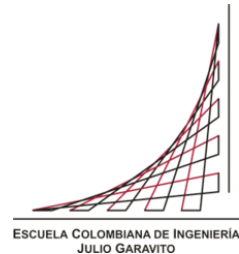


Maestría en Gestión de Información

Definición de una solución de recolección y visualización de información utilizando un sistema biométrico en una institución de educación superior

**Laura Niño Muñoz
Iván David Reina Espinosa**

**Bogotá D.C., Colombia
14 de Diciembre de 2018**

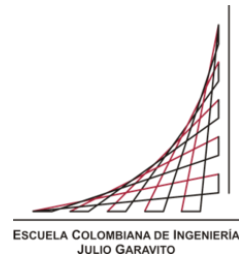


Definición de una solución de recolección y visualización de información utilizando un sistema biométrico en una institución de educación superior

Trabajo de Grado para optar al título de Magister
en Gestión de Información

Oswaldo Castillo Navetty
Doctor en Redes, Conocimientos y Organizaciones
Director

Bogotá D.C., Colombia
14 de Diciembre de 2018



Resumen

En la actualidad, varias instituciones de educación superior recolectan manualmente los datos de los participantes en diferentes actividades académicas que realizan. Estos datos son digitalizados y analizados por personal de la entidad, lo que hace que el proceso sea demorado y esté sujeto a errores humanos.

El objetivo del proyecto es definir una solución sistematizada que permita la recolección y la visualización de la información para agilizar la toma de decisiones.

El proyecto se centra en cuatro escenarios que se identificaron en una institución de educación superior. Para el desarrollo de estos escenarios se evaluó los resultados obtenidos por proyectos de investigación anteriores, la información que debía obtenerse, un sistema biométrico que apoyara la recolección de datos, el diseño de un modelo de datos, el análisis requerido y su visualización para la toma de decisiones.

Palabras clave: Visualización de información, biometría, autenticación.

Índice General

Resumen	3
Índice General	4
Índice de Figuras	6
Índice de Gráficos	7
Índice de Tablas	7
1 Introducción	9
2 Problemática (Justificación)	10
2.1.1 Pregunta de Investigación	12
3 Objetivos	13
3.1 Objetivo General	13
3.2 Objetivos Específicos	13
4 Alcance y Limitaciones	14
5 Marco Teórico	15
5.1 Inteligencia de negocios	15
5.1.1 Componentes de inteligencia de negocios	17
5.2 Proceso de extracción, transformación y carga (ETL)	18
5.2.1 Repositorio	19
5.2.2 Datamart	19
5.3 Explotación de la información	20
5.4 Visualización, interpretación y análisis de datos	20
5.5 Big Data	23
5.6 Biometría	25
5.6.1 Enrolamiento	27
5.6.2 Autenticación	27
5.6.3 Identificación	27

5.6.4	Matcher	27
5.6.5	Minucia.....	28
6	Estado del Arte	29
6.1	Análisis y visualización de datos	29
6.2	Biometría	33
7	Metodología	37
7.1	Entendimiento y búsqueda de información.....	38
7.2	Análisis y definición del diseño y desarrollo.....	38
7.3	Posibles soluciones	39
7.4	Resultados.....	39
8	Desarrollo	40
8.1	Arquitectura del sistema	40
8.1.1	Módulo de Enrolamiento	41
8.1.2	Módulo de Autenticación	42
8.1.3	Base de Datos.....	42
8.1.4	Matcher o Cotejador.....	43
8.1.5	Visualización	44
8.2	Comparación y Análisis herramientas de visualización	44
8.3	Análisis Sistemas de autenticación	49
8.3.1	Biometría dactilar	52
8.3.2	Reconocimiento facial	54
8.3.3	Reconocimiento del iris	57
8.3.4	Reconocimiento de voz	58
8.3.5	RFID.....	59
8.4	Escenarios.....	59

8.4.1	Asistencia a clases.....	60
8.4.2	Evaluaciones y exámenes finales	67
8.4.3	Charlas y conferencias.....	71
8.4.4	Simulacros de evacuación	75
9	Prueba Piloto	80
10	Costo – Beneficio	83
11	Conclusiones	84
12	Trabajos Futuros	86
13	Bibliografía	87
14	Anexos	92
	Anexo 1. Cronograma	92
	Anexo 2. VeriFinger SDK Brochure.....	92
	Anexo 3. Face Verificaction SDK Brochure	92
	Anexo 4. VeriEye SDK Brochure,.....	92
	Anexo 5. VeriSpeak SDK Brochure.....	92
	Anexo 6. NCheck Bio Attendance Brochure.....	92
	Anexo 7. SentiVeillance SDK Brochure.....	92

Índice de Figuras

Figura 1	Arquitectura de una solución de negocios. Elaboración propia	17
Figura 2	Escala de valor dentro del análisis de la Información. Fuente: Gartner.	23
Figura 3	Metodología. Elaboración propia	37
Figura 4	Arquitectura de la solución. Elaboración Propia	40
Figura 5	Modelo de datos. Elaboración Propia	43
Figura 6	Cuadrante Mágico de Gartner. Fuente: Gartner	45

Figura 7 Extracción de minucias. Elaboración propia	53
Figura 8 Diagrama de Flujo, Registro asistencia a clase. Elaboración propia	64
Figura 9 Dashboard del perfil de los estudiantes. Elaboración propia	65
Figura 10 Diagrama de Flujo Registro Asistencia Exámenes. Elaboración propia.....	69
Figura 11 Dashboard de asistencia a exámenes finales. Elaboración propia	70
Figura 12 Diagrama de Flujo Registro Asistencia a Conferencias. Elaboración Propia	73
Figura 13 Dashboard de asistencia a eventos. Elaboración propia	74
Figura 14 Diagrama de Flujo Registro Conteo Simulacros. Elaboración propia	77
Figura 15 Dashboard de participación en los simulacros de evacuación. Elaboración propia	78
Figura 16 Dashboard del piloto de asistencia a cursos. Elaboración propia.	81

Índice de Gráficos

Gráfico 1 Evaluación sistema biometrico vs manual. Recuperado Fingerprint-Based Attendance Management System.....	61
Gráfico 2 Absentismo por semestre. Recuperado IMPLEMENTATION OF AN EFFECTIVE	62

Índice de Tablas

Tabla 1 Comparación de herramientas de visualización de datos. Elaboración propia.....	47
Tabla 2 Requerimientos captura facial. Fuente: ICAO.....	56
Tabla 3 Costo-Beneficio del proyecto. Elaboración propia	83

1 Introducción

Gran parte del proceso educativo comprende la asistencia y participación en clases, conferencias y demás actividades extracurriculares que contribuyan con la formación de forma presencial, pero la mayoría de las universidades no presentan un sistema automatizado que lleve un registro y se evidencie que la identidad de la persona es la correcta.

Actualmente el registro de asistencia se hace con firmas en una hoja de papel que circula alrededor de la clase, o con el llamado a lista por parte del profesor, lo cual no es muy seguro y es bastante vulnerable, ya que es de fácil suplantación y cualquiera puede firmar por un compañero ausente, o en muchos casos el profesor opta por no registrar la asistencia para no gastar el tiempo de explicación del tema o realización de evaluaciones y talleres. De esta manera, se ve la necesidad de hacer uso de un sistema de información con autenticación, en donde se pueden mejorar aspectos como el tiempo que se invierte en un llamado a lista y la identificación plena del individuo en un examen.

En los diferentes eventos institucionales, se observa la importancia de la recolección de información de los asistentes y verificar su identidad, para lo cual se propone el uso de biometría, que está definido como un método basado en información, con el cual se pueden reconocer características físicas e incluso de comportamiento de una persona. Para esto es necesario implementar un mecanismo para hacer la captura inicial, también llamado proceso de enrolamiento, que se realiza para poder efectuar satisfactoriamente procesos de identificación y autenticación de personas, los cuales se basan en la realización de la comparación de las características en la que se analizan los rasgos y patrones para poder determinar si la información biométrica corresponde o no, a una persona en particular.

El proyecto sugiere el diseño de un sistema de autenticación multifactor, basado en el uso de factores de inherencia, en las instituciones de educación superior que permita verificar la asistencia a actividades académicas por parte de los estudiantes, y su integración con una solución de análisis de datos que le permita a la institución optimizar algunos procesos de toma de decisión como la asignación de recursos en términos de tiempo, planta física y personal, y verificar el grado de interés general a charlas y conferencias realizadas a los estudiantes para conocer el impacto que estas tienen en la institución. Por lo tanto, y con el fin de definir la mejor solución que integre herramientas de visualización de datos y los sistemas biométricos, se analizarán las diferentes metodologías y tecnologías que se han usado en proyectos que han abarcado problemáticas similares o afines.

En adelante el lector podrá conocer el motivo del uso de los sistemas biométricos en este proyecto, además se explicarán los beneficios del análisis de datos para poder visualizar la información relevante para la institución y tomar decisiones importantes al respecto, haciendo uso de inteligencia de negocios para unificar, transformar y visualizar los datos necesarios para que el sistema de información pueda ser de utilidad para la institución de educación superior.

2 Problemática (Justificación)

Hay muchos casos, como el registro de notas que ya se ha automatizado y se ha logrado obtener un mejor control de las mismas a través de las plataformas que tiene cada universidad, debido a que son datos más sensibles y cualquier modificación no autorizada de los mismos puede llegar a ser incluso perjudicial para la misma institución. Pero el registro de asistencia, no ha tenido ese mismo desarrollo y no se le ha dado la misma importancia, ya que se considera que es responsabilidad de cada estudiante asistir a las diferentes sesiones o eventos propuestos durante el semestre, y como se mencionó anteriormente,

al no tener un compromiso de los instructores por la cantidad de tiempo que toma hacer un chequeo manual, genera la inasistencia y desinterés por parte de los estudiantes que seguramente conllevará a perder el objetivo principal de una asignatura o evento académico, que es que los estudiantes puedan obtener conocimientos suficientes para poder utilizarlos en el desarrollo de su profesión y para beneficio de la sociedad.

Dicho esto, se puede evidenciar la oportunidad de crear un sistema que no sólo permite automatizar la participación de los estudiantes en una actividad académica, sino que a su vez pueda validar la identidad de cada individuo utilizando la biometría y poder así, evitar la suplantación de identidad cuando sea requerido por ejemplo en los exámenes y parciales. Si se utiliza un sistema como el anteriormente descrito, facilitará el trabajo de tomar asistencia a los instructores, ya que se realizará de una forma más rápida y esto hará que los estudiantes vuelvan a preocuparse por no faltar a clases, y podrán tener control de las personas que hagan parte de los eventos institucionales. (Trabelsi & Shuaib, 2011)

El sistema de información debe contar con el acceso a la información de los miembros de la comunidad académica, la cual debe ser suministrada por las instituciones de educación superior, esta información incluye, pero no se limita, a los datos básicos de las personas, en el caso de los estudiantes estos datos serían: nombres, programa al que pertenecen, historial académico, entre otros. Por lo anterior, es importante que, durante la selección de las tecnologías biométricas para plantear el sistema de control de asistencia, se tenga en cuenta la compatibilidad con los activos tecnológicos de la institución y que pueden llegar a estar integrados con el sistema, ya que este no solo necesitará acceso a las bases de datos para consultar información, sino que además requiere que se le garanticen permisos para el almacenamiento de la información biométrica de los estudiantes, recolectada durante el proceso de

enrolamiento, y los resultados de las autenticaciones durante la toma de asistencia.

Además de asegurar el mejor sistema biométrico que puede utilizarse en una institución y que asegure la calidad de los datos, es imprescindible una buena presentación de estos a los tomadores de decisiones, debido a que como menciona (Mohd Ali, Gupta, Lenka, & Krishna Nayak, 2016) para las personas es más fácil encontrar patrones en datos representados de manera visual que en datos sin procesar, para encontrar estos patrones y las correlaciones existentes entre los datos, es importante elegir cuidadosamente las dimensiones que se van a mostrar, ya que si se muestran todas la visualización puede ser muy densa y difícil de analizar y si se muestran muy pocas dimensiones se pueden estar perdiendo patrones y datos importantes para el negocio.

A partir de los datos con los que la institución de educación superior cuenta de cada estudiante, la información que se recolecta del sistema propuesto y la definición de la información que se quiere mostrar, el objetivo es obtener una bodega de datos que contenga toda la información relevante y necesaria para no perder los aspectos importantes para los tomadores de decisiones de la institución.

