

Pro Insurance Solutions

Cleansing Artificial Intelligence (CAI) System

OVERVIEW

The insurance industry in general is cautious about the adoption of new technology, and the ingestion of risk data quite often begins on hard-copy paper or in an Excel spreadsheet. This data may be entered in practically any format whatsoever before being delivered, for example, to an insurance broker. The broker may then add additional information such as the insurance coverage and terms, before pushing the data out to underwriters in the market to quote. This process, and the accompanying required data transformations, is a very manual, repetitive, and tedious one, especially due to the often liberal use of unstructured data and file types.

Pro Insurance Solutions Limited ("Pro") provides an outsourced service where they receive broker data (either directly or via an insurer) and create a template (manually) to represent that risk. Pro would then greatly enhance that data with other available information, providing to customers a generalized template more readily ingestible by 3rd party models to price that risk through computer simulations in order to determine the theoretical cost of coverage. However, these simulators required fixed data formats, while Pro and their clients received the data in random formats.

To solve these process and data issues, Pro, decided to build a new "Cleansing Artificial Intelligence" (CAI) system that would automate these data transformation and data cleaning processes. The goal was to reduce the total end-to-end processing time from several days to several hours or less. This would give Pro's clients a tremendous competitive advantage, and in fact be a major market disruptor. Atmosera, in partnership with Pro and AI firm Analycat, delivered the system in late 2019.

DETAILS

The core vision of CAI was to create a system that continuously monitors Input Events (input of spreadsheet risk data) and then semi-automates, via workflow processes, the tasks of cleansing, transforming, and generating output documents. The high-level conceptual architecture is as follows:

- Front end UI built using modern web technologies such as React and ASP.NET MVC.

- Hosting and services environments built using Microsoft Azure cloud technologies:
 - Azure Active Directory to authenticate and manage application users
 - Azure Web Apps to host the web applications
 - Azure Application Insights to capture application telemetry
- Data for the system hosted and processed using a combination of storage technologies:
 - Azure Blob storage to store worksheets in-progress and log files.
 - Cosmos DB to store information about clients, jobs, worksheet transformations, and other application state data.
 - Azure Key Vault (via use of Managed Service Identities) to store system secrets and access keys.
- Additional services:
 - Cognitive Services for language Translation.
 - Azure Search and SignalR.
- The high-level system functionality includes:
 - Basic Document Flow Management.
 - Email ingest and output components driven by an Azure Logic App, configured to trigger a workflow based on the receipt of incoming emails. The Logic App stores attachments in Workbook Storage, creates records in User/Client/Job Storage, and creates the necessary activity entries in Activity Log Storage.
 - Azure Functions are used to supplement the Logic App workflow where pre-built modules are not available.
 - End-User Application Core Features
 - Application Sign In to the Application Portal
 - User Management
 - Client Account Management
 - Data Import (Triage Interface)
 - Task Management
 - Preliminary Analysis of Documents

- Document Cleansing
- Data Transformation
- Geocoding Interface
- Output Generation
- Data Management
- Schema Management

