mint

Create tomorrow

r

Mint Vision Platform

05 August 2019

Created by Peter Reid

Table of Contents

1. Introduction - The Mint Vision Platform

The Mint Vision Platform is an enterprise-scale platform that brings the power of advanced AI computer vision into the organisation. The platform consumes real-time or recorded feeds from almost any source, and performs advanced analytics such as face recognition, object detection, and voice recognition.
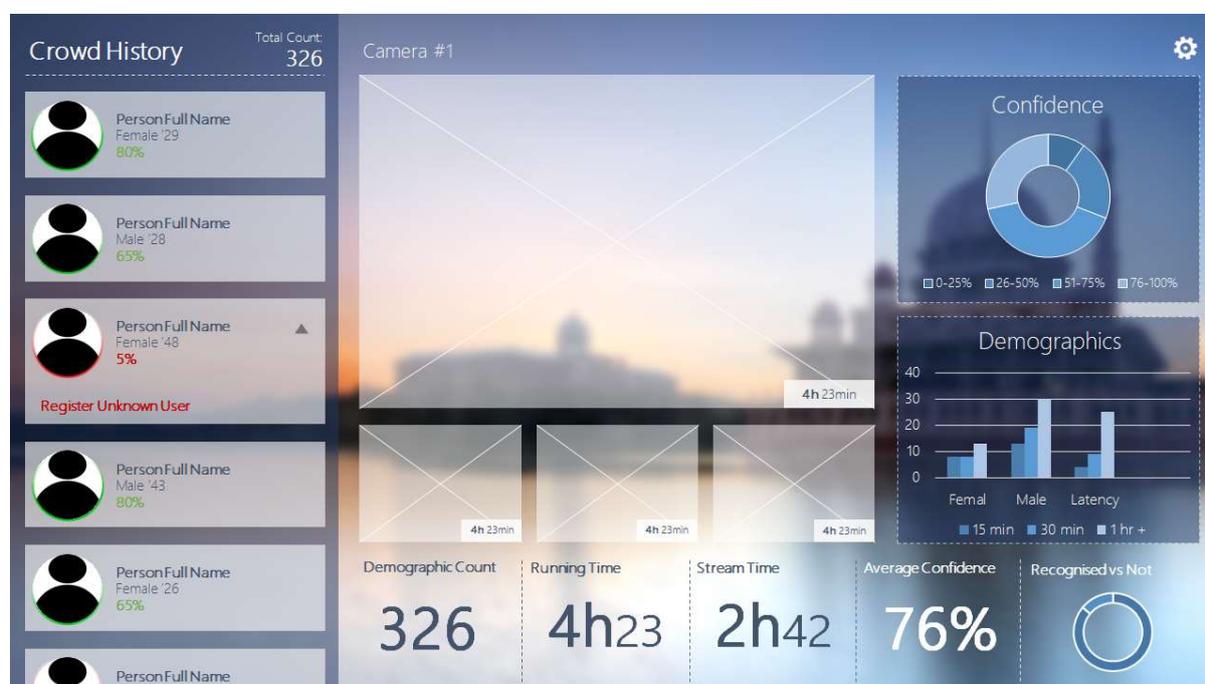
Figure 1: Crowd Tracking with the platform

Designed from the ground up to integrate into your business, the platform makes simple, easy-to-consume events and notifications available to business systems and processes.

1.1 Computer Vision is complex

Advanced Computer vision analytics has immense potential. In recent years advances in Computer Vision have exceeded human ability to find cancer tumours in MRI scans, read lips, detect suspicious behaviour outside of ATM's and detect obstacles in a road.

The potential for almost any business or sector is evident. However, implementing this type of analysis takes specialist skills and immense resources.

The four large cloud providers have teams of highly paid specialists implementing this type of work, but consuming in the real world is challenging.
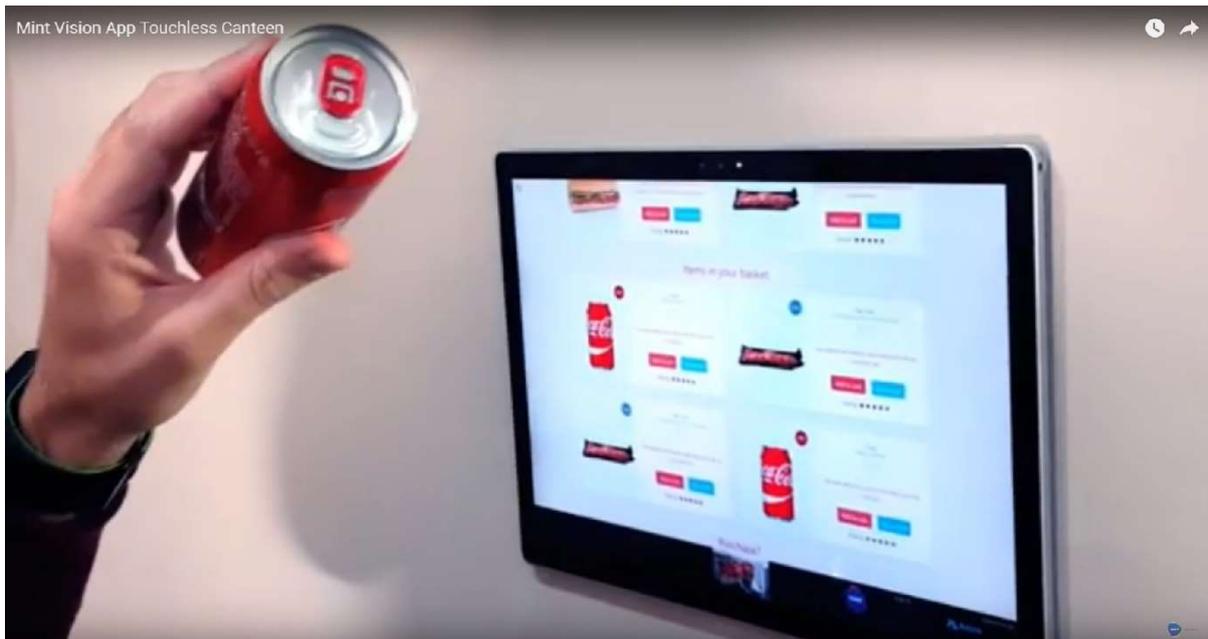
Figure 2: Adding items to a basket using Visual Recognition

For the consumer or enterprise, most of the incumbents are hardware manufacturers. This means that they only provide solutions that require their specific set of hardware, and moreover integration into business processes and systems is left to the purchaser.

Implementation of computer vision projects require a specific set of skills.

1.2 The Vision Platform

The Mint AI Solution is built with an Enterprise, scalable architecture to integrate into other systems and processes.

Able to consume any feed, from non-real-time sources such as video files and images, to real-time sources such as existing streaming cameras or even USB cameras and mobile devices, the platform consumes the feed, performs advanced analytics and provides simple, easy-to-action outcomes to other systems or processes.

The platform consists of an edge application, a windows 10 application that runs on any windows 10 machine including small embedded devices and tablets, and cloud-based processing. The application can run all-local, all-cloud or a hybrid combination of both, depending on the use-cases.
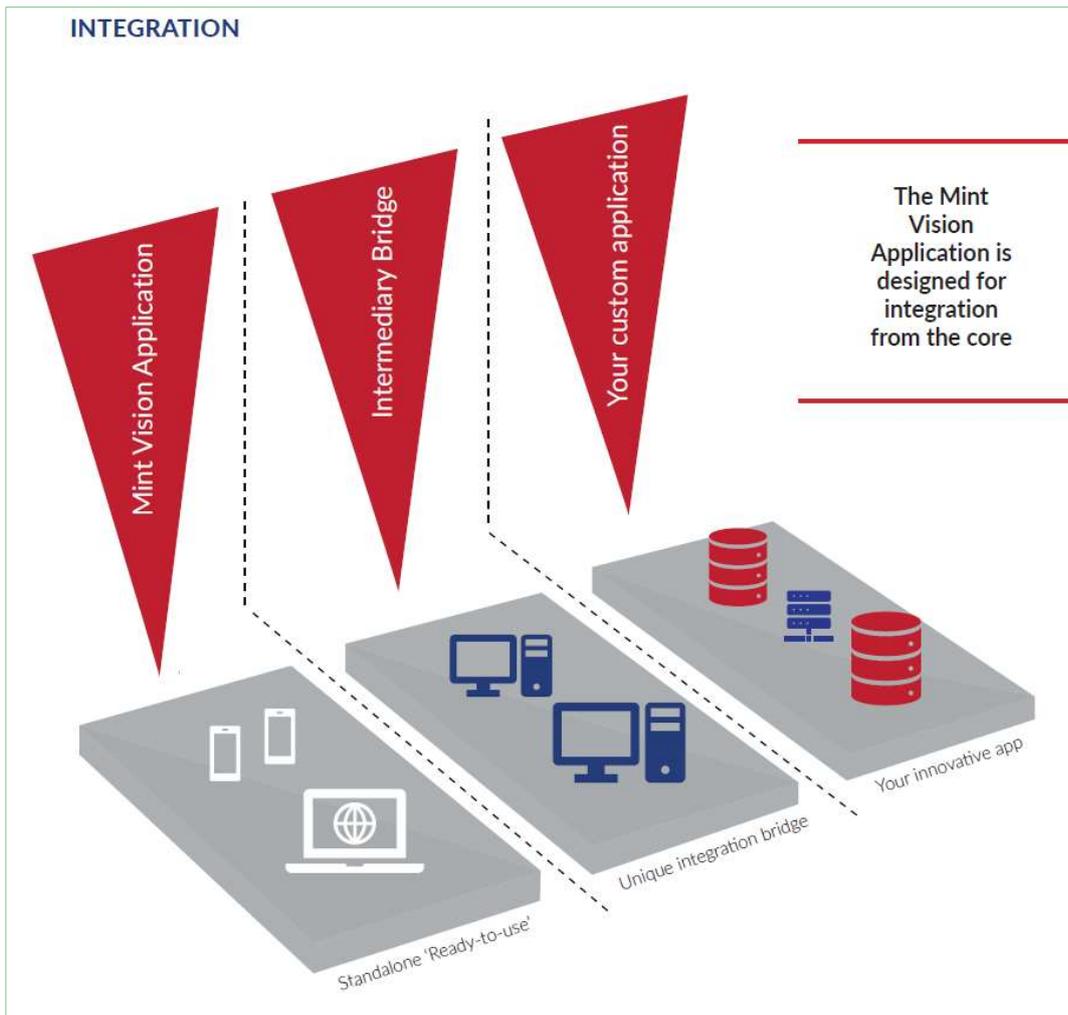
Figure 3: Designed for Integration