2.1.1 Pregunta de Investigación

¿Cuál es la mejor solución que se puede aplicar en una institución de educación superior que permita recolectar y visualizar la información obtenida en la asistencia a eventos institucionales, para los tomadores de decisiones, haciendo uso de un sistema biométrico?

3 Objetivos

3.1 Objetivo General

Definir una solución que realice la recolección y visualización de la información, obtenida mediante un sistema biométrico que garantice la correcta identificación de los estudiantes en las diferentes actividades dentro de una institución de educación superior y facilite la toma de decisiones.

3.2 Objetivos Específicos

- Determinar cuál es el mejor sistema biométrico en términos de costos, facilidad de implementación y precisión, para garantizar la identificación de los estudiantes y control de asistencia para diferentes tipos de eventos en una institución de educación superior.
- Recolectar, organizar y visualizar los datos del sistema y ponerlos disponibles para la toma de decisiones basado en cuatro escenarios que describen diferentes actividades que se realizan en las instituciones de educación superior.
- Detallar la implementación de la solución que integra el modelado de datos, los sistemas de visualización de datos y biometría para conocer los recursos necesarios, en cada escenario.

4. Alcance y Limitaciones

El alcance del proyecto es la definición de una solución de visualización de la información, utilizando un sistema biométrico, que pueda ser implementado en una institución de educación superior y para que pueda tanto tomar decisiones con los datos visualizados, al hacer el análisis de dicha información para beneficio propio.

Dado que no se tiene un contexto específico en el que una institución de educación superior pueda estar interesada en la implementación de este proyecto, se realizó un análisis de cuatro posibles escenarios, definiendo el tipo de biometría óptimo y la visualización de algunos datos que se pueden obtener, de acuerdo a la actividad en el que se tomen ese registro como la asistencia a clases, a eventos (conferencias, charlas, reuniones, etc.), a evaluaciones o el monitoreo de los estudiantes en un caso de evacuación. Como limitante se tiene el acceso a los datos de los estudiantes que pertenecen a la institución de educación superior, ya que, para obtener información puntual, se debe conocer con claridad si se cuenta con los datos básicos de la biografía, la biometría y con el perfil académico de cada uno de ellos.

5 Marco Teórico

A lo largo de este capítulo, se definirán los conceptos necesarios para el desarrollo de este proyecto que están enmarcados en tres temas principales: Inteligencia de negocios, análisis y visualización de datos y biometría. La descripción del funcionamiento y clases de sistemas biométricos harán parte del desarrollo y los anexos del documento.

Primero se hará una introducción a la inteligencia de negocios y se describirá cómo las tecnologías, procesos y aplicaciones pueden facilitar el uso de esos datos, generando conocimientos e información oportuna y confiable para responder a las situaciones que puedan presentarse en la institución educativa.

Posteriormente se hablará del análisis y visualización de los datos, se describirán algunas técnicas utilizadas actualmente y las ventajas que han brindado estos procesos en la toma de decisiones tanto en empresas como en instituciones educativas.

Por último, teniendo como base estos dos términos y considerando que para el desarrollo del proyecto se requiere un sistema de recolección de algunos datos de los asistentes a las actividades de la institución, se hará una introducción a la definición de biometría y por qué se ha convertido en una herramienta de gran importancia en la actualidad, además se aclaran algunos términos clave en los procesos y objetivos que tienen estos sistemas de identificación.

5.1 Inteligencia de negocios

En el mundo empresarial actual las organizaciones requieren estar en capacidad de adaptarse de manera ágil a los cambios en el mercado y en general cualquier situación que pueda afectar el negocio, por lo que para las

organizaciones es de gran importancia contar con herramientas que les permita conocer el estado de la organización, sus clientes y su entorno, de modo que quienes están encargados de tomar decisiones puedan hacerlo rápidamente y con la menor incertidumbre posible permitiéndole a la organización adaptarse o anticiparse a los cambios del mercado.

Una de estas herramientas es la implementación de soluciones de inteligencia de negocios, también conocida como BI, por sus siglas en inglés. Sobre estas soluciones varios autores han planteado diversas definiciones, por lo que a continuación se presentan dos de estas:

(Loshin, 2013) menciona la definición de inteligencia de negocios, del *Data Warehouse Institute*: “es la combinación de tecnología, herramientas y procesos que me permiten transformar los datos almacenados en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia comercial. La inteligencia de negocios debe ser parte de la estrategia empresarial, esta le permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y la capacidad de tomar buenas decisiones para así obtener mejores resultados”.

“Se entiende por *Business Intelligence* al conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización”(Curto Díaz & Conesa Caralt, 2010).

De acuerdo con las definiciones anteriores, se puede concluir que la inteligencia de negocios es un conjunto de procesos, metodologías, aplicaciones, tecnologías y mejores prácticas que permiten integrar los datos internos y externos del negocio depurados, transformados y consolidados en un repositorio para la generación de información y el perfilamiento de usuarios

que posteriormente se convertirá en explotación de la información a partir de la cual los usuarios podrán tomar decisiones sobre el negocio.

En este punto cabe resaltar que como lo menciona (Loshin, 2013) en una solución de inteligencia de negocio son más importantes los procesos y las personas que las herramientas tecnológicas que forman parte de la solución, esto debido a que son las personas las encargadas del análisis de los datos proporcionados por las herramientas y finalmente de la toma de las decisiones, pues el BI se encarga como tal de proporcionar el acceso a la información a los empleados de los diferentes niveles de la organización de modo que puedan identificar oportunidades y cumplir con sus labores de manera eficiente. Lo anterior conlleva al cumplimiento del objetivo principal de la inteligencia de negocios que es el aumento de la productividad en la organización.

5.1.1. Componentes de inteligencia de negocios

En la Figura 1 se pueden identificar los componentes que integran una solución de inteligencia de negocios.

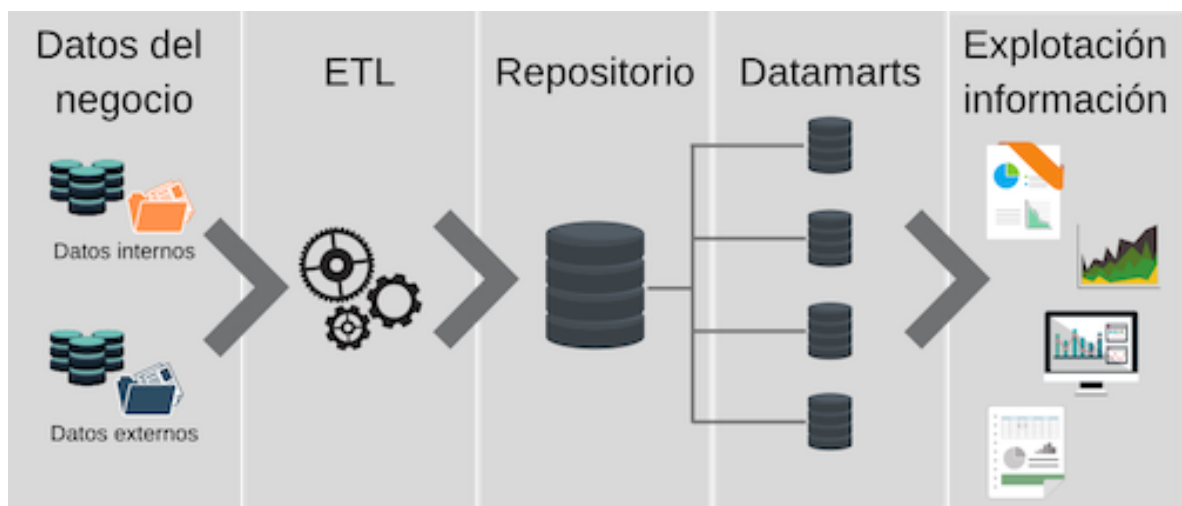


Figura 1 Arquitectura de una solución de negocios. Elaboración propia

5.1.1.1. Datos del negocio

En este componente se identifican las fuentes de datos internas y externas que va a utilizar la solución, a partir de los procesos de negocio que se quiere impactar. Estas fuentes pueden ser bases de datos, archivos generados por empleados de la organización en cualquier formato y los sistemas de información entre los que se encuentran los sistemas de soporte (ERP, CRM, entre otros), sistemas operacionales y los sistemas legados.

En el caso del proyecto de investigación las fuentes de datos del sistema de análisis provendrán de: sistemas propios de la institución en donde se encuentran los datos relacionados con los estudiantes, el personal, las actividades y la infraestructura de la institución, y del sistema biométrico con el cual se hará la recolección de datos relacionados con la asistencia a las actividades.

5.2 Proceso de extracción, transformación y carga (ETL)

El proceso de extracción, transformación y carga, también conocido como ETL por sus siglas en inglés, es el componente que permite tener los datos disponibles en un repositorio para su posterior uso. Este proceso como su nombre lo indica consiste en tres pasos:

- Extracción: Identificación de los datos que se deben extraer de las diferentes de datos, y la recolección de estos datos.
- Transformación: En este paso los datos se mueven a un área llamada *data staging area* en donde estos son organizados y se les aplican procesos de calidad de datos de acuerdo con las reglas particulares del negocio, de modo que los datos queden con una estructura apropiada para analizarlos y generar reportes

- Carga: Es el paso en el que los datos transformados son ingresados en el repositorio de destino.

5.2.1 Repositorio

Hace referencia al componente o la herramienta en la cual se almacenan los datos del negocio que ya están transformados para su análisis, y que se encuentra en un ambiente distinto al de los sistemas operacionales. Por lo general en la industria en esta etapa se usa un *Datawarehouse*, que es una base de datos corporativa que permite integrar la información proveniente de diferentes fuentes de datos, para que esta sea procesada y pueda ser analizada desde diferentes perspectivas. Este tipo de base de datos permite que la información sea consultada más rápido que en una base de datos relacional, esto principalmente por el tipo de estructuras que usa. (Sinnexus, n.d.-b)

5.2.2 Datamart

En este componente se hace el perfilamiento de los usuarios de la solución, de modo que cada usuario pueda ver la información que requiere según el rol que desempeñe en la empresa, esto debido a que como se mencionó anteriormente es indispensable que cada empleado tenga acceso a la información que le permita cumplir con sus labores de manera eficiente.

Para esto comúnmente se usan los *Datamarts*, que son bases de datos especializadas para el almacenamiento de la información de un grupo de usuarios en particular (grupos de roles, áreas de negocio, etc...) y permite hacer el análisis de esta información desde las perspectivas que afectan los procesos de negocio de interés o relacionados con ese grupo de usuarios. (Sinnexus, n.d.-a)

5.3 Explotación de la información

La explotación de la información es el componente final de una solución de inteligencia de negocios, en la que se hace la visualización de la información a través de aplicaciones, reportes, cuadros de mando, y en general cualquier artefacto o herramienta que le permita a los usuarios interactuar con la información para apoyar la toma de decisiones a partir de la identificación oportunidades y riesgos del negocio.

5.4 Visualización, interpretación y análisis de datos

En la actualidad las empresas cuentan con un gran volumen de datos de diferentes tipos relacionados con el negocio, sus sistemas transaccionales, sus plataformas digitales y sus clientes, sin embargo, para conocerlos, entenderlos y explotarlos para la toma de decisiones, se requiere que la empresa esté en la capacidad no solo de procesarlos sino que se puedan transformar en conocimiento al interpretarlos, a través de mecanismos de explotación de datos como la visualización de datos esto ya que como lo menciona (Mohd Ali et al., 2016) la manera más fácil para encontrar patrones y correlaciones entre datos e interpretar los datos es a través de representaciones gráficas que presenten los datos y sus dimensiones relevantes para el negocio.

Tal como lo definen en el informe Visualización de datos: definición, tecnología y herramientas (Iniciativa Aporta, 2016) la visualización de datos es la representación gráfica de información que tiene como propósito la interpretación y construcción de conocimiento a partir de los datos y la comunicación de la interpretación realizada a un conjunto de datos. Cabe resaltar que, aunque en la visualización de datos se representan generalmente variables cuantitativas y las correlaciones existentes entre ellas, es posible

utilizarse también para representar relaciones entre variables cuantitativas a través de gráficas como los grafos.