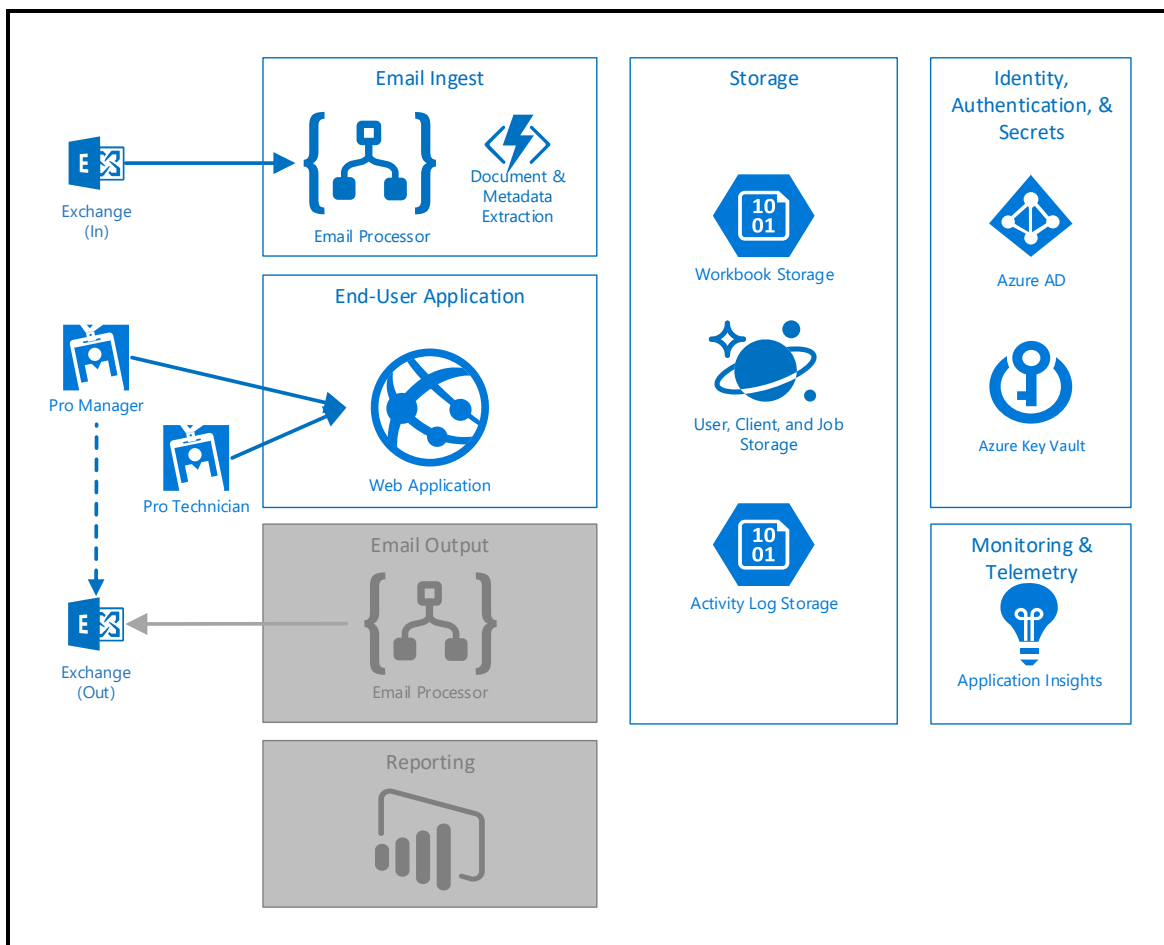


Figure 1 - High Level Architecture

The core tools used consisted of the following:

- **Power BI.** Power BI is a suite of business analytics tools that can be used to analyze data and share insights. Power BI reports can be connected to many

different data sources, produce reports, and publish them for an organization to consume on the web and across mobile devices.

- **Azure Blob Storage.** Azure Blob storage supports storing unstructured data in the cloud and can be used to store any type of text or binary data.
- **Azure Cosmos DB.** Azure Cosmos DB is a globally distributed, multi-model database service that offers storage across several well-defined consistency levels.
- **React.** A framework used for creating browser applications.
- **ASP.NET MVC.** A framework for building web APIs that are used to support the front-end browser application.
- **Also: Azure Search, Cognitive Services for Translation, and SignalR.**

Azure DevOps for Application Lifecycle Management

To ensure the application could be rapidly built, iterated on, and deployed, Azure DevOps was used for application lifecycle management, automating the build and release cycle. The team implemented continuous integration and continuous delivery processes using Azure DevOps Pipelines, ensuring code could be built, tested, and released on demand. Finally, Azure ARM templates were used as the solution for infrastructure-as-code, ensuring the entire environment could be released to Azure in a fully automated manner.

Transformation POC

ABCompanyTestData - Chinese

Worksheets: Europe, UK (selected), Transport

Select: Rows, Columns, Headings (selected)

To Be: Included (selected), Excluded

Town/City	Post Code	Address	Predominant Year Built	Predominant Number of Storeys	Buildings Value	Contents Value	Total
Chelmsford	CM1 1PD	1 Collier Road	circa 1900's	2	270000	7662.831750000012	277662.831
	CM1 1LL	1 Gwydir Street		n/a	0	0	0
	CB1 1PT	11 Collier Road	1973	2	265000	7662.831750000012	272662.831
	CB1 1PT	117 Catherine Street	1973/2012	n/a	0	5108.5545	5108.5545
	CB1 2LZ	18 Eckersley Road		Two floors	0	0	0
	CB1 2LZ	17 Collier Road	circa 1900's	Single storey	265000	7662.831750000012	272662.831
	CM1 1TD	22 Thoday Street	2015	n/a	0	7662.831749999999	7662.83174
	SE18 6PF	23 Collier Road	Partly demolished, not being replaced	2	265000	7662.831750000012	272662.831
	CR0 0YN	24 Great Eastern Street	1890	Single storey	0	2554.27725	2554.27725
	KT2 6NH	25a Newmarket Road	Refurbished 2011/2012	Single storey	0	7662.831749999999	7662.83174
	TW3 3HW	267 Mill Road	2015	Single storey	0	7662.831749999999	7662.83174
	SE1 0LH	29 Abbey Road	New build, completed 2012	n/a	0	6385.693125000001	6385.69312
		37 Collier Road	Refurbished 2006/2007	2	250000	6385.693124999999	256385.693
		37 Tenison Road	Refurbished 2006/2007	n/a	0	15325.663499999999	15325.6634
				Overall Totals	1315000	89399.70375000003	1404399.70