1.3 Public Display of Private Information

The platform can be used to display private information, such as Payslips, personal learning records and leave days – but in a public area. Since the tablet can automatically unlock whenever the correct person is present but lock again if any other faces are present in the viewfinder or when the current face disappears, it locks again.

This enables non-desk-bound workers to still have access to private information. Kiosks can now display personal or private information in a safe way.

2. Privacy

As an entity that stores, processes and retrieves personal information, including at times biometric, medical and other information, Mint takes their role of *custodian* of *your* personal information very seriously.

We are of the view that we have been provided with personal information in confidence, with trust and for a specific purpose, and consider that we must care for this information with utmost care.

Although in most cases the business is appointed to manage this information on behalf of a customer, and the customer bears the greatest burden of caring for private information, Mint Intelligent Insights take their custodia role just as seriously

To this end, the Mint Intelligent Insights business goes to extraordinary lengths to ensure strict privacy controls, including

- All data on-the-wire and at-rest is encrypted
- Consent is gained for all registrations, and is clear and transparent
- Every member of the business undergoes Privacy training annually
- Board-level Risk Registers on data storage and management are maintained and reviewed monthly
- Customers are guided on how to appoint Mint as a Data Provider so that compliance and ethics are maintained

Privacy and consent are built into the product at the core and Mint considers it a privilege to process the data subject's personal information, not a right.

3. The Mint Vision Platform in action

The Mint Vision platform can be used for a multitude of applications, from collecting feedback at a conference to opening turnstiles at a university. Below are some of the examples of where the platform has had an impact.

3.1 Crowd Tracking

Below is a sample, blank image of the application performing Crowd Tracking. In this mode, the application and platform will run in anonymous mode and collect information about anonymous faces. Information such as approximate age, gender and face expressions is collected into a database, and a PowerBI Dashboard is made available.
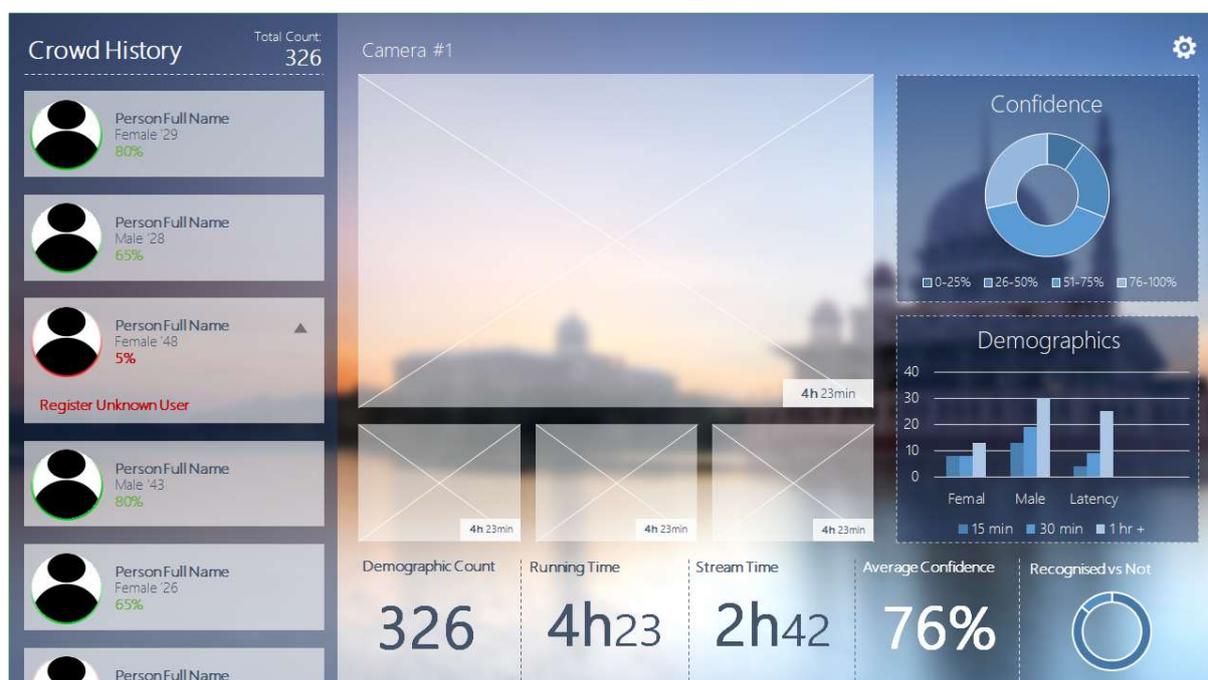
Figure 4: Crowd Tracking Screen

An example of the dashboard is displayed below:
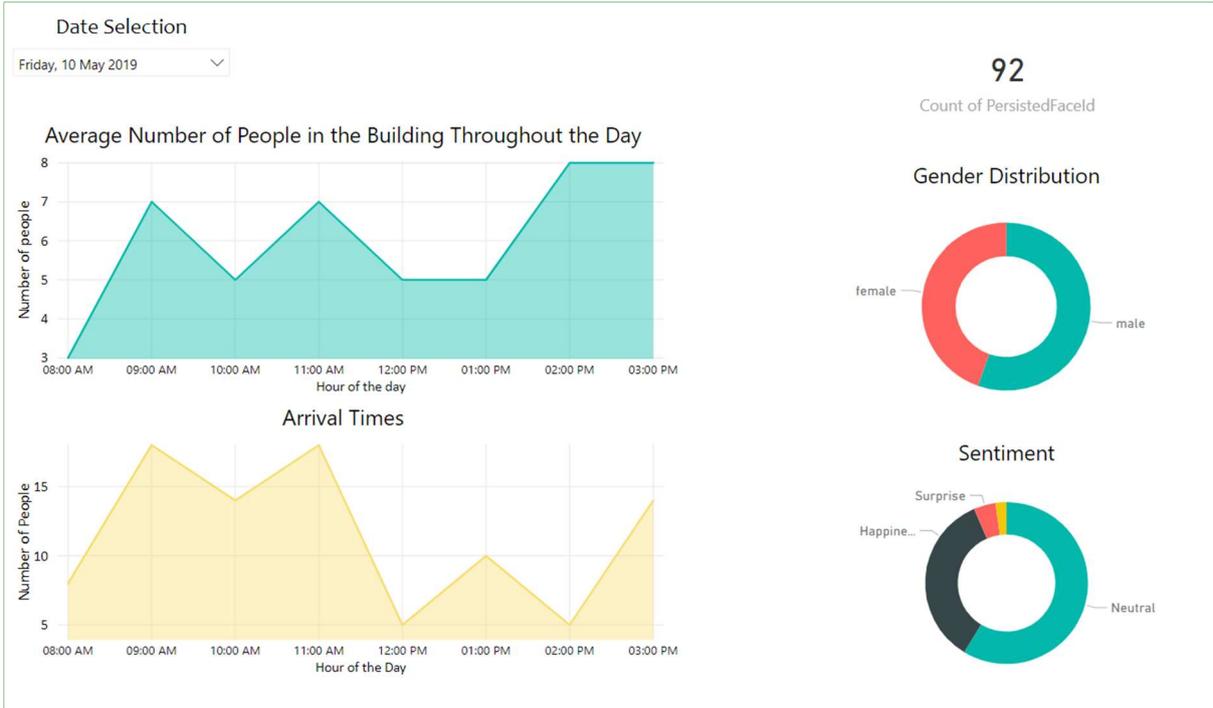
Figure 5: Sample Crowd Tracking Report