Hoy en día existe una gran cantidad de herramientas que permiten construir conocimiento a partir de la visualización de datos a través de distintos tipos de gráficas, para los datos estructurados especialmente se adoptaron la clasificación de tipos de visualización propuesta por (Iniciativa Aporta, 2016) que se lista a continuación:

- Elementos básicos de representación de datos: Este es el caso más sencillo de visualización. Algunos ejemplos de este tipo de visualización son las gráficas (Barras, líneas, columnas, puntos, etc.) y los mapas de burbujas o los mapas de calor.
- Cuadros de mando: Utilizados principalmente en las empresas para el análisis de un conjunto de datos y la toma de decisiones, los cuadros de mando están compuestos por visualizaciones de elementos básicos de datos que tienen relación entre ellos.
- Infografías: Este tipo de visualización permite hacer una construcción narrativa a partir de un conjunto de datos, por lo tanto, son construidas a partir de la combinación de elementos visuales como leyendas, símbolos, dibujos, etc. Elementos que permiten contar una historia y que a diferencia de los cuadros de mando no se enfocan en proporcionar al usuario elementos que permitan hacer un análisis de las variables presentadas.

Para los procesos de visualización de datos no solo es importante conocer los diferentes tipos de gráficas que existen y cómo se usan, sino que es importante considerar algunos principios que permitan presentar los datos de manera adecuada, (DataCentric, 2017) lista los siguientes principios en su artículo

¿Qué es el *Data Visualization* o cómo interpretar grandes cantidades de datos?

- Simplificar: presentar la cantidad de variables adecuadas.
- Comparar: para facilitar la comprensión y evitar errores de memoria, es importante poder visualizar todos los datos en una imagen.
- Dónde y cómo mostrar la información que se quiere destacar para captar la atención.
- Diversificar: diferentes representaciones de los mismos datos pueden llevar a nuevas conclusiones
- Buscar el por qué: la meta de la visualización de datos no es encontrar qué ocurre, si no dar una forma visual fácilmente interpretable a la mirada de datos de los que dispone una compañía

Una parte importante de la explotación de datos es la interpretación y análisis de los mismos, esto se puede realizar a través de diferentes herramientas que le permiten al usuario generar sus reportes y/o cuadros de mando y adicionalmente ofrecen la opción de encontrar *insights* de los datos y la realización de cálculos avanzados con el fin de facilitar su análisis. En este punto cabe resaltar que para el proceso de análisis de los datos existe una escala de valor que le permite a la organización encaminar sus esfuerzos en la implementación de tecnologías y decisiones estratégicas que le permitan optimizar sus procesos y los servicios que prestan a sus clientes. Como se puede ver en la figura 3 los primeros niveles la analítica descriptiva y la analítica diagnóstica le proporcionan a la organización la información necesaria para saber exactamente lo que sucede de modo que se puedan tomar las decisiones necesarias para garantizar el buen rendimiento en sus procesos y el servicio que prestan a sus clientes. Mientras que en los

siguientes niveles el enfoque es la predicción de eventos y cómo lograr que estos ocurran de manera que la organización se vea beneficiada bien sea adquiriendo una ventaja sobre sus competidores o mitigando un riesgo del negocio.

Analytic Value Escalator

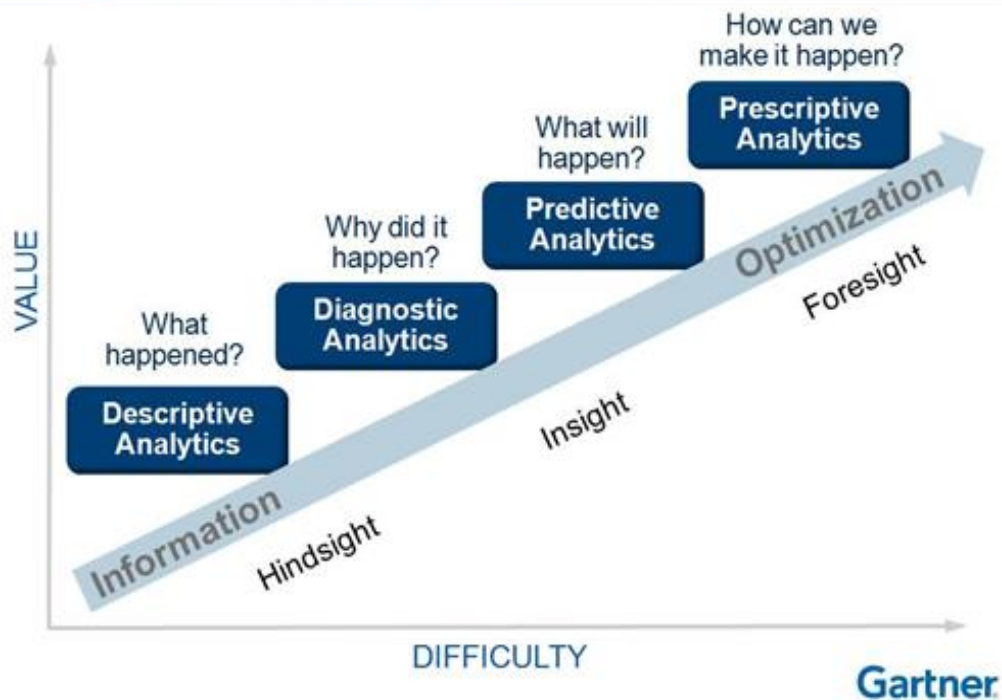


Figura 2 Escala de valor dentro del análisis de la Información. Fuente: Gartner.

5.5 Big Data

Si bien el concepto de *Big Data* es relativamente reciente, la necesidad de la humanidad por almacenar diferentes tipos de datos que les permitiera llevar el control de sus negocios y de la población, así como el deseo de predecir el futuro para adelantarse a una situación en particular y tomar una decisión acertada, viene de años e incluso varios siglos atrás; a lo largo del tiempo se han desarrollado diferentes técnicas y herramientas tecnológicas que permitan cuantificar, codificar, almacenar información y predecir eventos.

El concepto de *Big Data* hace referencia a un conjunto de herramientas y técnicas para el análisis de información, que permiten solucionar problemas que generalmente involucran un gran volumen de datos, en diferentes formatos, provenientes de diferentes fuentes, que cambian a una gran velocidad y que los sistemas tradicionales no podrían abordar de manera eficiente. Estas herramientas son aplicables en cualquier modelo de negocio y por lo general están asociadas a temas de mercadeo y analítica digital, está última principalmente para optimizar los canales digitales de las compañías; la retención de clientes, y el reconocimiento de nuevas oportunidades de negocio, pues se logra perfilar los clientes y mejorar las promociones y productos o servicios que se les ofrecen.

Actualmente se tiene la concepción errónea de que el éxito del *Big Data* está en el volumen de información, en la capacidad que nos brinda para cuantificar cosas, y se cree que la información que esta herramienta muestra es la verdad absoluta y objetiva de una situación en particular. El *Big Data* no solo es aplicable a cualquier modelo de negocio, sino que no depende del tamaño de las compañías ni del tamaño de sus conjuntos de datos, está directamente relacionado con la forma en la que se interpretan y desechan datos, bien sea por consideraciones propias de quien recolecta y analiza la información o porque los datos no son producto de modelos cuantitativos o no se ajustan a un modelo en particular. En conclusión, el éxito de un proyecto en el que se involucre *Big Data* no radica en el volumen de datos con el que se cuente, sino en la calidad de los datos pues lo que para una compañía puede ser un conjunto de datos grande para otra puede representar una cifra insignificante.

Además, se debe tener en cuenta que para obtener datos de calidad que permita a un gerente tomar una decisión adecuada frente a una situación en específico, se deben validar las fuentes de las que se está obteniendo la información de modo que se reduzca la probabilidad de que esta sea alterada por las personas, hacer revisión de los algoritmos que se están aplicando,

validar los datos con los que se están entrenando estos algoritmos y a lo largo del proceso tener en cuenta diferentes perspectivas del problema, para esto, como lo sugiere (Wang, 2016), es importante contar con personal calificado en soluciones de *Big Data* y etnógrafos o investigadores que puedan recolectar *Thick data* (datos densos) que hacen referencia a datos humanos, tales como: historias, emociones, interacciones, datos que no se pueden cuantificar de modo que los datos cualitativos se puedan incluir en los modelos cuantitativos y así obtener mejores resultados.

5.6 Biometría

A finales del siglo XIX con el descubrimiento del carácter distintivo de las huellas digitales varias entidades legales iniciaron la recopilación de las huellas de delincuentes en una base de datos de modo que si estas aparecían en la escena de un crimen se pudiera reconocer al criminal a través de la comparación de las huellas obtenidas en la escena y las guardadas previamente en la base de datos. Es gracias a esto que los procesos biométricos han ido ganando popularidad, al punto que en la actualidad son usados en cualquier tipo de proceso que requiera hacer el reconocimiento de una persona en la mayoría de los casos a través de sus características corporales como lo son los rasgos faciales, la voz, las huellas dactilares, entre otros.

La biometría ha ido evolucionando y en la actualidad no solamente se consideran las características fisiológicas de una persona para hacer el reconocimiento biométrico de la misma, sino que se usan también características de la conducta de la persona, sin embargo, tal como menciona (Jain, Ross, & Prabhakar, 2004) estas características deben cumplir con los siguientes requisitos:

- Universalidad: cada persona debe tener la característica.

- Carácter distintivo: cualquiera de las dos personas debe ser lo suficientemente diferente en términos de la característica.
- Permanencia: la característica debe ser suficientemente invariante (con respecto al criterio de coincidencia) durante un período de tiempo.
- Coleccionabilidad: la característica puede medirse cuantitativamente.

En general cuando se habla de sistemas biométricos se hace referencia a un sistema que es capaz de reconocer o verificar la identidad de una persona a través de una característica corporal o de comportamiento específica de una persona. Para realizar el reconocimiento o la verificación de identidad el sistema inicialmente hace un proceso de extracción y almacenamiento de los patrones encontrados en la característica que se va a registrar, este proceso es conocido como registro o enrolamiento y finalmente durante la autenticación se realiza nuevamente un proceso de extracción del patrón y se compara con el patrón almacenado en la base de datos. Este proceso entrega un resultado que indica el porcentaje de coincidencia entre los patrones, lo cual permite definir si una persona fue o no reconocida.

Algunos de los requisitos con los que debe cumplir el sistema biométrico son los siguientes: Precisión de reconocimiento, velocidad y recursos especificados, ser inofensivo para los usuarios, ser aceptado por la población a la que va dirigido y ser lo suficientemente robusto frente a varios métodos fraudulentos y ataques. (Jain et al., 2004)

Por tanto, la biometría se refiere a todos los procesos utilizados para reconocer, autenticar e identificar a una persona según características físicas o comportamentales, de acuerdo a plantillas o patrones que son únicos e irrepetibles, que se encuentran almacenados en bases de datos.

5.6.1 Enrolamiento

Es la captura de información biométrica que es almacenada en una base de datos, para que pueda ser utilizada al momento de una identificación o autenticación. En un sistema de seguridad biométrica es el proceso inicial de recopilación de muestras de datos biométricos de una persona, que se almacena en una base de datos, creando una plantilla de referencia que representa la identidad de un usuario para su posterior comparación. (C.O, A.O, O.O, & E.O, 2013)

5.6.2 Autenticación

El término autenticación en referencia a biometría es el proceso por el cual se comprueba que una persona es la que dice ser. De acuerdo a la compañía internacional francesa Morpho (IDEMIA), consiste en verificar que los datos biométricos, que se registran son los mismos datos que se encuentran registrados en las bases de datos (IDEMIA, n.d.)

5.6.3 Identificación

Por otra parte la identificación según Morpho (IDEMIA, n.d.), es el proceso en el cual se realiza la comparación de la información biométrica de una persona contra toda la almacenada en una base de datos para identificar si ya se encuentra registrada.

5.6.4 Matcher

Un matcher consiste en tomar una muestra biométrica, para compararla con un patrón o plantilla almacenado en una base de datos con el fin de autenticar o no a una persona.

Para tener una definición más exacta de lo que es un matcher, se recurre a la definición dada por Technovelgy, donde se considera que es la coincidencia

biométrica, que puede consistir en una puntuación y designa el grado de similitud de una muestra biométrica y una plantilla de referencia almacenada en una base de datos. (Technovelgy, n.d.)

5.6.5 Minucia

“La minucia es un punto de interés de la huella, que se caracteriza por su posición (x, y) y un ángulo ya sea de una terminación o bifurcación. Una minucia de terminación es donde la cresta termina y no continúa en otro lado y de bifurcación es donde una cresta se divide” (Rosales Cruz, 2001)

De esta manera se entiende las minucias son puntos donde los bordes terminan o se dividen y es lo que permite dar a una huella dactilar características únicas que permiten identificar a un usuario.

6 Estado del Arte

A continuación, se describe el estado actual de soluciones similares a las planteadas en el proyecto, en términos de análisis y recolección de datos.