Fig 2 - Data Selection: Selecting data of interest from ingest worksheets

CAI - CLEANSING

Tasks Documents Administration Dave Baskin

SOURCE					TRANSFORM	TARGET				
S1: Worksheet	S2: ADDRESS	S3: CITY	S4: STATE	S5: ZIP CODE		T1: POL ID	T2: LOC ID	T3: LOC NAME	T4: INCEPT DATE	T5: EXPIRE DATE
1 Demo SOV	101 West ? Park...	Birmingham	AL	35211	T1 = "DEMO"	1 DEMO	1	101 West ? Park...	01/01/2019	31/12/2019
2 Demo SOV	1230 Inverness C...	Birmingham	AL	35242	T2 = ROWID	2 DEMO	2	1230 Inverness C...	01/01/2019	31/12/2019
3 Demo SOV	1400 STARTFOR...	Stratford	CT	06615	T3 = S2	3 DEMO	3	1400 STARTFOR...	01/01/2019	31/12/2019
4 Demo SOV	1300 N. W. 1st A...	Boca Raton	FL	33432	T4 = "01/01/2019"	4 DEMO	4	1300 N. W. 1st A...	01/01/2019	31/12/2019
5 Demo SOV	1200 Court St	Clearwater	FL	33756	T5 = "31/12/2019"	5 DEMO	5	1200 Court St	01/01/2019	31/12/2019
6 Demo SOV	6238 CREEL RD	Theodore	AL	36582	X2 = GEO(X1, "resolution")	6 DEMO	6	6238 CREEL RD	01/01/2019	31/12/2019
7 Demo SOV	1410 Zero St	Fort Smith	AR	72901	T6 = GEO(X1, "street address")	7 DEMO	7	1410 Zero St	01/01/2019	31/12/2019
8 Demo SOV	4100 W GALVES...	Chandler	AZ	85226	T7 = GEO(X1, "city")	8 DEMO	8	4100 W GALVES...	01/01/2019	31/12/2019
9 Demo SOV	524 Day Street	San Francisco	CA	94107	T8 = GEO(X1, "county")	9 DEMO	9	524 Day Street	01/01/2019	31/12/2019
10 Demo SOV	520 3rd street	San Francisco	CA	94107	T9 = GEO(X1, "state")	10 DEMO	10	520 3rd street	01/01/2019	31/12/2019
11 Demo SOV	8200 W Colfax A...	Lakewood	CO	80214	T10 = GEO(X1, "post code/zip")	11 DEMO	11	8200 W Colfax A...	01/01/2019	31/12/2019
12 Demo SOV	5105 Johnson R...	Coconut Creek	FL	33073	T11 = GEO(X1, "country")	12 DEMO	12	5105 Johnson R...	01/01/2019	31/12/2019
13 Demo SOV	10777 Phillips H...	Jacksonville	FL	32256	T12 = GEO(X1, "latitude")	13 DEMO	13	10777 Phillips H...	01/01/2019	31/12/2019
14 Demo SOV	922 Northside D...	Statesboro	GA	30458	T13 = GEO(X1, "longitude")	14 DEMO	14	922 Northside D...	01/01/2019	31/12/2019
15 Demo SOV	73-172 Aulepe St	Kailua-Kona	HI	96740	T14 = GEO(X1, "resolution")	15 DEMO	15	73-172 Aulepe St	01/01/2019	31/12/2019
16 Demo SOV	560 Bessie Cole...	Chicago	IL	60666	T15 = LOOKUP("CONSTRUCTIONTYPE (FRAME, BLOCK)", S15)	16 DEMO	16	560 Bessie Cole...	01/01/2019	31/12/2019
17 Demo SOV	1021 Burkemont...	Morganton	NC	28655	T16 = LOOKUP("occupancy", S14)	17 DEMO	17	1021 Burkemont...	01/01/2019	31/12/2019
18 Demo SOV	4161 W Henriett...	Rochester	NY	14623	T17 = S17	18 DEMO	18	4161 W Henriett...	01/01/2019	31/12/2019
19 Demo SOV	6624 GARTH RD	Baytown	TX	77521	T18 = FIRST(S13)	19 DEMO	19	6624 GARTH RD	01/01/2019	31/12/2019
20 Demo SOV	1225 INTERSTAT...	Beaumont	TX	77701	T21 = S16	20 DEMO	20	1225 INTERSTAT...	01/01/2019	31/12/2019
					T22 = "SOFT"					
					T23 = S11					
					T24 = IF(SAME(S10, ""), 0, S10) + S9					
					X3 = IF(SAME(S10, ""), 0, S10) + S9 + S11					
					T25 = PRORATA(X3, Se+7)					
					T27 = 0.01					
					T28 = "USD"					
					T98 = 0					
					+ Add ...					

