge

Legislation states that after 20 years government departments need to determine if an asset should be stored or not. With vast volumes of ‘born digital records,’ the National Archives could no longer manually identify assets to be stored.

They wanted to advise other government departments on how technology could assist them in identifying and selecting relevant assets more quickly and ensure assets were not overlooked.

## Adatis Approach

The National Archives decided to work with Adatis in order to conduct the research and build a proof of concept into machine learning techniques, to classify digital assets that should be selected for preservation.

Adatis used Microsoft Cognitive services to search the documents for key phrases and entities and used Databricks to train the machine learning models.

## Customer Outcome

The research paper and proof of concept developed using Azure machine learning technologies will enable the government department to faster identify and select digital assets and documents that need to be stored.

It will also ensure that critical assets are not missed that need to be stored in line with legislation and reduce the time, overhead costs and manual effort required to identify these assets

# Find out more about the Adatis Knowledge Mining Solution

- Call for more information: +44 (0)1252 267 777
- Ask a question via email: [enquiries@adatis.co.uk](mailto:enquiries@adatis.co.uk)
- Learn more: <https://adatis.co.uk/>
- Check out our blogs: <https://adatis.co.uk/blog/>
- Check out our blog on [How to get started with Azure Cognitive Search](#)



Microsoft  
Partner



Gold Cloud Platform  
Gold Data Analytics  
Gold Data Platform  
Gold DevOps  
Gold Datacenter