En el caso de la recolección de datos se contextualiza, clasifica y categoriza los diferentes sistemas biométricos que han sido probados en otras instituciones para este propósito. Mientras que, para el análisis de datos, se evalúan los proyectos desde el punto de vista de la explotación de los datos, los usos y los beneficios que se han encontrado en este tipo de proyectos en instituciones educativas, así como los datos y aspectos a tener en cuenta dentro del análisis de los datos.

6.1 Análisis y visualización de datos

Durante la revisión de soluciones similares a la que se plantea en este documento, se encontró que en su mayoría las soluciones que se hacen para automatizar el proceso de registro de asistencia, se centran en resolver el problema de garantizar la identidad de la persona que intenta ingresar a un evento, agilizar el proceso de toma de asistencia, minimizar los errores humanos en el momento de transferir los datos de asistencia a un sistema principal y en reducir las maneras en las que los estudiantes vulneran los sistemas manuales de registro de asistencia bien sea haciéndose pasar por otro estudiante o asistiendo sólo una parte del tiempo al evento.

Este es el caso del sistema de asistencia presentado por (Abas, Tuck, & Dahlui, 2014) en donde describen un escenario en el que la asistencia a las clases es tomada de manera manual, lo que hace que para los maestros sea difícil llevar el registro de asistentes cuando los grupos son muy grandes y no tengan manera de saber si los estudiantes se quedan hasta que finaliza la clase; otro problema que identifican en el proceso es que la información de asistencia después de ser recolectada debe ser ingresada en un sistema

principal de manera manual, lo cual hace el proceso sensible a errores humanos. A lo largo del documento los autores explican con detalle la manera en la que el sistema atacará el problema, a través de la definición de los flujos que el sistema de asistencia que van a desarrollar seguirá y las tecnologías que van a usar en la implementación del sistema, el cual además de registrar los datos de asistencia y garantizar que los asistentes a clase permanecen dentro del salón durante la sesión proporciona una interfaz gráfica en donde se presentan los resultados que obtuvo el sistema, entre los datos que se muestran se encuentran: Código del estudiante, fecha y hora, estado que indica sí asistió o no a la clase, entre otros. A pesar de que el sistema está en capacidad de mostrar esos datos en forma de tablas y gráficas el documento no enfoca sus conclusiones en la explotación de estos artefactos por parte de los usuarios lo que impide saber los posibles usos y beneficios que tiene para una institución este tipo de reportes.

El proyecto presentado por (Sulaiman, Bakar, Noor, & Abdullah, 2015) por su parte expone el caso en el que la institución ofrece a los estudiantes actividades enfocadas en mejorar las habilidades blandas de los estudiantes, sin embargo el método que usan para tomar la asistencia es poco efectivo y se presentan casos de fraude en los que estudiantes que no asisten a las actividades aparecen registrados como asistentes, el sistema además de recopilar los datos de asistencia que son tomados al inicio y al finalizar la actividad proporciona una funcionalidad para el análisis de los datos, en el que se puede validar la cantidad de estudiantes que asistieron y se quedaron durante toda la actividad, la opción de listar los estudiantes ausentes de la actividad y los datos pueden ser exportados a Excel. Sin embargo, al igual que el sistema de asistencia descrito anteriormente el documento donde se describe el proyecto y sus conclusiones no se enfocan en describir los usos que les dan a los datos que recolectan y a los reportes que el sistema permite.

En el artículo *“Improving the quantity and quality of attendance data to enhance student retention”* (Bowen, Price, Lloyd, & Thomas, 2005) los autores describen las mejoras realizadas a un sistema de asistencia manual a través de la implementación y el uso de un sistema electrónico de monitoreo de asistencia llamado Uni-Nanny® el cual permitió eliminar errores humanos en los registros de asistencia haciendo que se lograra capturar más datos y que estos fueran consistentes, a diferencia de los dos sistemas descritos anteriormente este documento en sus conclusiones habla sobre el uso de los datos para validar el rendimiento de los estudiantes respecto a su asistencia y la posibilidad de acceso a los datos tomados por el software Uni-Nanny® por parte de diferentes miembros del personal de la institución para definir el uso de espacios y gestionar recursos. Una parte interesante del artículo es que se tuvo en cuenta los requerimientos del personal de la institución y de sus estudiantes de modo que el proceso de captura de asistencia fuera cómodo para los estudiantes y el personal obtuviera la información que requería para cumplir con sus labores, además en las conclusiones se hace referencia a la acogida que tuvo el sistema por parte de los estudiantes y resaltan tres motivos para que esto sucediera:

- Los estudiantes no quieren o no saben qué deben hacer para solicitar ayuda, este sistema de monitoreo de asistencia podría ayudar a iniciar servicios que le brinden soporte al estudiante.
- El sistema destacó problemas de la institución que pueden ocasionar que los estudiantes se retiren, y permite a la institución abrir un canal de comunicación permanente con los estudiantes en los que se reciba realimentación sobre las razones que tienen para no asistir a las clases.
- El sistema provee a los estudiantes una motivación adicional para asistir a los eventos.

Otro aspecto importante que menciona este artículo es el tratamiento que se le debe dar a esta información, que debe ser tratada como información sensible y confidencial que ha de ser usada para apoyar a los estudiantes en lugar de usarlo para vigilar o controlarlos. En relación con lo anterior y sumado al enfoque que tiene el proyecto que es encontrar la manera en la que una educación de institución superior puede aprovechar el análisis de los datos de asistencia de los estudiantes se hizo la revisión de algunas investigaciones orientadas a determinar si existe una relación entre el rendimiento académico de los estudiantes y el porcentaje de asistencia de estos a las clases, en su artículo *“Student Attendance: Is It Important, and What Do Students Think?”* (Muir, 2009) concluye que como resultado de su investigación que los estudiantes que asisten con regularidad a clase obtengan mejores notas, lo cual mencionan que es acorde con las investigaciones realizadas en el ámbito de la enseñanza y el estudio en instituciones de educación superior en los que se evidencia la relación entre la asistencia de los estudiantes a las clases y la motivación que tienen para hacerlo, encontrando que quienes están más motivados tienen mejores niveles de asistencia.

Basado en la revisión realizada a los artículos de los autores (Muir, 2009) y (Levshankova, Hirons, Kirton, Knighting, & Jinks, 2018) se encontró que para el análisis de los asistencia es importante que además de contar con los datos de asistencia el sistema tenga acceso al sexo, la edad y las calificaciones de los estudiantes de modo que se puedan encontrar patrones que ayuden a identificar el comportamiento de una población en particular, en este sentido ambos autores resaltan que además de la extracción de los datos es importante involucrar a los estudiantes de modo que se pueda identificar la razón por la que se presentan las inasistencias a las clases, en el caso de los dos artículos encontraron que en su mayoría ocurrían cuando el estudiante estaba enfermo, sin embargo esta razón puede variar en el contexto de las instituciones de educación superior colombianas. La identificación de la razón

de las inasistencias es importante debido a que al conocer la razón por la que un grupo de estudiantes no asisten a una clase o un estudiante deja de asistir a la mayoría de clases que ha inscrito para cursar durante el semestre, la institución puede mejorar su oferta académica y establecer o mejorar sus canales de comunicación con los estudiantes de modo que se pueda dar contexto a los resultados que aparecen en los reportes y la institución pueda hacer una mejora constante en sus programas de acuerdo a las necesidades identificadas con sus usuarios.

6.2 Biometría

En el artículo publicado por (Lukas, Mitra, Desanti, & Krisnadi, 2016) en la conferencia internacional de convergencia de tecnologías de información y comunicación ICTC, por sus siglas en inglés, se describe el desarrollo de técnicas para la detección facial de los estudiantes en un salón de clase. Este es un sistema biométrico bastante común de utilizar, ya que actualmente se encuentran en el mercado mejores cámaras, incluso en dispositivos móviles, en los que la autenticación facial se puede realizar a través de la implementación de aplicaciones móviles que capturen la información biométrica del estudiante, y la cantidad de algoritmos de código abierto que se encuentran en el mercado capaces de recolectar una buena cantidad de información biométrica, en este caso del rostro, lo cual permite que la precisión de las validaciones sea bastante buena. Sin embargo, para obtener resultados óptimos este tipo de biometría requiere que, tanto para el enrolamiento como para la validación, el usuario cuente con unas condiciones de ambiente especiales que garanticen la calidad de los datos biométricos que se recolectarán. Algunas de estas condiciones son: las condiciones de luz (intensidad, contraluz, escasez de luz...), la calidad de la cámara del dispositivo, la posición de la cámara respecto al rostro (cercanía al dispositivo, rotación de 15 grados o más...), el uso de accesorios como lentes o sombreros, entre otras. Este es un aspecto bastante importante a tener en

cuenta ya que muchas de estas condiciones son difíciles de controlar en las aulas.

Otro sistema biométrico que es usado comúnmente para la autenticación es el reconocimiento de voz, el cual es presentado como una opción económica por (Yang, Song, Ren, & Huang, 2016) en el artículo "*An automated student attendance tracking system based on voiceprint and location*", debido a que no se necesita de dispositivos de autenticación especializados, sin embargo, se necesita un software para el tratamiento de señales bastante robusto. En este documento además se presenta un sistema de reconocimiento de voz en los celulares de los estudiantes, lo cual puede beneficiar el proyecto en términos de costos, ya que la institución no incurriría en gastos que involucren la adquisición de dispositivo. De igual manera se encontraron aspectos poco convenientes, como el hecho de que existe la posibilidad de que no siempre los estudiantes cuenten con un dispositivo móvil con el que puedan realizar las autenticaciones; los costos de desarrollo y mantenimiento de una aplicación móvil que soporte al menos los dos sistemas operativos de dispositivos móviles más comunes en el mercado (iOS y Android); y el impacto de la calidad del micrófono de los dispositivos móviles en el proceso de reconocimiento, lo cual implicaría calibrar el motor de reconocimiento de voz, ya que por lo general estos son más fáciles de calibrar para el idioma inglés y por lo tanto, son más precisos en ese idioma.

Por otra parte, se encuentran los lectores de huella digitales, que en la actualidad es el sistema de biometría más usado en las organizaciones independientemente del sector al que pertenezcan, generalmente en procesos de control de acceso, en el cual puede llevarse el registro de asistencia a clases por parte de los estudiantes. Una de las grandes ventajas que ofrece esta tecnología, es la precisión que brinda pues se estima que se pueden obtener resultados de más del 95% de confiabilidad, aunque al igual que en los otros tipos de biométricos mencionados con anterioridad, se requieren

algunas consideraciones durante la implementación del sistema como restricciones de edad, y las enfermedades de la piel a la hora de realizar las capturas dactilares, tal como lo menciona (Adeniji, Scott, & Phumzile, 2016). Haciendo uso del enrolamiento de huellas de los estudiantes y módulos transmisores que realizan la comparación de la huella capturada con los registros, en el artículo *“Biometric Attendance”* (Said et al., 2014) se muestran ejemplos de resultados en interfaces para proveer datos a los tomadores de decisiones, lo cual brinda flexibilidad para la integración de una solución de análisis de datos que a su vez permita la explotación de los datos biométricos recolectados por el sistema, y junto con otros datos de la institución educativa, se logre identificar patrones en la asistencia de los estudiantes, por programa, semestre y asignatura que permita tomar acciones frente a situaciones problema como la deserción en las instituciones educativas, optimizar el uso de recursos para la asignación de salones y profesores, y así ofrecer a los estudiantes un mejor servicio.

En el artículo *“Design and Implementation of Iris Recognition Based Attendance Management System”* (Khatun, Haque, Ahmed, & Rahman, 2015) se aborda el control de asistencia de los estudiantes a través del reconocimiento del iris, el cual es considerado el proceso biométrico más confiable, preciso y eficiente debido a las características internas del iris. Al igual que para el proceso biométrico de reconocimiento de voz, no se necesitan dispositivos biométricos avanzados, pero si algoritmos complejos para, en este caso, el tratamiento de imágenes y la detección de sus características. A pesar de que los dispositivos requeridos para este sistema no son considerados como avanzados, si son especializados y tienen costos bastante elevados que en el caso del proyecto tendrían que ser asumidos por la institución educativa. Adicionalmente tanto el proceso de enrolamiento como de autenticación, es bastante invasivo y tendría que hacerse estudiante por

estudiante lo cual resulta poco conveniente para dar solución a la problemática que desarrolla el proyecto de investigación.