Fig 3 - Transforming origin data into standard form using transformation formulas

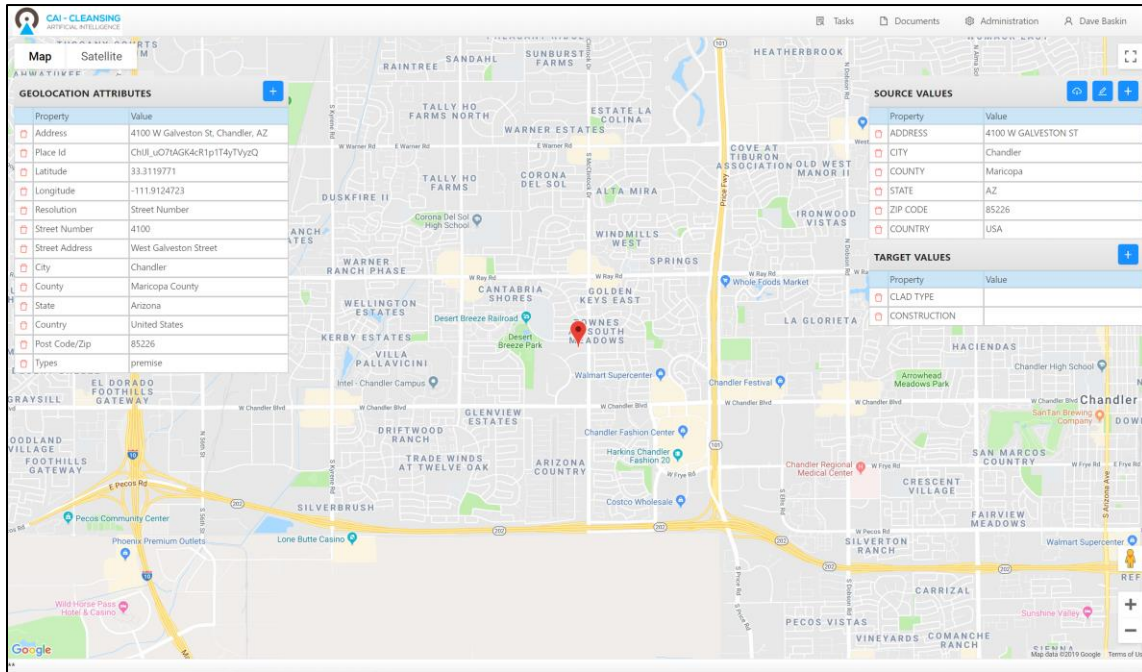


Fig 4 - Capturing geocoding information for an individual property

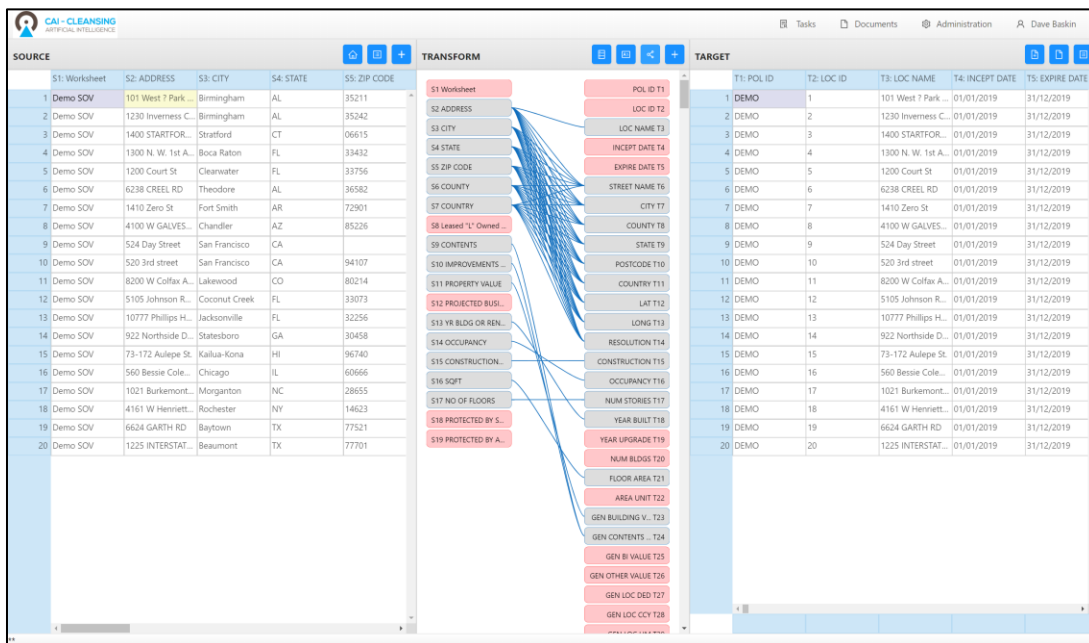


Fig 5 - Mapping relationships between source columns and target columns

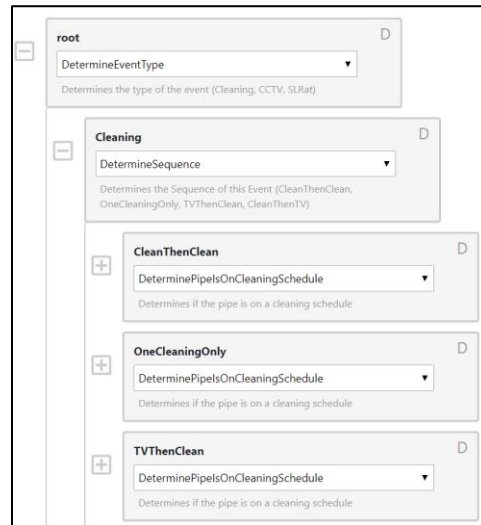


Figure 6 - Decision Tree Builder allows full configuration of processing logic

WHY ATMOSERA?

Pro needed a partner that was expert at providing the architectural skills needed to take the complex business logic required and translate that into a technical design that was flexible and performant, as well as the implementation expertise to develop the production system.

- Atmosera has deep experience in developing intelligent and configurable decision solutions.
- Atmosera is a Microsoft Gold Cloud Platform, Application Development, Data Platform, and Data Analytics partner
- Core focus and skillset in React for Modern Web Development.
- Atmosera is a recognized leader in software architecture and implementation on the web, mobile, and cloud platforms.

About Atmosera

Atmosera was founded by a team of technology experts who are renowned in the industry, with the goal of providing the highest level of custom software development, cloud,

architecture, and training expertise to our clients. Our mission is to help customers “build better software, faster,” and to do this we have structured our company as a dedicated team of technology experts accompanied by the top tier of software engineers. Along the way, Atmosera has developed a reputation as a world class provider of software technology services.