At a recent customer, videos of 16 stations at an event were submitted for analysis. Over 176,000 face recognition events were captured.
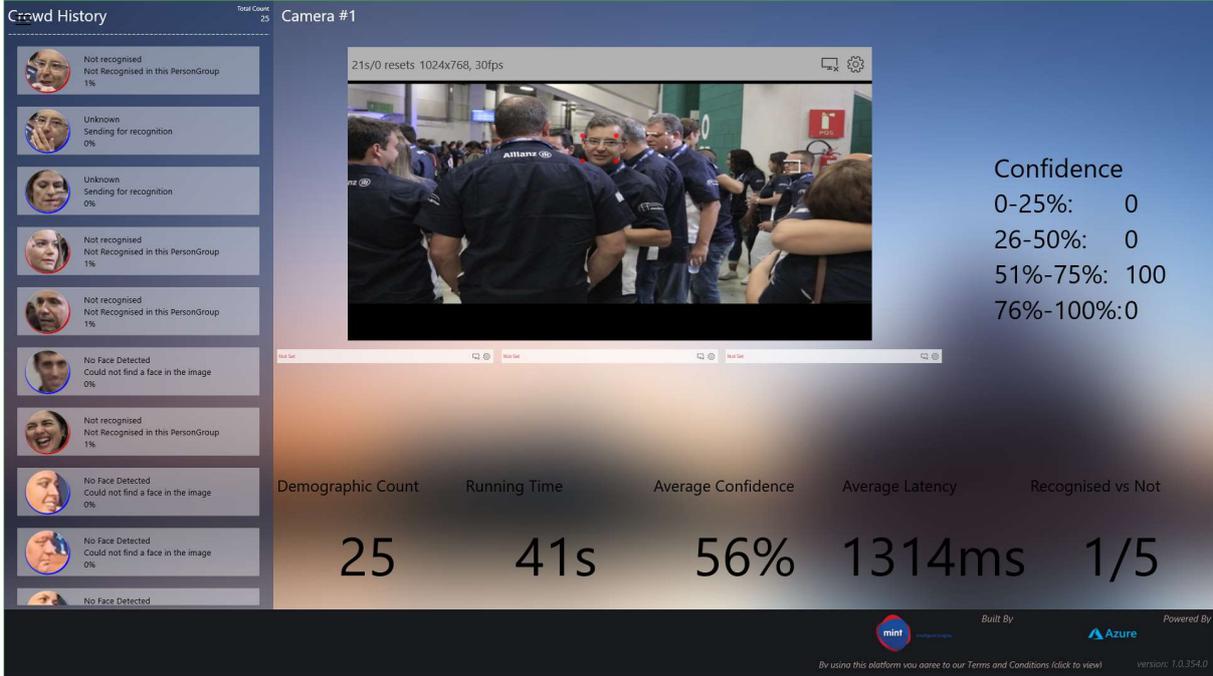

Figure 6: Event Tracking

3.2 Registration

Regardless of the use-case, when *recognition* is required then users must register (and accept terms and conditions) before they can use the system. When registering, users have the option to scan their drivers' license or ID card. Their details are extracted from the image and sent to Home Affairs (if they are South African Citizens) for verification.
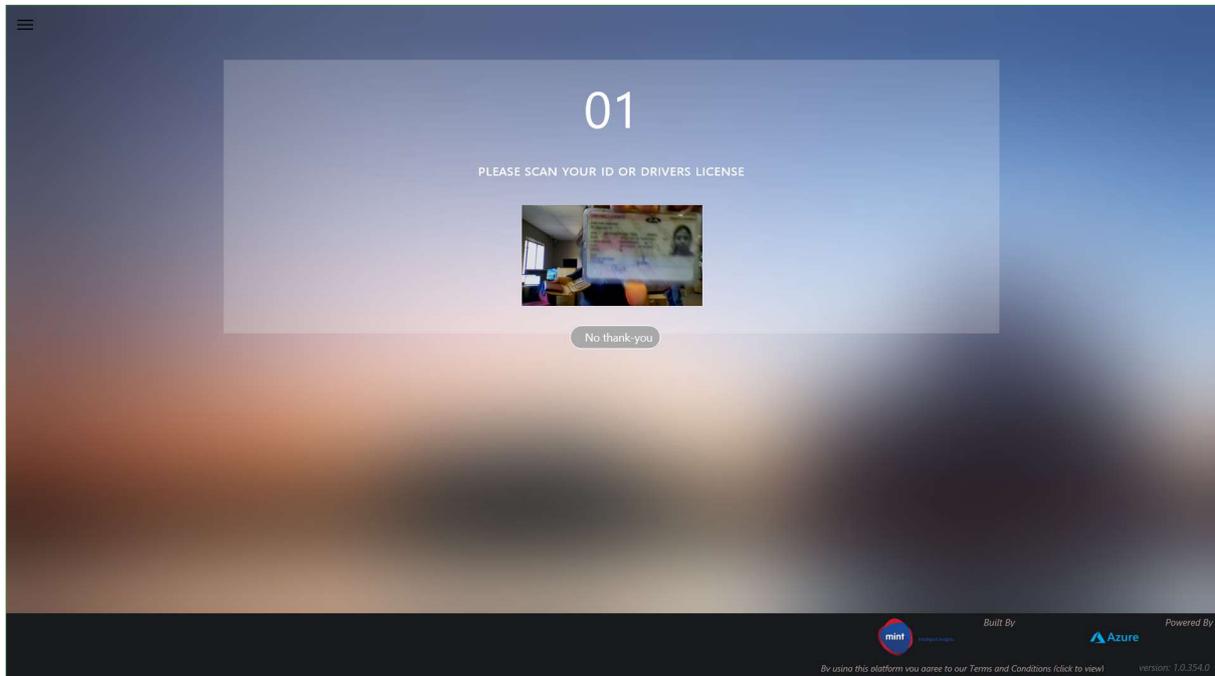
Figure 7: Drivers license being scanned

Once their details are confirmed, several images are captured of them to complete the registration.
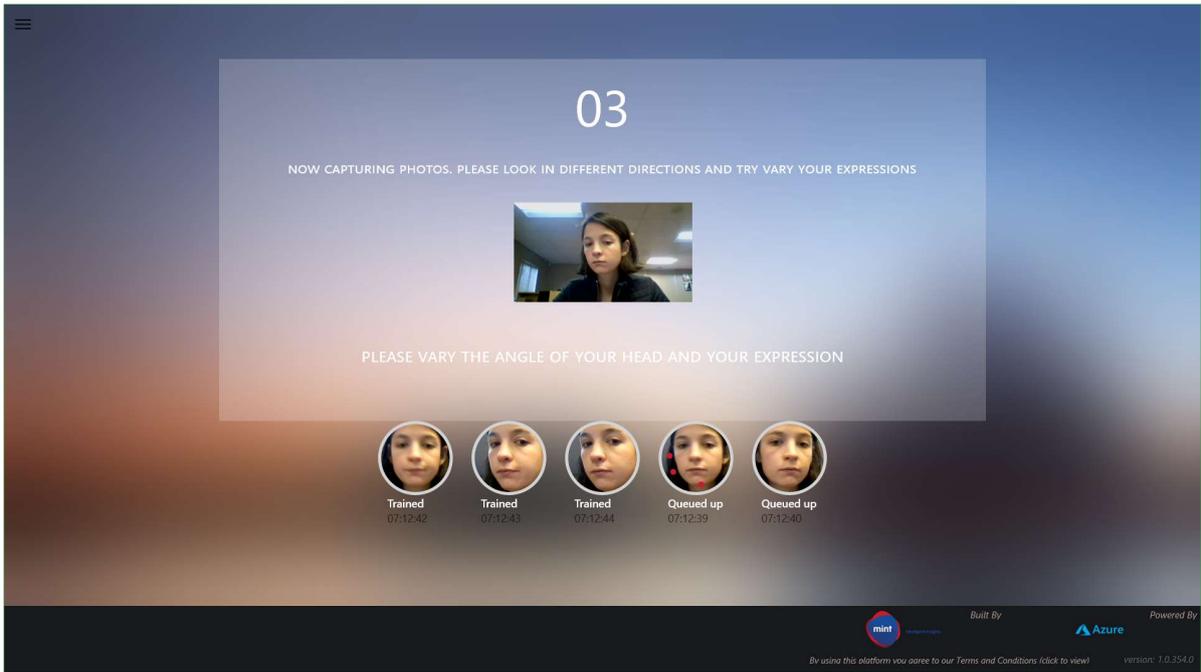
Figure 8: Capturing initial images