Finalmente se encontraron otros proyectos en donde a través de la implementación de tecnologías adicionales a los sistemas de biometría, se busca llevar control de la asistencia de los estudiantes asegurando la identidad de estos, es el caso del documento "*Student attendance monitoring at the university using NFC*" (Benyo, Sódor, Doktor, & Fördös, 2012) en el que se presenta una tecnología no tan novedosa como el lector NFC para el uso de tarjetas compatibles, que contengan almacenadas las huellas del estudiante, de manera que se puedan hacer verificaciones cruzadas tanto de las huellas dactilares, como del documento de identificación del estudiante para garantizar su plena y correcta identificación. En el caso de (Eze Peter, Joe-Uzuegbu, Uzoechi, & Opara, 2013) se encontró que describen un sistema de validación biométrica muy similar al mencionado anteriormente, el cual se diferencia esencialmente en el hecho de que proponen el uso de una mayor cantidad de datos del estudiante y de la inscripción para realizar un monitoreo constante (en tiempo real) en los sistemas de identificación y verificación.

Para la búsqueda de la información y beneficios de cada uno de los sistemas biométricos que se pueden implementar en una institución educativa, se tuvieron en cuenta algunas características claves a la hora de elegirlos; que indiquen el nivel de asertividad que posee aproximadamente el tipo de biometría implementado, si se diseñó o si se requiere el desarrollo de un software complejo para poder utilizar el dispositivo biométrico respectivo, que se especifiquen las restricciones que posee cada tipo de biometría, y finalmente un ejemplo del diseño que están planteando y los resultados que se obtuvieron. Ya que el desarrollo de la problemática de cada uno de los artículos científicos se enfoca en una institución educativa para verificar la asistencia de los estudiantes, es clave poder conocer y comparar los resultados de cada artículo para verificar la viabilidad de la implementación de

cada sistema biométrico o una mezcla de los mismos, teniendo en cuenta características adicionales a las ya mencionadas como la movilidad que debe emplear cada estudiante para la autenticación, precios o conocer si la institución educativa ya cuenta con los dispositivos biométricos necesarios.

7 Metodología

El proyecto de grado está enfocado en el logro de tres objetivos específicos que le permitieron al grupo de trabajo proponer una solución. Para lograrlo fue indispensable tener una comunicación constante con un representante de una institución de educación superior que permitió la identificación de los casos de uso o escenarios en los que la implementación de este sistema aporta valor; y el desarrollo de un cronograma en el que se definieron las fases y las actividades necesarias para la consecución de estos objetivos, evidenciadas en el Anexo No.1: Cronograma. Como se evidencia en la Figura 1, el proyecto se realizó en 4 fases que se describen a continuación:



Figura 3 Metodología. Elaboración propia

7.1 Entendimiento y búsqueda de información

El desarrollo de esta investigación inicia con la definición, entendimiento y búsqueda de información científica de la problemática identificada, para esta fase se utilizaron los métodos de búsqueda de bibliometría en las bases de datos que facilita la Escuela Colombiana de Ingeniería, SCOPUS; y el uso de SCIMAGO, que permitió encontrar un ranking de revistas científicas y su importancia en la comunidad. A partir de la información recolectada de las bases de datos, se realizó la construcción del estado del arte que posteriormente facilitó la definición del sistema a desarrollar.

En esta fase se buscó tener la mayor información sobre los diferentes sistemas biométricos y las herramientas para visualizar y analizar datos que se han implementado o probado en otras instituciones educativas y que se pueden tomar de base para identificar y proponer un mejor análisis de esos datos.

7.2 Análisis y definición del diseño y desarrollo

Teniendo como base la información recogida en la fase de entendimiento y búsqueda bibliográfica, junto con la experiencia y conocimientos de los integrantes, se exponen cada una de las biometrías que pueden ser utilizadas en una institución educativa y que se pueden conseguir fácilmente en el mercado, para luego, hacer un análisis de ventajas y desventajas, en donde se tiene en cuenta el costo de las licencias, los tiempos de captura, la vulnerabilidad y asertividad en la identificación de los estudiantes. Dado que cada sistema biométrico tiene un mejor desempeño en ciertas situaciones, se plantean 4 escenarios aplicados a una institución educativa en donde se sugiere un sistema de información.

En la visualización de información se debe implementar un proceso de ETL para almacenar la información con la que ya cuenta la institución que es requerida en el proceso, junto con el tipo de datos que recolecta el sistema

biométrico, y poder construir una bodega de datos completa, para luego mostrarle la información relevante a los tomadores de decisiones y puedan efectuar un respectivo análisis de la misma.

7.3 Posibles soluciones

En esta fase se hizo una muestra de los sistemas biométricos escogidos, demostrando sus características, el funcionamiento y los datos que se pueden obtener de los mismos, para posteriormente realizar un ejemplo de un posible análisis de la información y las decisiones que se pueden tomar a partir de este.

El objetivo de esta fase es dar a conocer cómo sería el proceso que cada estudiante tendría que hacer al momento de autenticarse, cuánto tiempo tomaría la captura de biometría, qué datos se van a capturar y qué se puede hacer con la información recolectada, en cada uno de los 4 escenarios propuestos.

7.4 Resultados

En esta fase se da el cierre de la investigación a través de los resultados obtenidos durante todo el proceso, se expondrán todos los beneficios tanto del sistema de recolección de datos biométricos para la asistencia, como del análisis de la información; se mostrará la forma en que se pueden implementar, los costos y los requerimientos físicos y tecnológicos necesarios para el funcionamiento.

Además, se darán las conclusiones y recomendaciones necesarias para continuar o implementar el proyecto en alguna institución de educación superior.

8 Desarrollo

La definición del sistema de análisis de información y de autenticación, se realizó en cuatro escenarios posibles en los que la implementación puede aportar valor a la institución de educación superior, estos escenarios son la asistencia a clases, exámenes, conferencias o charlas y simulacros de evacuación.

Antes de exponer cada escenario se realiza una descripción detallada de la arquitectura del sistema, y posteriormente para sugerir un sistema de información que pueda ser implementado en cada escenario, se realizó una comparación detallada para conocer con certeza las ventajas y desventajas que puede otorgar las herramientas de visualización de datos y de cada tipo de biometría.

8.1 Arquitectura del sistema

En los contextos planteados las propuestas están enmarcadas en la arquitectura del sistema que se muestra en la Figura 4 y que se explica a continuación:

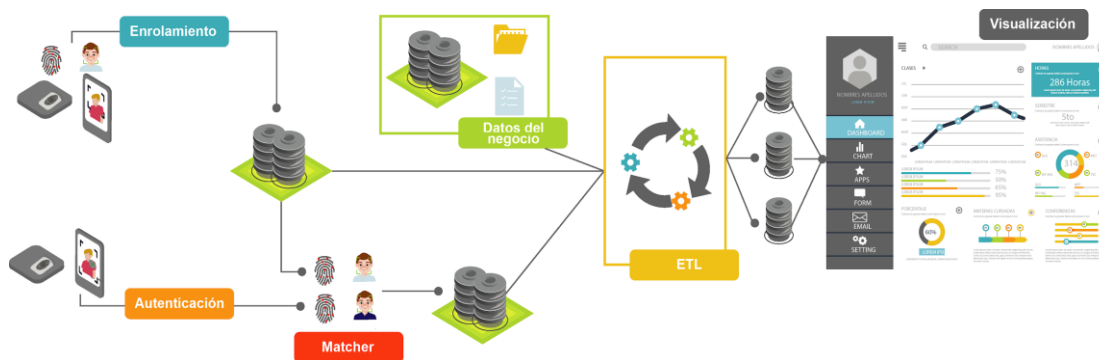


Figura 4 Arquitectura de la solución. Elaboración Propia

8.1.1 Módulo de Enrolamiento

Este es uno de los pasos en los que el proyecto de investigación debe tener especial atención, teniendo en cuenta que el éxito de la precisión durante el proceso de autenticación del sistema está basado en la calidad de la información biométrica recolectada durante los enrolamientos, y que la propuesta es el diseño de un sistema de autenticación multifactor. Para ello se debe gestionar una logística por parte de la institución de educación superior, en la que se encarguen de ambientar espacios que cumplan con todos los requisitos para tener información biométrica de calidad, por lo cual se deben seguir todos los lineamientos especificados por las normas regulatorias y las recomendaciones de los fabricantes de los dispositivos.

Una vez hecha la extracción de la minucia en el enrolamiento, dicha información debe ser almacenada en la base de datos junto con la información de la persona que se va a registrar, motivo por el cual, este proceso siempre debe ser ejecutado por un administrador de la institución de educación superior que debe tener presente que está tratando con información sensible de cada estudiante.

Siempre es recomendable al realizar un enrolamiento, realizar más de una captura para poder comparar cuál de ellas tiene mejor calidad y de cuál se puede obtener más minucias para luego tener una mejor comparación al momento de realizar la autenticación.

El sistema de enrolamiento debe poder contar con una función que no permita que se registren biometrías duplicadas, es decir, que en la base de datos se registre más de una persona con la misma información biométrica. Para poder realizar este proceso de manera correcta, se debe implementar un método de identificación para validar si el registro de biometría que se desea enrolar ya existe en la base de datos.

8.1.2 Módulo de Autenticación

Este es el módulo base de toda la solución y al igual que en el enrolamiento, las capturas biométricas deben seguir los lineamientos indicados para obtener una respuesta óptima e identificar a cada persona que intente acceder al sistema a registrar su asistencia.

Cada persona que desee ingresar a un aula de clase o a un evento de la institución de educación superior, deberá ingresar el número de identificación con el cual se encuentra registrado en las bases de datos donde se almacena la biometría, para que luego se compare con la minucia que se acaba de capturar del estudiante.

8.1.3 Base de Datos

Al contar con una herramienta de inteligencia de negocios, se almacenarán todos los datos de interés en un repositorio central, tanto los que pueda poseer la institución de educación superior, como los que se recolectarán utilizando el sistema de autenticación. Se sugiere utilizar bases de datos relacionales para asociar toda la información antes mencionada.

También se debe crear una tabla en donde se almacenen todas las transacciones que generen los registros de asistencia a los eventos de la institución, en donde se almacenen todos los datos necesarios para mostrar la información de interés.

El modelo planteado será el que se muestra en la Figura 5, en donde se muestra la información que se espera obtener por parte de las instituciones de educación superior, y se añadirán las tablas de biometría y de transacciones. En la de biometría se almacenarán toda la información obtenida en los enrolamientos y en la de transacciones se guardarán cada de los registros de

asistencia con la información de la persona que la generó y el resultado de la autenticación, así como las fechas en que fue ejecutada.

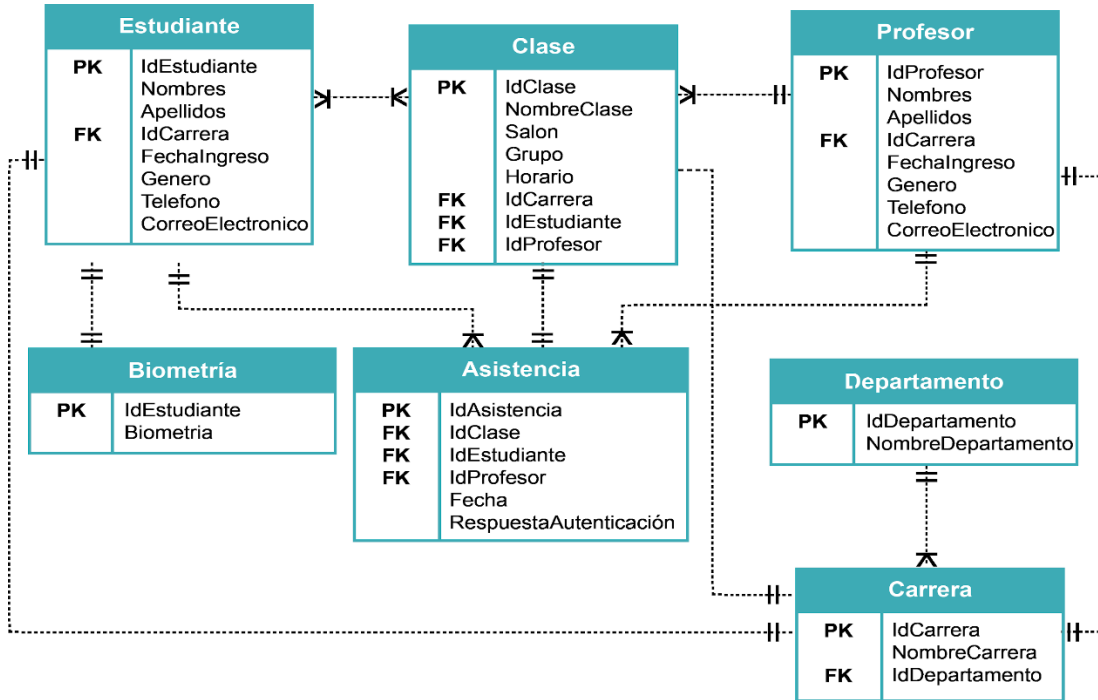


Figura 5 Modelo de datos. Elaboración Propia

8.1.4 Matcher o Cotejador

Todo sistema biométrico debe contar con un matcher o cotejador, el cual como se describió en el marco teórico, es el encargado de comparar 2 minucias de una misma biometría y arrojar un resultado de dicho cotejo.