In the training world, we are one of the top trainers on the Microsoft platform in the U.S., including for Microsoft itself where we train their product team developers and testers in Seattle and around the world. On the Consulting side, Atmosera is recognized as a leader in mission critical architecture design and development, and has worked with major companies on large enterprise systems around the country and worldwide. We have even worked with Microsoft to develop some of their products and architectures.

Atmosera is a Microsoft Gold Cloud Platform, Data Platform, AppDev, DevOps, and Data Analytics partner. In 2018, we were awarded the IAMCP World Wide Microsoft Channel Partner of the Year. In 2019 we won the MSUS Partner Award for Intelligent Cloud: DevOps.

About Analycat

Analycat was founded in 2015 as the parent company for catastrophe modeling companies around the world that have used Analycat software under various brands since 2008. Analycat’s Asian distributor, Catalytics has the vast majority of the cat modeling market in Indonesia, Thailand, Vietnam, Singapore, the Philippines and Malaysia using our software. Analycat represents Atmosera in the UK, and works in partnership using their domain knowledge in the insurance industry to streamline deliverables in creating bespoke first-class solutions.

About Pro

Pro is the principal subsidiary of Pro Global Holdings Limited, an international service provider and consulting group focused on the insurance and reinsurance industries. The Pro group maintains offices in London, Cologne, Zurich, New York and Buenos Aires and is supported by regional centers across the globe, and maintains 4 main lines of business: Technical Outsourcing, Legacy Operations, Risk/Compliance, and Operational Consulting. Pro is highly embedded in the insurance industry, with a portfolio of over 160 industry-leading clients.