3.3 Public Clinics

3.3.1 Queue Management

The solution can be used for clinic management. The solution allows for registration of users, and then Queue Management for patients.
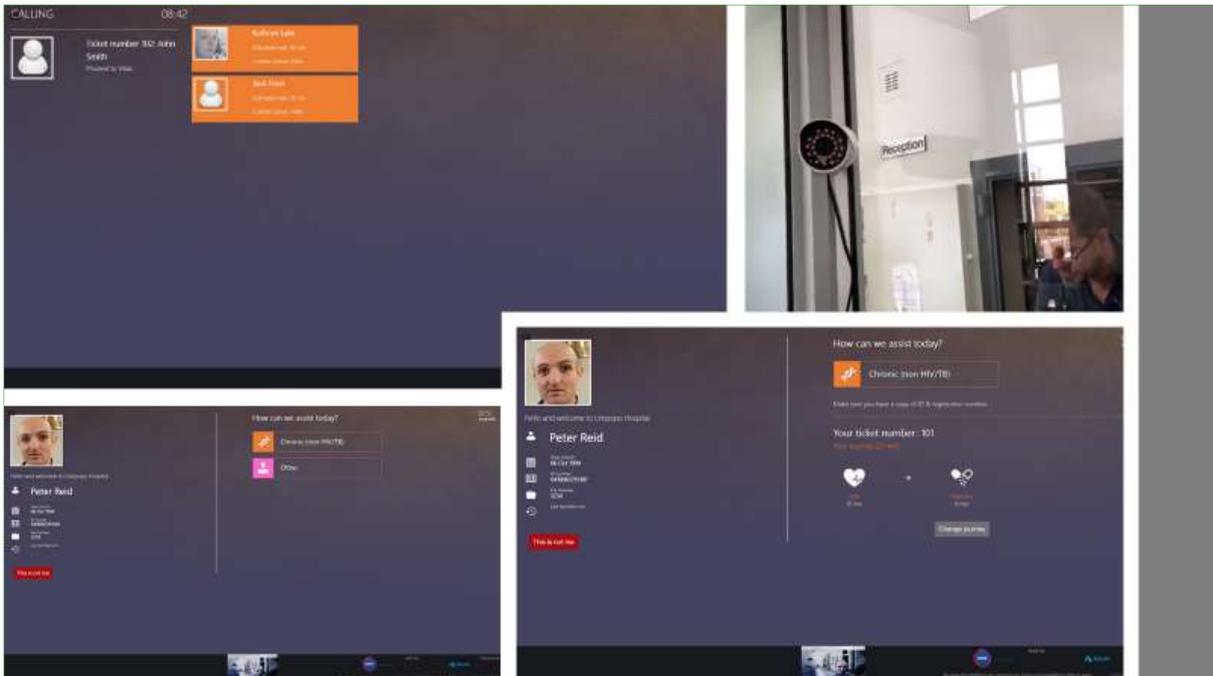
Figure 9: Basic Queue Management at a clinic

3.3.2 Patient Notes and Dispensing

The solution also can capture notes against a patient, including dispensing of medication to that patient once their face is recognised
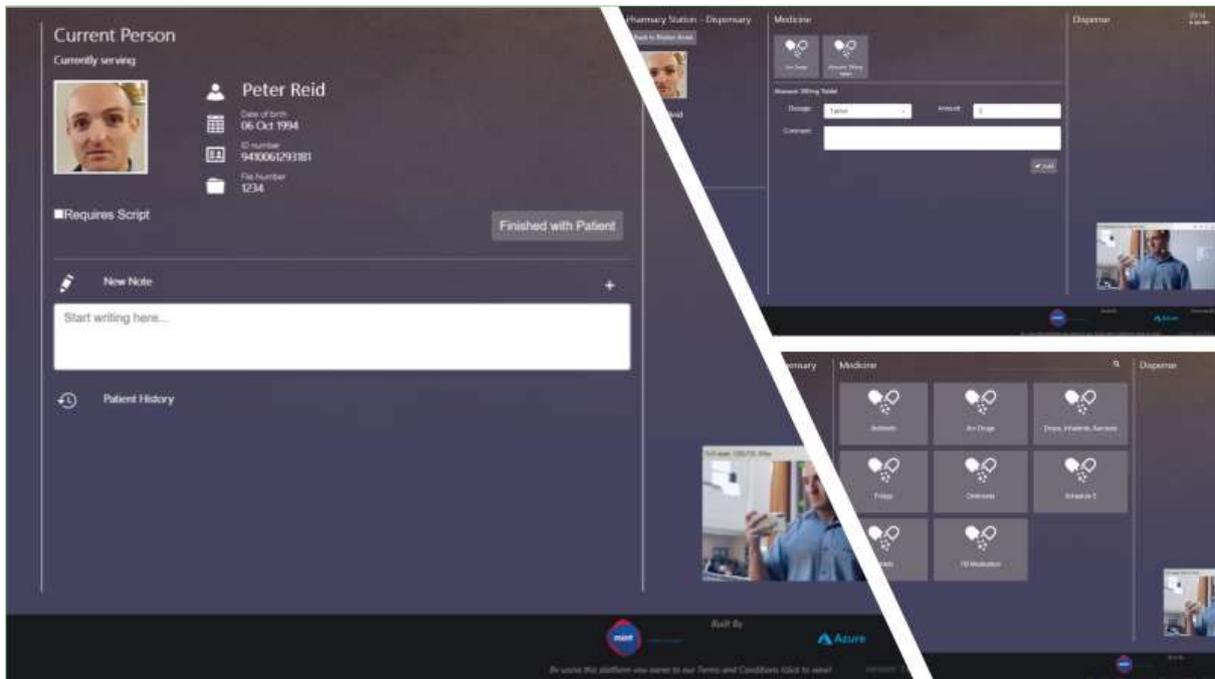
Figure 10: Capturing patient notes and dispensing medication

3.3.3 Doctor Greeting

The system can also display private, patient information to doctors in a public space, by only showing them information targeted to them when their face is visible in the viewfinder.
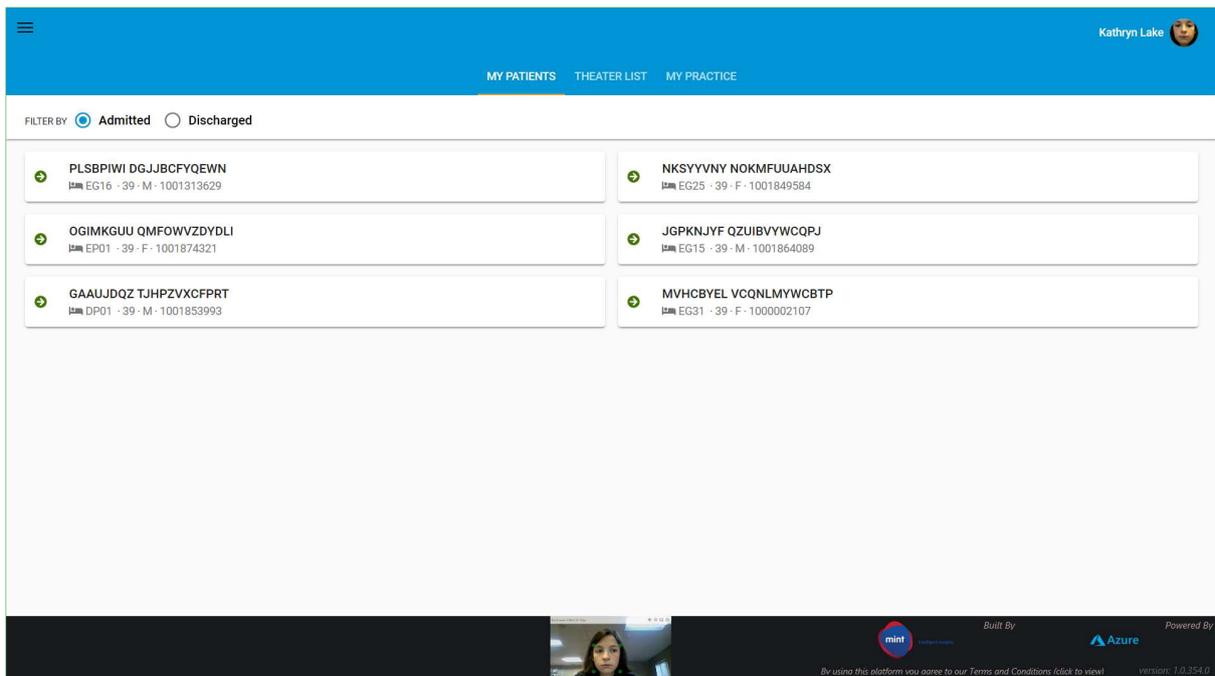
Figure 11: Patient infromation displayed to a doctor