Al realizar la comparación, verifica las minucias con detalle para evaluar puntos y características comunes para finalmente obtener un puntaje final de cotejo que define si el usuario fue autenticado o no.

Cada proveedor tiene su propio algoritmo para hacer la comparación biométrica y también tienen diferentes rangos de evaluación de puntaje, por lo que es importante en esta solución poder contar con un cotejador fiable obtener la menor cantidad de falsos rechazos o falsas aceptaciones.

8.1.5 Visualización

La visualización de datos es la representación visual es el proceso de búsqueda, interpretación, análisis y comprensión de datos textuales y numéricos para la toma de decisiones, que permite ahorrar tiempo y tener un conocimiento a profundidad que permita la toma de decisiones.

Es la parte más importante de la solución y todos los componentes descritos anteriormente son la base para contar con datos fiables para ser visualizados y facilitar la toma de decisiones de los integrantes de la institución de educación superior, que mejoren las condiciones académicas y del negocio.

8.2 Comparación y Análisis herramientas de visualización

Para la selección de la herramienta de visualización de datos se tuvo en cuenta el reporte *“Magic Quadrant for Analytics and Business Intelligence Platforms”* (Howson et al., 2018) de la empresa Gartner, teniendo en cuenta que esta es una compañía reconocida a nivel mundial por sus reportes sobre nuevas tecnologías en los que investigan y analizan el comportamiento y las tendencias de los mercados.

El cuadrante mágico las tecnologías analizadas se ubican en uno de los sectores del cuadrante dependiendo de dos variables, una es la integridad de visión que está relacionada con el conocimiento que tienen los proveedores sobre el mercado para crear oportunidades que generen valor al cliente y al negocio, y la segunda es la capacidad de ejecución que indica la habilidad del proveedor para ejecutar su visión de mercado. Los sectores en los que pueden ubicarse los proveedores analizados son: Líderes, visionarios, jugadores de nicho y retadores. (Gartner, 2018)



Figura 6 Cuadrante Mágico de Gartner. Fuente: Gartner

En el caso del reporte “*Magic Quadrant for Analytics and Business Intelligence Platforms*” como se puede ver en la Figura 6 fueron denominados como líderes del mercado las empresas Microsoft, Tableau y Qlik con sus productos Power BI, Tableau y QlikView respectivamente. Cabe resaltar que el reporte evalúa las herramientas bajo 5 casos de uso y 15 habilidades críticas para una plataforma de analítica y BI entre las que se destacan las siguientes por su afinidad con las necesidades del proyecto:

- Autoservicio de preparación de datos
- *Dashboards* analíticos
- Exploración visual interactiva

- Facilidad de uso, atractivo visual e integración de flujos de trabajo

	QlikView	Tableau	Power BI
MODO DE DESPLIEGUE			
Cloud	SI	SI	SI
On Premise	SI	SI	SI
DISPOSITIVOS SOPORTADOS			
Windows	SI	SI	SI
MacOS	SI	SI	NO
Web	SI	SI	SI
Android	SI	SI	SI
iOS	SI	SI	SI
INTEGRACIÓN			
REST API	NO	SI	SI
Lenguajes soportados	JavaScript, .Net	JavaScript	JavaScript, .Net, Python
FUNCIONALIDADES			
Usuarios	Técnicos	Técnicos Analistas	Técnicos No técnicos
Análisis de datos	SI	SI	SI
Visualización de datos	SI	SI	SI
Dashboard	SI	SI	SI
Trabajo colaborativo	SI	SI	SI
Links públicos	SI	SI	SI
Integración con R	NO	SI	SI
Soporte múltiples fuentes de datos	SI	SI	SI
SOPORTE			
Soporte telefónico	NO	SI	NO
Soporte en línea	SI	SI	SI
Base de conocimiento	NO	SI	SI
Tutoriales	SI	SI	SI
PRECIO (USD)			

Cuenta gratuita	Personal Gratuita	Pública Gratuita	Desktop Gratuita
Cuenta paga	Enterprise Por Cotización	ON PREMISE/PUBLIC CLOUD Tableau Viewer \$70/usuario/mes Tableau Explorer \$35/usuario/mes Tableau Creator \$12/usuario/mes HOSTED BY TABLEAU Tableau Viewer \$70/usuario/mes Tableau Explorer \$42/usuario/mes Tableau Creator \$15/usuario/mes	Pro \$9.99/usuario/mes Premium \$9.99/usuario/mes \$4.995/nodo/mes

Tabla 1 Comparación de herramientas de visualización de datos. Elaboración propia

El reporte también da una breve descripción y se enfoca en exponer los puntos fuertes y débiles de cada una de las herramientas, lo que permite enfocar la búsqueda de herramientas que cumplen con las necesidades específicas del proyecto, teniendo en cuenta lo anterior se hizo una tabla de comparación de las herramientas limitada a Power BI, Tableau y QlikView (ver Tabla 1), la información que aparece sobre cada uno de los productos se obtuvo de sus páginas web oficiales y se seleccionaron los siguientes aspectos que el equipo de desarrollo del proyecto considero relevantes: Modo de despliegue e integración, dispositivos soportados, funcionalidades disponibles, el servicio de soporte y los precios de las licencias.

Una de las razones más importantes para que dentro de la comparación sólo se tuvieran en cuenta estas tres herramientas es principalmente la facilidad de uso que ofrecen a los usuarios finales, especialmente en los casos de Power BI y Tableau en los que se destacan la facilidad con la que los usuarios pueden hacer un reporte o un *dashboard* con sólo arrastrar y soltar el gráfico que deseen sin necesidad de que los usuarios tengan conocimientos técnicos o de programación, ya que un punto importante en este proyecto es lograr que

cualquier persona interesada en los datos de asistencia a los eventos, las clases y el rendimiento de los estudiantes pueda generar reportes personalizados de acuerdo a sus necesidades para el cumplimiento de sus labores dentro de la entidad educativa.

En general en la actualidad la mayoría de herramientas de visualización de datos que se ofrecen en el mercado de analítica y BI proporcionan la función de crear *dashboards* analíticos con contenidos que permiten a los usuarios encontrar relaciones y patrones de comportamiento entre los datos a través de la exploración interactiva de los diferentes tipos de gráficas que ofrecen las herramientas, sin embargo, en este aspecto nuevamente están en la delantera las herramientas posicionadas como líderes en el cuadrante de Gartner, especialmente Power BI.

Un gran diferenciador que tiene Tableau frente a sus dos competidores es la gran cantidad de fuentes de datos con la que es posible conectarse para extraer, transformar y cargar datos, haciendo que en algunos casos las empresas medianas que adquieren este producto no requieran comprar o desarrollar software adicional que ejecute los procesos de ETL de datos para su posterior uso en el diseño de *dashboards* y reportes.

Además de la revisión del reporte de Gartner y la documentación de las herramientas, el equipo instaló e hizo uso de las versiones gratuitas de las tres herramientas durante una semana para evaluar la facilidad de uso de la herramienta desde una perspectiva distinta a la proporcionada por los expertos en sus reportes, y aunque en la Tabla 1 las tres herramientas están muy parejas en cuanto a las funcionalidades que ofrecen, el equipo seleccionó Power BI para el desarrollo del proyecto debido a la facilidad de uso, que permite apoyar la construcción de *dashboards* a través de la formulación de preguntas sobre los datos, también ofrece la opción de encontrar patrones o relaciones entre los datos a través de dos funcionalidades de análisis, una de

manera general sobre todos los datos y otra específica sobre un grupo de datos de una gráfica.

Power BI además ofrece una gran cantidad de recursos gráficos a través de la tienda y cuenta con la capacidad de integración con R permitiendo el uso de gráficas personalizadas en cualquier reporte o *dashboard*, la herramienta también cuenta con una gran cantidad de información que se puede conseguir a través de sus servicios de soporte, sus bases de conocimiento y comunidades por lo que la gran mayoría de dudas que surgen al hacer uso de este tipo de aplicaciones ya se encuentran resueltas en esta documentación. Otro aspecto importante para la selección de la herramienta fue el precio de las licencias, que en el caso de Power BI es considerablemente más bajo que los precios que manejan sus competidores. Para esta herramienta Microsoft ofrece dos tipos de licencia paga, Power BI *Pro* en la que el licenciamiento es individual y está enfocada en usuarios que requieren crear, compartir y ver contenido y Power BI *Premium* que licencia el contenido generado y puede ser visto por la cantidad de usuarios que se requiera sin incurrir en costos adicionales, sin embargo, estos usuarios no podrán crear contenido. (Esteves Wild, 2017)

Debido a que el sistema que se propone en este proyecto no está dirigido a una gran cantidad de usuarios que requieran crear y visualizar contenidos se sugiere el uso de licenciamiento *Pro* que tiene un costo de \$9.99 USD usuario/mes.

8.3 Análisis Sistemas de autenticación

En la actualidad cada vez más organizaciones, independientemente de la industria a la que pertenezcan, implementan sistemas de autenticación en uno o más procesos en los que requieran validar la identidad de sus empleados, proveedores o clientes. Esta validación se puede hacer a través de tres tipos

de factores de autenticación, que son: Factor de Conocimiento (*Algo que el usuario sabe*), Factor de Posesión (*Algo que el usuario tiene*) y Factor Inherente (*Algo que el usuario es*).

Por lo general las organizaciones usan sistemas de autenticación con un esquema clásico en el que se implementa lo que se denomina primer factor de autenticación, que normalmente hace referencia a un factor de conocimiento, como lo son las contraseñas, sin embargo, algunas organizaciones principalmente las gubernamentales y financieras, han decidido implementar segundos factores de autenticación a través de la implementación de esquemas de autenticación fuerte y multifactor aumentado así la seguridad en sus transacciones.

Es importante aclarar que un esquema de autenticación fuerte no necesariamente es multifactor. Es decir, que cuando se solicita durante el proceso de autenticación *Algo que el usuario sabe + Algo que el usuario sabe*, se está hablando de únicamente autenticación fuerte; a menos que se solicite adicionalmente, *Algo que el usuario tiene* o *Algo que el usuario es*, lo cual garantiza un mayor nivel de seguridad que además de continuar siendo autenticación fuerte, también es multifactor.

El esquema de autenticación multifactor más común está compuesto por factores de conocimiento y de posesión, esto debido a que los costos en la implementación y los dispositivos requeridos son menores respecto al de los factores inherentes, y a que hasta hace algunos años el mercado no contaba con algoritmos y dispositivos para autenticación biométrica (factores inherentes) que garantizaran un alto porcentaje de confiabilidad.

Sin embargo, en los últimos años la popularidad y adaptación de los sistemas biométricos para validar la identidad de las personas en diferentes tipos de organizaciones ha crecido, debido a las facilidades técnicas que existen

actualmente para integrar estos sistemas a casi cualquier proceso en el que la organización considere indispensable garantizar la autenticidad de la identidad de una persona y también, a la reducción de costos en los dispositivos requeridos para el registro y la autenticación de la información biométrica a través del uso de los dispositivos móviles de los clientes y empleados de las organizaciones.

Otra ventaja de este tipo de autenticación frente a los otros factores es que son difíciles de falsificar y no requieren recordarse, ni dispositivos adicionales a los necesarios para el proceso de autenticación, no obstante, tal como lo menciona (Sutcu, Tabassi, Sencar, & Memon, 2013) la autenticación biométrica es un factor que supone nuevos retos, como la robustez que requiere el sistema respecto a las variaciones en la información biométrica capturada por los dispositivos, y la definición de la cantidad de información biométrica que se va a recolectar. Estos puntos son elementos clave a la hora de diseñar y/o evaluar algoritmos y dispositivos de autenticación biométrica.

Algunos de los errores más comunes al momento de implementar un sistema de autenticación, es el no tener en cuenta que la captura de información debe ser minuciosa, es decir, dependiendo del tipo de biometría que se quiera utilizar, se deben cumplir algunos criterios para que se obtengan los resultados esperados. En el momento en que se presentan fallos, puede ser debido a que o los enrolamientos (información biométrica que se utiliza como referencia al momento de hacer un cotejo), o las capturas de autenticación o identificación, no fueron realizados bajo los lineamientos óptimos.

Cabe aclarar que existen posibilidades en que no se obtenga un resultado exitoso, dado que cada sistema de autenticación tiene un rango de error en el que presenta ya sea un falso rechazo o una falsa aceptación, en otras palabras, que una persona que ya ha sido enrolada correctamente trata de autenticarse y tiene como resultado que no ha sido identificado; o el caso de

falsa aceptación, que una persona trata de suplantar a otra y obtiene un resultado positivo. Con estos factores se puede conocer qué sistema de autenticación es más fiable a la hora de cotejar, y dependiendo del uso que se le quiera dar, puede llegar a ser delicado si se trata de transacciones que requieran alta seguridad.

Cada sistema de autenticación debe contar siempre con un grado de seguridad, ya que es información sensible de cada persona; por ello, los diferentes sistemas biométricos cuentan ya sea con una codificación antes de proceder a ser cotejado, o poseen distintos formatos de captura para proteger el contenido de la biometría.

Una característica importante del uso de un sistema biométrico es que captura los rasgos físicos o de comportamiento de una persona, en este caso se almacenarán propiedades inherentes de los estudiantes; pero, hay un motivo por el cual se plantea esta solución para una institución educativa con estudiantes universitarios, y es debido a que los niños con edad inferior a los 14 años están en constante crecimiento y con ello, van a ir cambiando sus características físicas. Por otro lado, la biometría en la gente mayor va desapareciendo a medida que pasan los años, por ejemplo las huellas se van desvaneciendo y los patrones faciales van cambiando; por esta razón, en lugares como Canadá tienen como requerimiento que las personas deben enrolarse siempre y cuando cumplan con los rangos de edad de entre 14 y 79 años de edad (Gobierno de Canadá, n.d.)

8.3.1 Biometría dactilar

El primero de ellos es el de biometría dactilar, es la más utilizada debido a su robustez, fiabilidad y fácil implementación, existen en el mercado muchos proveedores y compañías que ofrecen este tipo de tecnología ya sea para control de acceso, aplicaciones web o móvil o autorización de transacciones.

En este proyecto únicamente trataremos los dispositivos de captura de una única huella, dado que es el de más fácil implementación en un aula de clases, porque los otros existentes toman la información biométrica de diferentes formas como por ejemplo el 4-4-2 (dedos de la mano derecha, dedos de la mano izquierda, pulgares), lo cual causaría que la autenticación de cada estudiante sea más demorada.

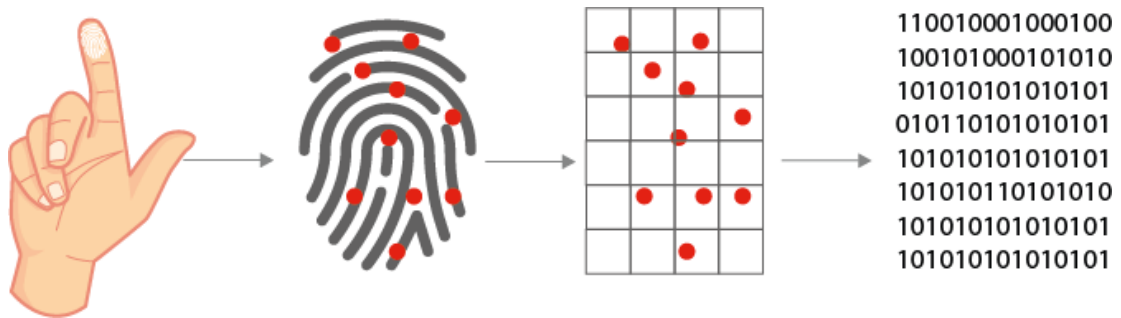


Figura 7 Extracción de minucias. Elaboración propia

Como se muestra en la Figura 7, este sistema funciona a partir de un lector de huella digital que se encarga de tomar la imagen para comparar los patrones de esa huella con los que se tienen registrados desde el enrolamiento. Estos patrones son conocidos con el concepto de minucia, “que es un punto de interés de la huella, que se caracteriza por su posición (x, y) y un ángulo ya sea de una terminación o bifurcación” (Rosales Cruz, 2001) Para obtener una coincidencia, el sistema del lector no necesita encontrar el patrón completo en la muestra, simplemente debe encontrar un número suficiente de minucias (que dependen de la programación del lector) en ambas imágenes, y dependiendo de la cantidad de patrones que encuentre similares, le asignará un puntaje al cotejo.

“La idea básica es medir las posiciones relativas de las minucias, es decir la distancia media entre crestas y el número máximo de muestras que es de diez. Para ello el primer paso es alinear el patrón con la muestra de la base de datos, a continuación, se convierten los dos patrones alineados en dos cadenas de

minucias que, en el proceso de comparación, minimizan una función de coste, llamada distancia de edición. Finalmente, el coste obtenido al comparar los dos patrones proporciona un valor, que indica el grado de similitud entre las cadenas de minucias y que determina si la autenticación es correcta o no". (Simon, García Gomar, Sánchez Asenjo, L Sánchez Bote, & Ortega García, 2018)

Hay varios tipos de dispositivos y fabricantes por lo cual fácil de conseguir en el mercado, es un sistema no invasivo, que tiene una captura rápida, no necesita condiciones especiales para que el usuario realice su autenticación y en casos específicos, por ejemplo, personas que sufran de enfermedades dermatológicas se puede graduar para que el puntaje de autenticación sea más flexible y pueda efectuar la captura biométrica.

Desafortunadamente al ser uno de los sistemas más utilizados en el mercado, existen muchas técnicas que logran falsificar la autenticación, entre ellas se pueden encontrar la copia de huellas por pegamento, látex, silicona y otros materiales, que han hecho vulnerable el sistema y ahora cualquier fabricante de dispositivos dactilares biométricos debe cumplir con la norma de detención de huella viva, lo que significa que no solo captura la imagen, sino también las pulsaciones. Además, puede tener problemas de reconocimiento cuando se presentan heridas o callosidades que dificultan el proceso de digitalización de sus huellas.

8.3.2 Reconocimiento facial

El siguiente sistema de autenticación es el de reconocimiento facial, que también es uno de los más implementados, de hecho, se pueden encontrar algunos matchers online en el que realizan cotejos de huellas y biometría facial de buena calidad.

La gran diferencia con la biometría dactilar para el usuario final es en el momento de realizar la captura, porque para las huellas únicamente se necesita un dispositivo, pero las fotos para el reconocimiento facial deben cumplir algunos criterios para que se considere una captura óptima, estos criterios se encuentran descritos en la norma internacional ISO/IEC 19794-5, creada por el Comité Internacional para los Estándares de Información y Tecnología, y adoptada por la Organización Internacional de la Aviación Civil (ICAO), la norma incluye más de 15 requisitos, que se dividen en tres categorías principales, “especificaciones de escena (pose facial, expresión, uso de accesorios), características fotográficas (iluminación, nitidez, distancia focal, exposición, saturación) y atributos propios de las imágenes digitales (resolución, nivel de compresión, formato de los archivos, formas de almacenamiento)”. (Vazquez, Chang, Rizo, & Morales-González, 2012)

En la Tabla 2 se mencionan los parámetros de cada categoría.

Especificaciones de escena	
Iluminación	La luz debe estar distribuida uniformemente en el rostro, el cual no debe estar afectado por sobras, ni regiones brillantes.
Ojos	Ojos abiertos, visibles, sin ojos rojos y mirada al frente
Boca	La boca debe estar cerrada y tener una expresión neutral
Postura	Posición frontal
Fondo	Fondo uniforme
Espeuelos	No reflejos y solo se pueden utilizar gafas que no oscurezcan los ojos, transparentes de manera que las pupilas y el iris queden visibles
Características fotográficas	
Posición	Cabeza ubicada en el centro
Distancia	El largo de la cabeza debe comprender entre el 70%-80% del alto de la imagen. La razón entre el ancho de la cabeza y la imagen debe encontrarse entre 4:7 y 1:”

Color	Los colores deben ser naturales, no saturados y con un contraste adecuado
Atributos imágenes digitales	
Nitidez	Las imágenes deben estar enfocadas y no borrosas
Resolución	La distancia entre el centro de los ojos debe ser de mínimo 60 pixeles.

Tabla 2 Requerimientos captura facial. Fuente: ICAO

A partir de estas especificaciones técnicas de la fotografía se desarrolla el proceso de reconocimiento facial que está dividido en 4 fases y que es similar en todos los sistemas biométricos, detección de la cara, alineación extracción de las características y reconocimiento.

El proceso inicia con la captura de la imagen que puede adquirirse por celulares, cámaras o video; con la imagen obtenida se realiza la fase de detección, que es el proceso por el cual se separa el área de la cara del fondo de la imagen, con la imagen se realiza un proceso de alineación de los componentes de la cara, es decir los ojos, la nariz, las orejas y la boca. “Posteriormente, se produce la normalización con respecto a las transformaciones geométricas, como el tamaño y la pose, además de las propiedades fotométricas, como la iluminación y la escala de grises”, para realizar una extracción de características en el que se obtiene información biométrica con respecto a variaciones geométricas y fotométricas y se generando un patrón que se compara con los vectores de características extraídos de las caras de la base de datos y el matcher da un puntaje que determina la similitud de las imágenes.

Entre los beneficios de este método de identificación encontramos la facilidad y rapidez de la captura, no tiene contacto con el dispositivo, es decir que no es invasivo, permite la captura de personas en movimiento, es un sistema que

puede camuflarse y es difícil de vulnerar ya que no es fácil falsificar características físicas de las personas.

Pero a la hora de identificar a una persona a partir del reconocimiento facial, existen una serie de dificultades como la imprecisión de la identificación, requiere imágenes de buena calidad, buena iluminación, se puede vulnerar cuando la persona ha tenido accidentes o cirugías y se hace muy difícil cuando la variabilidad entre individuos es muy pequeña (por ejemplo, es el caso de familiares o, especialmente, es el caso de los gemelos), o cuando al comparar dos imágenes de un mismo individuo es muy distinta. (Jha, 2007)

8.3.3 Reconocimiento del iris

Esta forma de autenticación se hace por medio de un proceso de adquisición riguroso, el sistema captura la imagen a nivel de grises mediante *scanners* que capturan la superficie del ojo, posteriormente se realiza una etapa de preprocesado en el que se aísla la estructura del iris a partir de la detección de los contornos exterior e interior del mismo, después se hace una extracción de características de la información relativa a los sectores izquierdo y derecho dividiendo la captura de la imagen en cuatro sectores ya que no se puede controlar la abertura del párpado y finalmente se realizará una reordenación de la información resultante mediante un muestreo tanto en radio como en ángulo de la misma, obteniendo una imagen cuadrada generando el vector patrón que se utiliza para aplicar el algoritmo de *matching* basado en la mínima distancia entre el patrón de *test* y todos los modelos. (Duró, 2014)

Este tipo de biometría es el más seguro que existe en el mercado, ya que el iris es predominantemente plano, y su configuración geométrica es controlada únicamente por dos músculos complementarios haciendo que su forma sea mucho más predecible que otros sistemas; es órgano interno que está bien protegido contra daños y el desgaste contrario a lo que sucede con las huellas

dactilares y uno de sus aspectos más importantes es la facilidad para detectar un fraude, ya que con la captura de dos imágenes consecutivas se puede comprobar la correlación y la diferencia de tamaño de la pupila.

A comparación de los dos primeros sistemas de autenticación, el reconocimiento de iris tiene un costo muy elevado, su tiempo de captura es de más de 60 segundos, es susceptible a la mala calidad de la imagen y aunque no es invasivo, tiene poca popularidad por la poca distancia entre el usuario y el dispositivo de captura.

8.3.4 Reconocimiento de voz

El reconocimiento de voz es el reconocimiento de patrones a partir de las características de la señal de voz, para hacer estos patrones existen muchas técnicas disponibles tales como MFCC, LPCC, LPC, Prosódico, etc. Una de las características que se tienen en cuenta para la creación de patrones son los espectros de la señal. Después de la extracción de los patrones, hay dos fases, la de inscripción y la de verificación; en la primera fase se extraen impresiones de voz de la grabación y en la verificación se comparan calidad, duración, dinámica, intensidad y tono de la señal con la muestra de impresiones con las plantillas almacenadas, seleccionando la que mejor coincidencia tiene. (Singh, Agrawal, & Khan, 2018).