3.4 Greeting Screens

The solution can also implement custom screens to greet users at a building reception, including calling the host to notify them that a person has arrived.
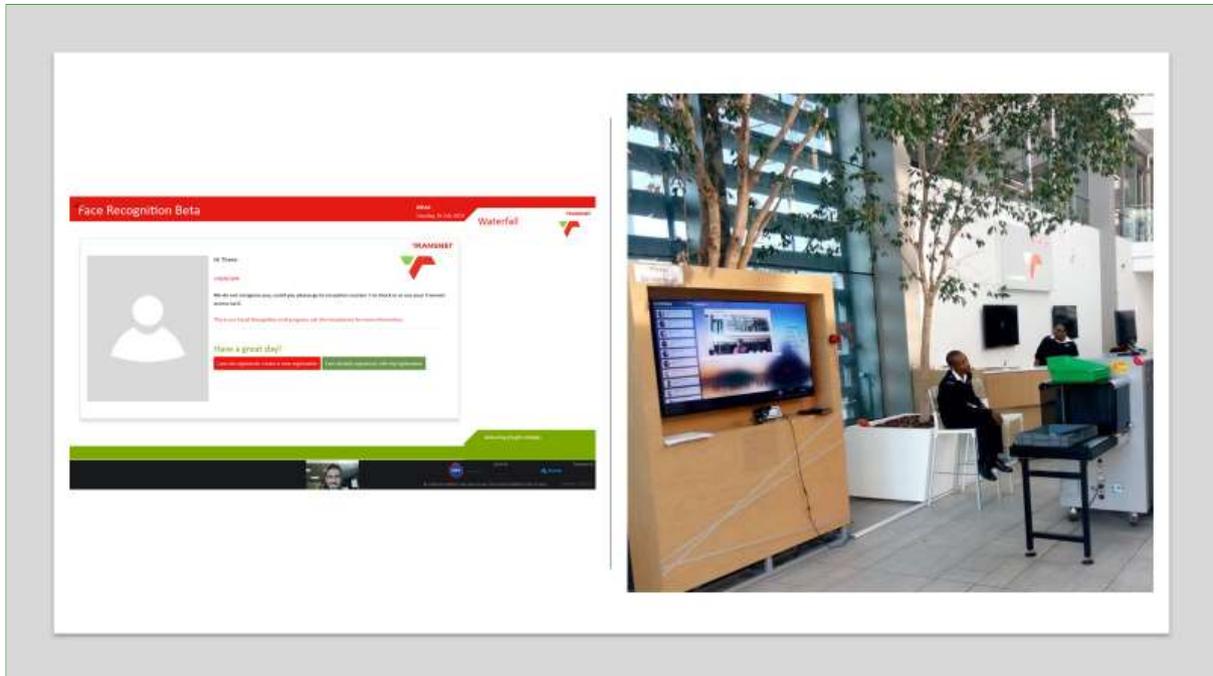
Figure 12: Solution greeting customers and visitors
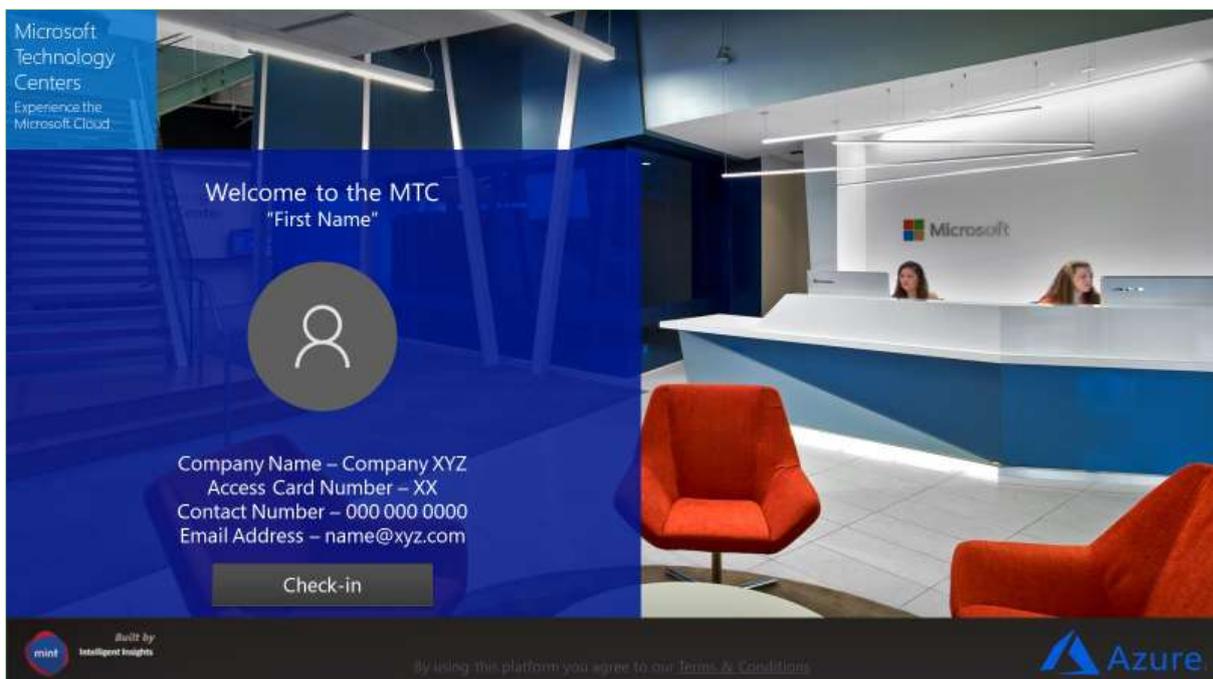
Figure 13: Greeting screen for visitors

The system can also look up from Outlook what current meeting rooms are busy and who the attendees are, and using that can determine which meeting room to send the user to and which other members of their company are present.

*Figure 14: Personalised Reception Greeting*

3.5 Conference Registration, greeting and surveys

The application can be used for registration at conferences, including displaying information to each delegate such as their personalised agenda. The application can also extract information from delegates, for example customized feedback forms. As the user approaches the tablet, their details are displayed to them.

Figure 15: Personalised agenda at a conference

The users can interact with the system and capture their information, for example feedback information.
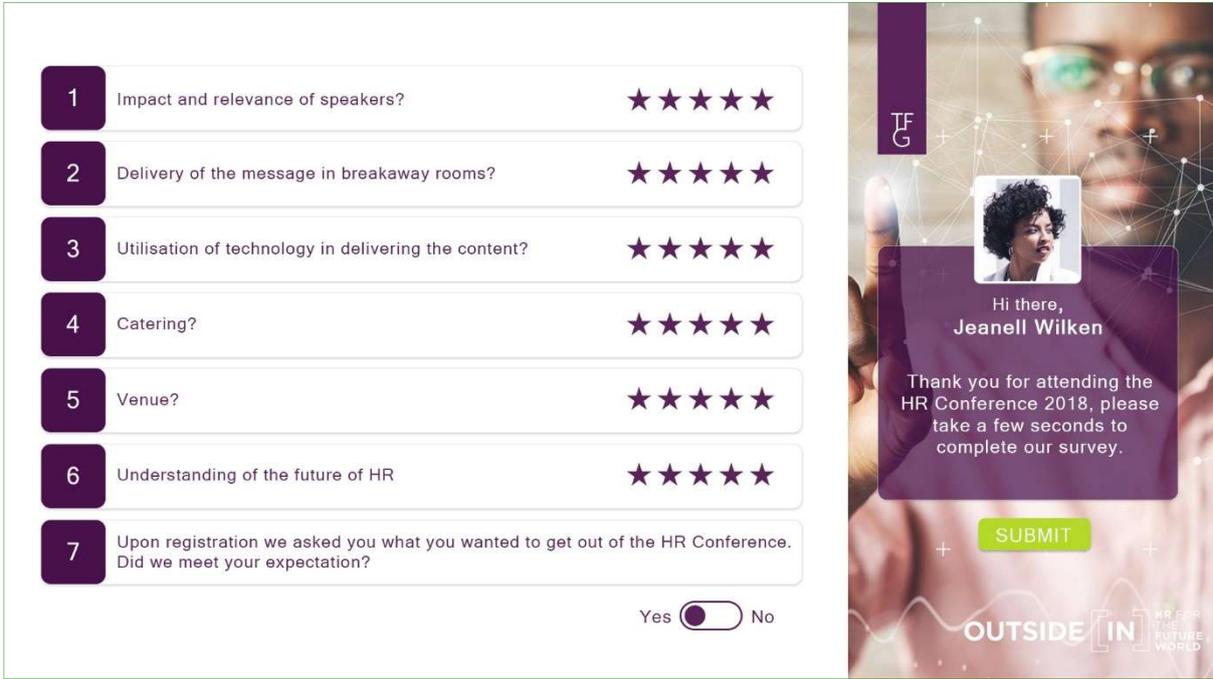
Figure 16: Feedback Survey with Face Recognition

3.5.1 Celebrity Greeting

The Mint Vision Application was used to greet finalists at the Sunday Times Top 100 Awards in 2018
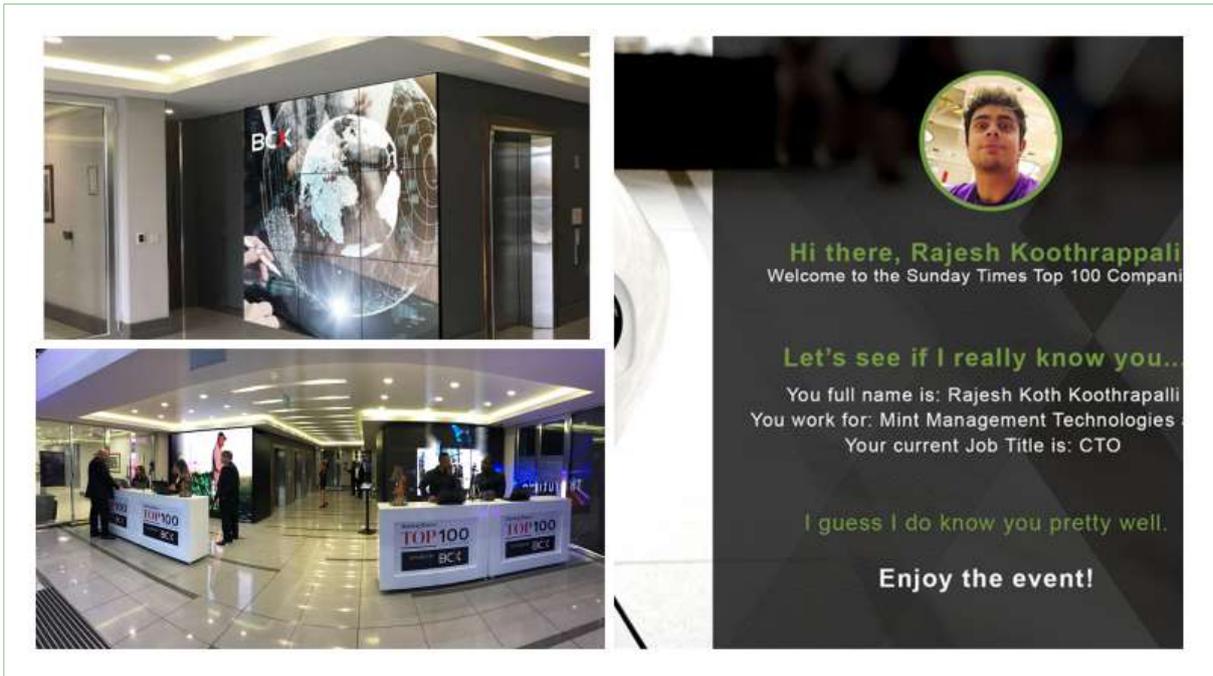
Figure 17: Greeting celebrities

3.6 Triggering External Evens – interacting with the world

The Mint Vision Platform has extensive plugins and integration points to allow for the triggering of external hardware such as doors, lights – and even coffee machines.
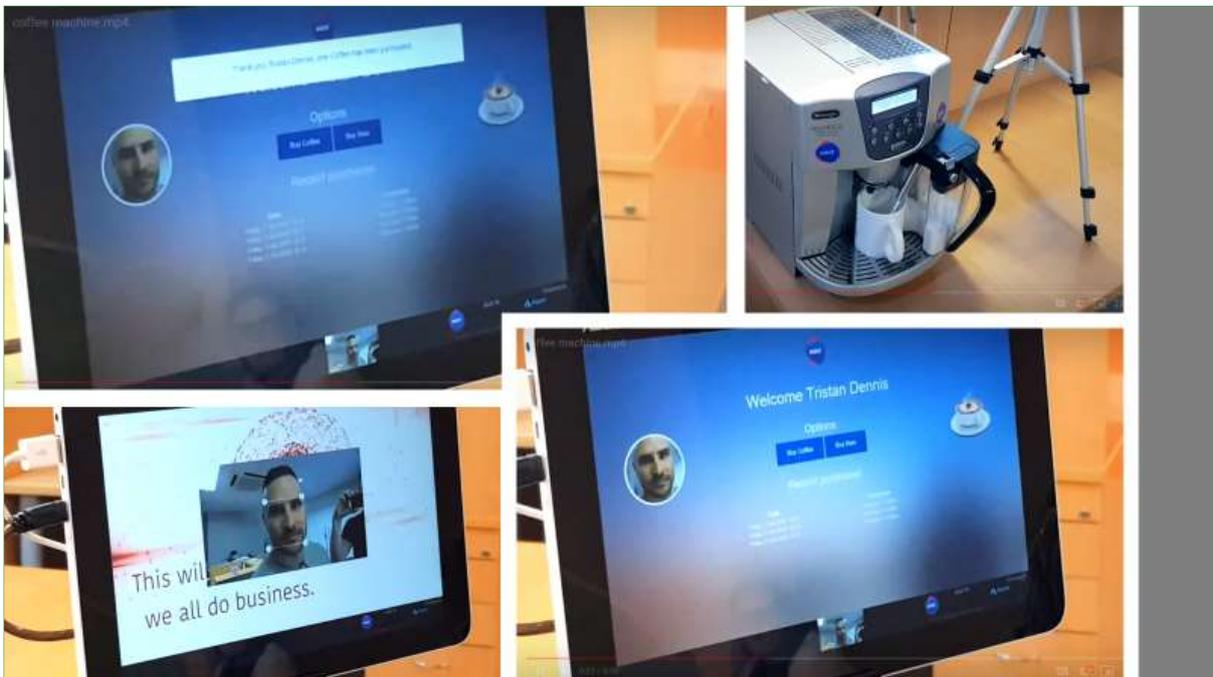
Figure 18: The Platform making coffee at Mint

The platform can be configured to trigger anything from basic lights to complex door triggers

Figure 19: A No-Touch Access Control Pilot

4. Platform Description

4.1 Turn Physical Actions into Business Process triggers

In a business or public office, business processes are triggered by actions: an arriving customer, an email, new deliveries, etc.

The Mint Vision Platform gives you the ability to use any physical motion or recognised object into a business input: someone dropping something, a customer waking into the reception, someone holding up a coke in a shop.

4.2 Architecture

In its' simplest form, the application runs on a single stand-alone device or computer and does all the processing locally. Images are ingested from a USB camera or video file or other feed and processed internally in combination with Azure. The results are pushed into a database in the cloud, available to be consumed in one of the standard reports or for the customer themselves to build custom reports.

In an enterprise environment, the application can run with multiple parallel processing units, where each unit is offloaded to a different processing location specialised specifically for what it does.
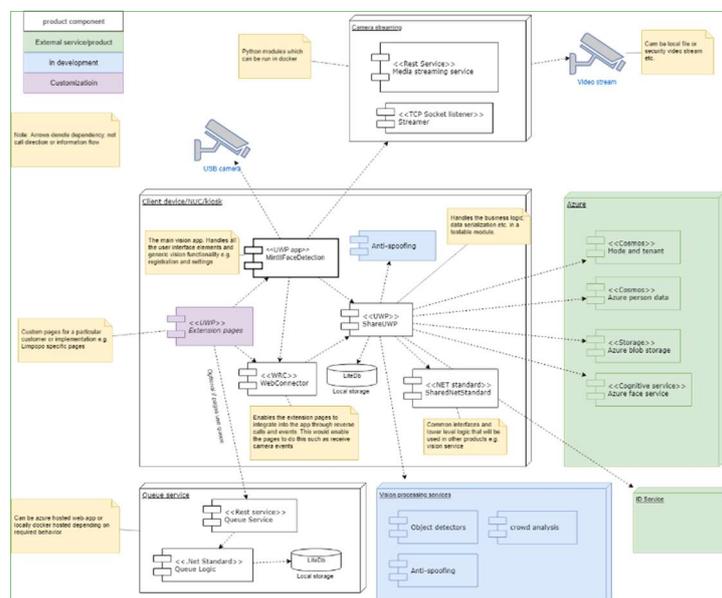

Figure 20: Pluggable Architecture

4.3 Feeds

The platform can consume feeds from any video source, including

- Streaming cameras
- USB cameras
- Video files
- Still images/sets of still images
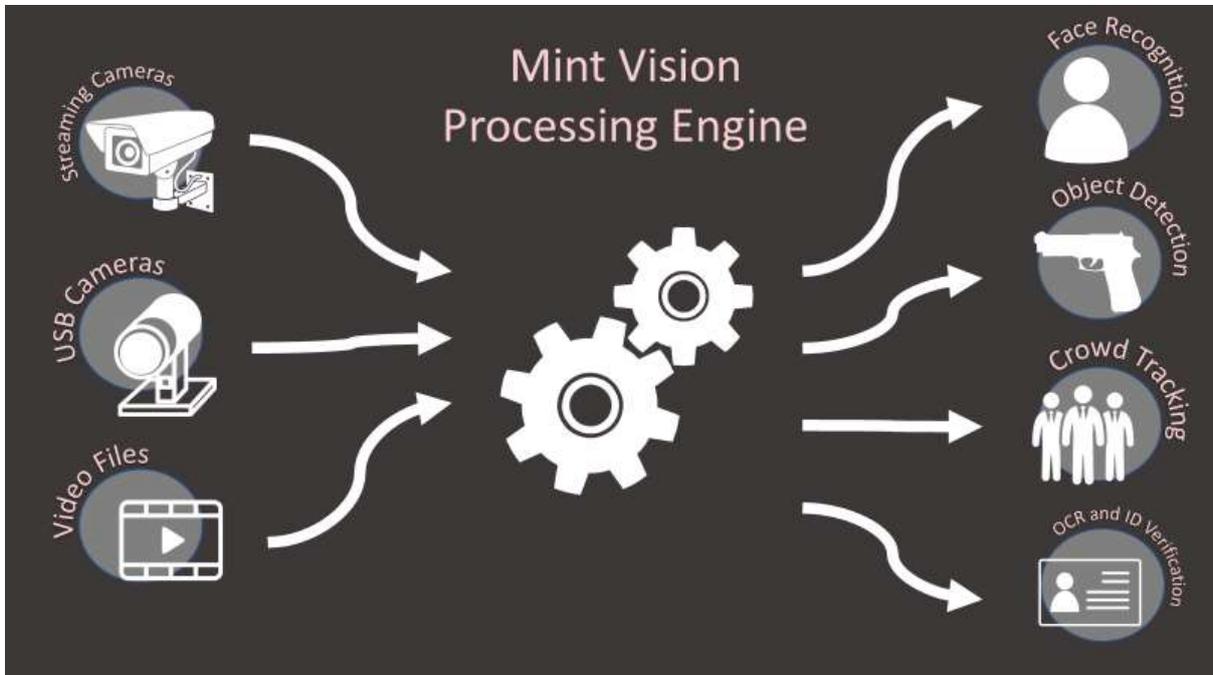
Figure 21: Ingestion and Processing Engine

The ingestion engine is designed to be decoupled from the rest of the platform; that is, how you process the data is independent of how the feed is ingested into the platform. This means that any processing can be done on any feed; for example, if we wish to monitor a conveyor belt and act on detecting unusual objects, whether we process a live stream, or a video file is immaterial. This means that

- Testing is dramatically simplified, as video files can be used to test the solution until acceptable results are achieved and the solution is moved into the real-world
- The software is independent of any hardware, allowing the solution to work with existing infrastructure or for the customer to choose preferred suppliers without being forced to work within a specific framework or set of hardware.

4.4 Processing Engine

The Processing Engine performs the following functions:

- Face Identification
  For example, in Crowd Tracking, the platform can identify and monitor the number of faces, approximate age and Gender, emotional status, etc.
- Face Recognition
  The platform can recognise known, or trained faces, in a feed – giving an idea of confidence and automatically re-training if necessary.

- Object Recognition
  The platform can identify, or recognise, objects within the camera frame.
- Sequence Recognition
  The platform can recognise sequences of events such as making a tackle in a sports game, or an accident on a road from traffic cameras.
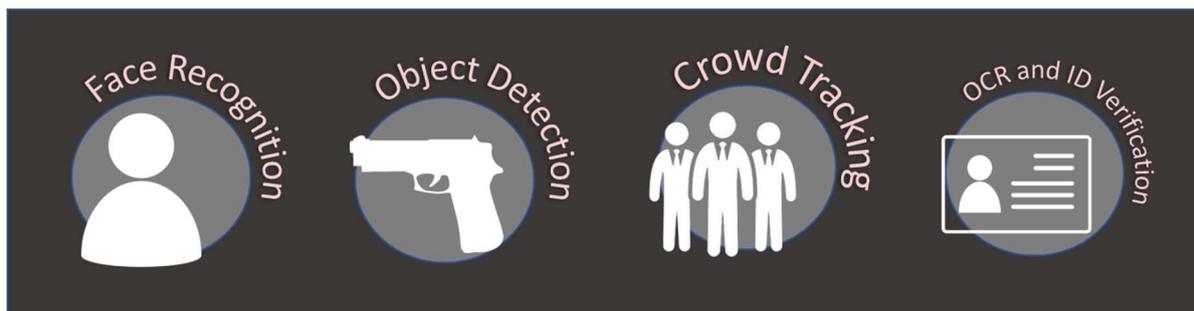
Figure 22: Trigger Events

4.5 Recognition events

Once the platform has recognised objects, events, sequences or people, it can feed the results into any business process or system. This includes the ability to extract OCR information from forms and receipts that a user may hold up to the camera.

Physical triggers can range from opening a door, stopping a conveyor belt or running a coffee machine, to name a few.
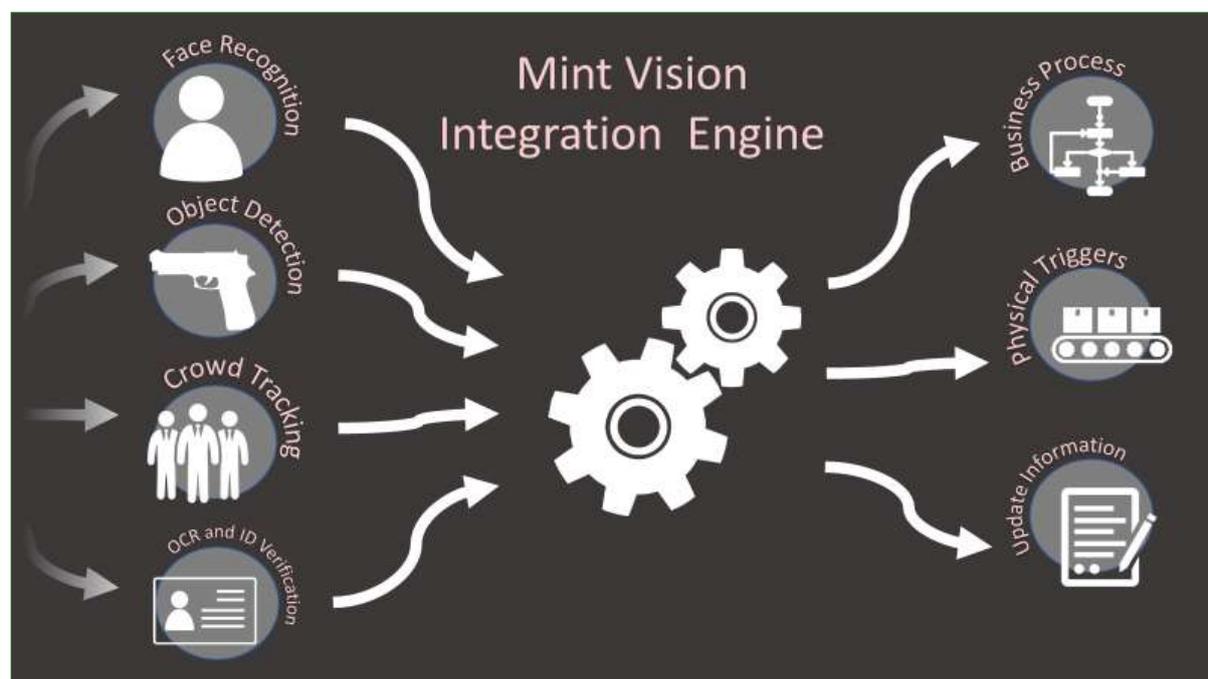
Figure 23: Business Process Triggering

4.6 Hosting – Web Integration

The application hosts any web page or web application, and feeds events into the page regarding Face Recognition, Voice Recognition, Object Detection or other updates. The pages require little to no integration

The Shell hosts any web application, and is fed voice and face events from the platform. Integration is simple with our API and ensures secure integration and very little code changes to get this working.
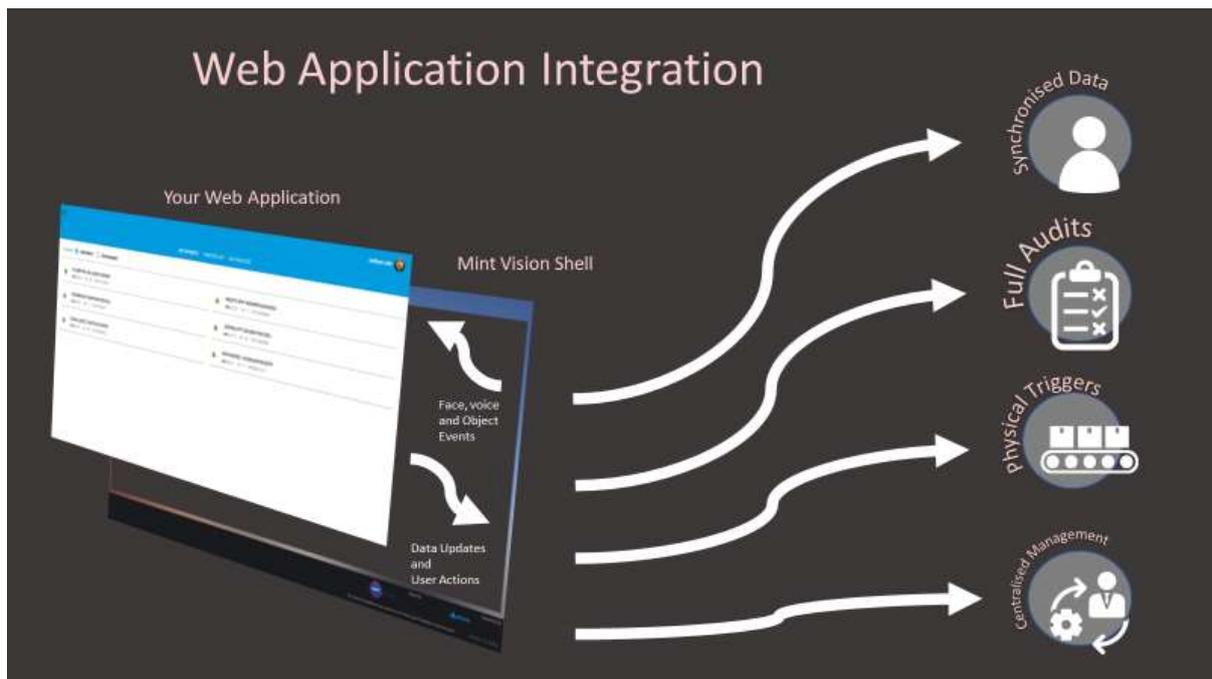
Figure 24: Shell Page Integration

4.7 Features and Tiers

The product is licensed on features and tiers. The below table describes these features and how they relate to the product

Table 1: Feature List of the platform

Feature	Details
Simple and secure person registration	The ability to securely register new users, by scanning of their ID card, Drivers license, other identification or manually entering their details. The registration process should take between 20 and 25 seconds.
Any camera or video stream as input	The application can take any stream source, including USB cameras, Video files, and Streaming/Network cameras.
Simple management & monitoring	Designed to be deployed in the Enterprise environment, the application has features that allow for easy

	management of the devices such as remote "heart-beat" monitoring and automatic version updates.
Access control	Integration with building management systems to trigger turnstiles, sliding doors, etc.
Recognized crowd tracking	Tracking of known and anonymous faces through open areas with existing or new cameras. The ability to recognise trained individuals, or pick up repeated faces and generalised markers such as emotion and gender.
Person information, that travels with their face	The system can track customised and detailed information about a person, such as medical records or student information. The platform automatically synchronises this information across all systems and devices in the deployment.
Integration	The Platform is home-grown by our own experienced team of integration developers, and has been designed from the ground-up to be integrated into customer systems. The platform exposes API's and event hooks to allow for easy integration.
Custom forms (not simple text)	The platform has the ability to configure custom input forms for recognised people, such as patient medical forms or student feedback forms. This information is automatically synchronised across all devices in the cluster.
Responsive and great UX for users	The platform has been designed to be used by end-users with little to no training, and as such has a fast, responsive, easy-to-use interface.
Simple deployment & configuration	The platform and Application can be deployed and configured with little to no training and configuration.
Spoof detection	On enterprise applications the platform can detect "liveness" or spoofing, so that you can be sure that live people are using the system.
Image/camera auto-correction	The system automatically corrects for lighting and contrast issues with images, along with more complex corrections such as Histogram correction.
User management	The application allows for secure and simple management of users
Compliant, secure system that doesn't compromise user data	The system can securely transfer and store data. Secure APIs are used and encryption is used on-the-wire and at-rest.
Accuracy/2nd factor authentication/active learning	The system employs several methods for accuracy improvement including Active Learning, second-factor authentication and user-group-limiting features.
Custom content kiosk	The platform can host custom content for each scenario, that displays different user information depending on the customer and user. This is an easy-to-configure and easy-to-implement customisation
General (anonymous) crowd tracking	The platform can be used to track crowds in videos, camera feeds or other input. Instant PowerBI Reports are available on the results.
Queue management	The platform has a built—in Queue management system that uses Face Recognition to neatly queue people.

|	
Low cost hardware	Where required, the platform can run on low-cost hardware such as a Raspberry Pi or other SoC-type boards.
Object detectors	The platform can detect objects in frames, such as hard-hats, syringes, guns, or any other type of object that can be trained into the system.
Alerts of banned people	The ability to set an alert if a camera sees a person from a banned list
Vision as a service	A cloud-based, REST API allowing submission of faces to be processed as a service without the front-end UI being deployed.

5. Licensing

The Mint Vision Platform is charged per transaction. However, to cater for organisations that can currently only spend on a CapEx model, there are alternatives:

- Once-off software license (enterprise)
- Per-station license
- Per-transaction license

5.1 Licensing structure

The product has three "Tiers" that we license

Enterprise
- The full product, designed for large implementations and full integration

Standard
- Most of the features of the platform. Excludes advanced features like Liveness detection, and no integration

Basic
- Basic anonymous crowd-tracking or customized kiosk-displays
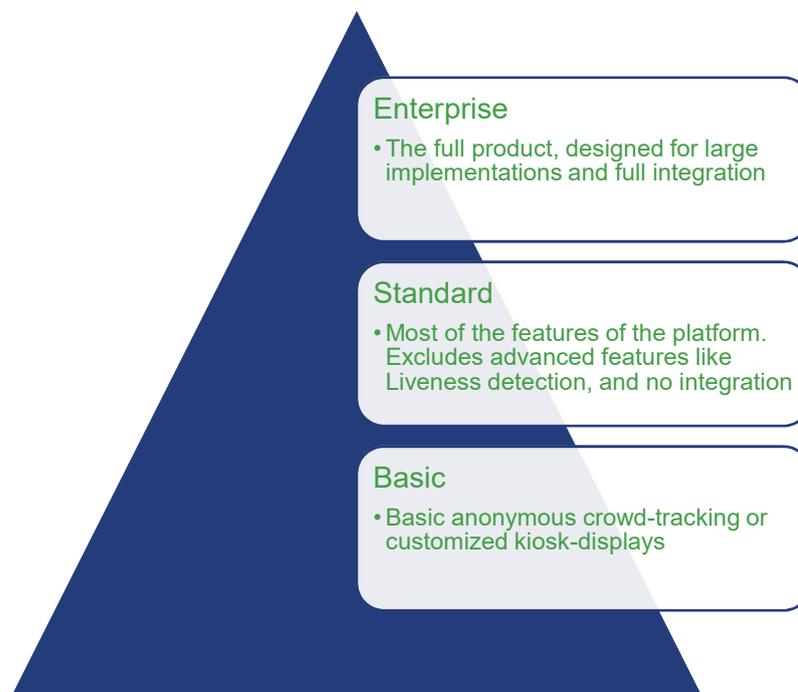
5.1.1 Features

The below list details the features for each version

Table 2: Licensing Features

Feature	Details	Enterprise	Standard	Basic
Simple/secure person registration	The ability to securely register new users, by scanning of their ID card, Drivers license, other identification or manually entering their details. The registration process should take between 20 and 25 seconds.	Yes	Yes	No

Any camera or video stream as input	The application can take any stream source, including USB cameras, Video files, and Streaming/Network cameras.	Yes	Yes	USB/file only
Simple management & monitoring	Designed to be deployed in the Enterprise environment, the application has features that allow for easy management of the devices such as remote "heart-beat" monitoring and automatic version updates.			
Access control	Integration with building management systems to trigger turnstiles, sliding doors, etc.	Yes	No	No
Recognized crowd tracking	Tracking of known and anonymous faces through open areas with existing or new cameras. The ability to recognise trained individuals, or pick up repeated faces and generalised markers such as emotion and gender.	Yes	Yes	No
Person data that travels with their face	The system can track customised and detailed information about a person, such as medical records or student information. The platform automatically synchronises this information across all systems and devices in the deployment.	Yes	Yes	No
Integration	The Platform is home-grown by our own experienced team of integration developers and has been designed from the ground-up to be integrated into customer systems. The platform exposes API's and event hooks to allow for easy integration.	Yes	No	No
Custom forms (not simple text)	The platform has the ability to configure custom input forms for recognised people, such as patient medical forms or student feedback forms. This information is automatically synchronised across all devices in the cluster.	Yes	No	No
Responsive and great UX for users	The platform has been designed to be used by end-users with little to	Yes	Yes	Yes

	no training, and as such has a fast, responsive, easy-to-use interface.			
Simple deployment & configuration	The platform and Application can be deployed and configured with little to no training and configuration.	Yes	No	No
Spoof detection	On enterprise applications the platform can detect "liveness" or spoofing, so that you can be sure that live people are using the system.	Yes	No	No
Image/camera auto-correction	The system automatically corrects for lighting and contrast issues with images, along with more complex corrections such as Histogram correction.	Yes	Yes	Yes
User management	The application allows for secure and simple management of users	Yes	Yes	No
Compliant, secure system that doesn't compromise user data	The system can securely transfer and store data. Secure APIs are used, and encryption is used on-the-wire and at-rest.	Yes	Yes	Yes
Accuracy/2nd factor authentication/active learning	The system employs several methods for accuracy improvement including Active Learning, second-factor authentication and user-group-limiting features.	Yes	Yes	No
Custom content kiosk	The platform can host custom content for each scenario, that displays different user information depending on the customer and user. This is an easy-to-configure and easy-to-implement customisation	Yes	Yes	No
General (anon) crowd tracking	The platform can be used to track crowds in videos, camera feeds or other input. Instant PowerBI Reports are available on the results.	Yes	Yes	Yes
Queue management	The platform has a built—in Queue management system that uses	Yes	No	No

	Face Recognition to neatly queue people.			
Low cost hardware	Where required, the platform can run on low-cost hardware such as a Raspberry Pi or other SoC-type boards.	Yes	No	No
Object detectors	The platform can detect objects in frames, such as hard-hats, syringes, guns, or any other type of object that can be trained into the system.	Yes	No	No
Alerts of banned people	The ability to set an alert if a camera sees a person from a banned list	Yes	No	No
Vision as a service	A cloud-based, REST API allowing submission of faces to be processed as a service without the front-end UI being deployed.	Service license	Service license	Service license

Each "tier" can be licensed with a standard, once-off license or a per-impressions license fee.

For License Quotes, please contact Mint directly.