Es una tecnología biométrica precisa y fiable, ya que la voz es única e irrepetible, permite que los clientes interactúen con el sistema y aquí lo importante no es lo que dices sino cómo lo dices y es difícil de vulnerar, ya que no se puede robar una huella vocal o suplantarla sin que el sistema detecte que se está haciendo uso de una grabación.

Al haber interferencia en el ambiente puede causar un error en la autenticación, se requiere de un desarrollo muy preciso y el tiempo de

verificación tiende a ser más alto que el de otros sistemas mencionados anteriormente.

8.3.5 RFID

A diferencia de los otros sistemas de autenticación el RFID o identificador por radiofrecuencia, no es un sistema biométrico, sino una tecnología que utiliza frecuencias de radio para identificar objeto físico, en este caso la identidad de una persona, su funcionamiento es simple. El usuario desliza la tarjeta a través de un lector que lee la etiqueta que contiene los datos de identificación del objeto al que se encuentra adherida, generando una señal de radiofrecuencia con dichos datos la compara con la información almacenada en la base de datos y, si corresponde, toma la asistencia. (Mshelia, Alkali, Isuwa, Dibal, & Onogwu, 2017).

En cuanto a las ventajas de este sistema tenemos lecturas rápidas y precisas, reducción de procesamiento de datos y mayor seguridad, no depende de las condiciones del ambiente, es una operación sin contacto, además cuenta con una gran variedad de formatos y tamaños.

El principal inconveniente de esta tecnología es el precio, ya que tiene que reponerse si se reporta una pérdida, además existe la posibilidad de que la lectura de datos sea defectuosa o que si se utiliza 2 lectores a la vez para una misma tarjeta RFID ésta no puede dar una información correcta ya que los dispositivos lectores cruzarán sus ondas y la tarjeta no es capaz de responder a 2 consultas simultáneas.

8.4 Escenarios

De acuerdo con el análisis de las herramientas de visualización y de los sistemas de autenticación, se exponen 4 escenarios, (asistencia a clases, evaluaciones, conferencias y simulacros de evacuación) donde se describe,

en cada caso, el sistema de información acorde con el contexto, las ventajas y desventajas que tiene cada sistema escogido para poder evidenciar de qué manera se puede implementar y qué beneficios se van a obtener.

Para elegir el sistema biométrico que más se acopla a los requerimientos de cada escenario, se debe conocer las especificaciones de cada biometría teniendo en cuenta parámetros como: velocidad de captura, precio de la licencia, precisión (bajo índice de falso rechazo y falsa aceptación), tamaño de información que va a generar y vulnerabilidad. Se decidió tener como punto de referencia todas las soluciones de la empresa Neurotechnology, en los cuales implementan todos los sistemas biométricos tratados en este documento y que se evidencian en los anexos mencionados a continuación: Anexo 2. *VeriFinger SDK Brochure*, en el cual se implementa una solución basada en identificación dactilar; Anexo 3. *Face Verification SDK Brochure*, donde se utiliza biometría facial; Anexo 4. *VeriEye SDK Brochure*, que hace uso del reconocimiento de iris para su posterior identificación y por último el Anexo 5. *VeriSpeak SDK Brochure* que está diseñado con algoritmos de reconocimiento e identificación de voz.

8.4.1 Asistencia a clases

8.4.1.1 Objetivos:

- Captura de asistencia rápida en comparación con la manual.
- Validar la identidad del estudiante que asiste a clase.
- Registro general e individual de asistencia.
- Visualización de los registros recolectados por diferentes filtros (estudiante, materia, departamento, carrera).

8.4.1.2 Sistema de autenticación sugerido:

Biometría dactilar.

De acuerdo a los resultados de los documentos consultados y la experiencia de los integrantes del grupo, el sistema que se propone es el de biometría dactilar porque es el más fiable y se caracteriza por tener una captura no invasiva y rápida como se evidencia en el estudio realizado (C.O et al., 2013), en el que se desarrolló un sistema de autenticación dactilar para 87 estudiantes y 30 empleados, teniendo un porcentaje de 2,56% en falso rechazos y falsa aceptación y un tiempo de aproximadamente 8 minutos por medio del sistema biométrico en comparación a los 36 minutos por medio del sistema manual, que toma el sistema biométrico en comparación con el sistema manual, como se muestra en la Gráfico 1.

Table 2a. Details of the evaluation

	Successful Verification	Unsuccessful Verification
Staff	30	0
Student	84	3
Total	114	3

Table 2b. Details of the evaluation

No of Attendee	Successful Verification	Unsuccessful Verification	Accuracy
117	114	3	97.4%

Table 3. Time taken for Verification

Type of system	No of attendee	Total time in seconds	Total time in minutes	Average Execution Time in seconds
Fingerprint-Based Attendance System	117	502.41	8.37	4.29
Manual Attendance System	117	2161.55	36.02	18.48

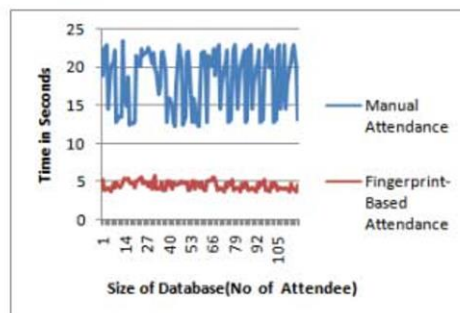


Gráfico 1 Evaluación sistema biométrico vs manual. Recuperado Fingerprint-Based Attendance Management System

De esta manera se evidencia que al aplicar esta solución se obtendrá no solo una disminución en el tiempo de tomas de asistencia como se muestra en la Gráfico 1, sino también se ahorraría en el tiempo que el instructor se toma en subir este registro al sistema de la institución de educación superior, eliminando los posibles errores humanos al momento de digitar.

Adicionalmente, como se mencionó en la problemática, al no tener un compromiso de los instructores por el tiempo que toma hacer un chequeo manual, que genera la inasistencia y desinterés por parte de los estudiantes al utilizar este sistema se puede lograr un mejor compromiso por parte de los estudiantes como lo demuestra el artículo (Trabelsi & Shuaib, 2011) en el que se evaluó el número de ausentes durante seis semestres, tres meses utilizando el sistema manual y los siguientes 3 semestres, Gráfico 2 utilizando el sistema biométrico; y la percepción de los estudiantes sobre cómo el tener un sistema biométrico para el registro de asistencias influye en su decisión de ir o no a clase.

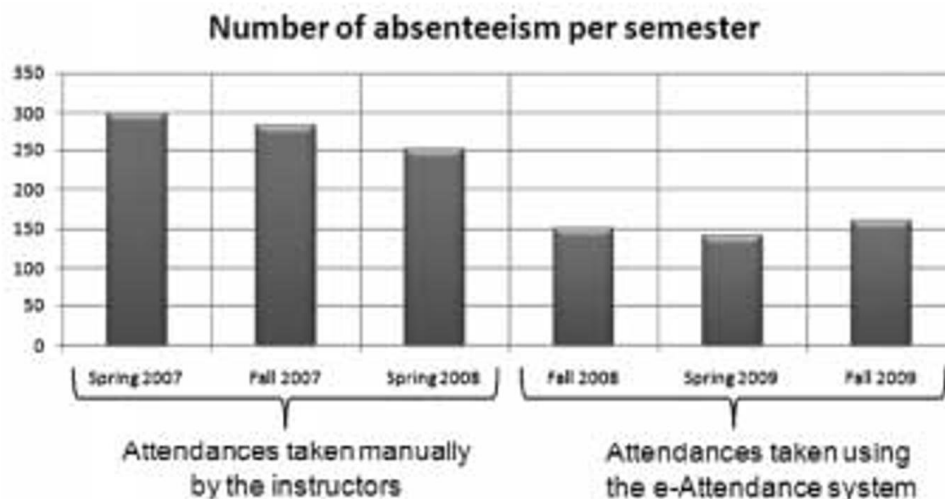


Gráfico 2 Absentismo por semestre. Recuperado IMPLEMENTATION OF AN EFFECTIVE

8.4.1.2.1 Costo licencia sistema biométrico:

\$382 USD. Múltiples equipos. Fuente: Anexo 2

8.4.1.3 Implementación

Para la realización de este registro se tendrá que poner en cada salón un huellero de fácil acceso y además de una interfaz gráfica para que cada profesor pueda ingresar la clase que se va a iniciar y se pueda revisar el resultado de la autenticación de cada estudiante.

8.4.1.4 Diagrama de flujo de actividades:

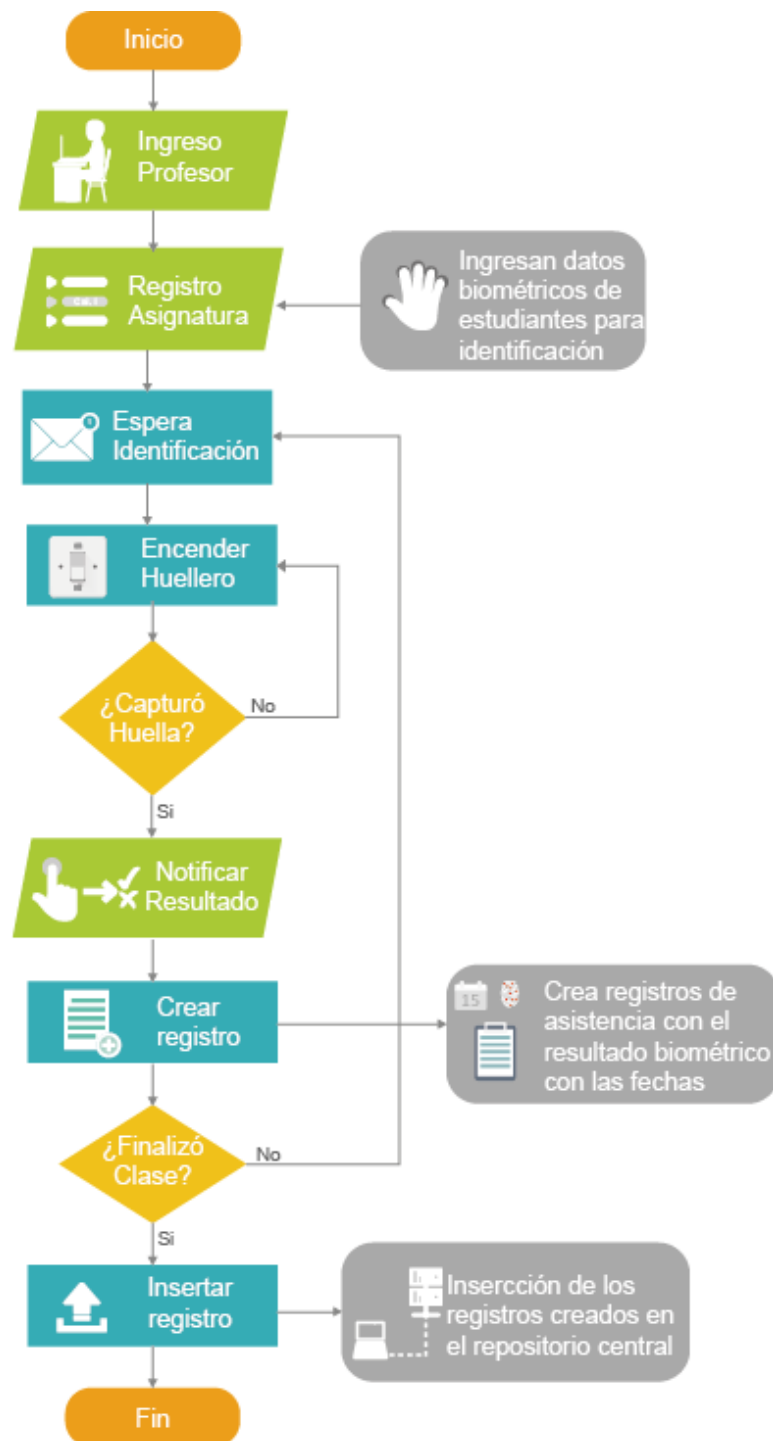


Figura 8 Diagrama de Flujo, Registro asistencia a clase. Elaboración propia

8.4.1.5 Análisis y explotación de los datos:

Para el análisis y la explotación de los datos relacionados con la asistencia a clase se propone la visualización de la información se haga en un *dashboard* en el que se pueda perfilar a los estudiantes que no asisten a clase con regularidad (*ver figura 9*), esto con el fin de tener un mejor conocimiento de la población y así poder establecer si existe un patrón que haga aumentar o disminuir el porcentaje de ausencias a clases.

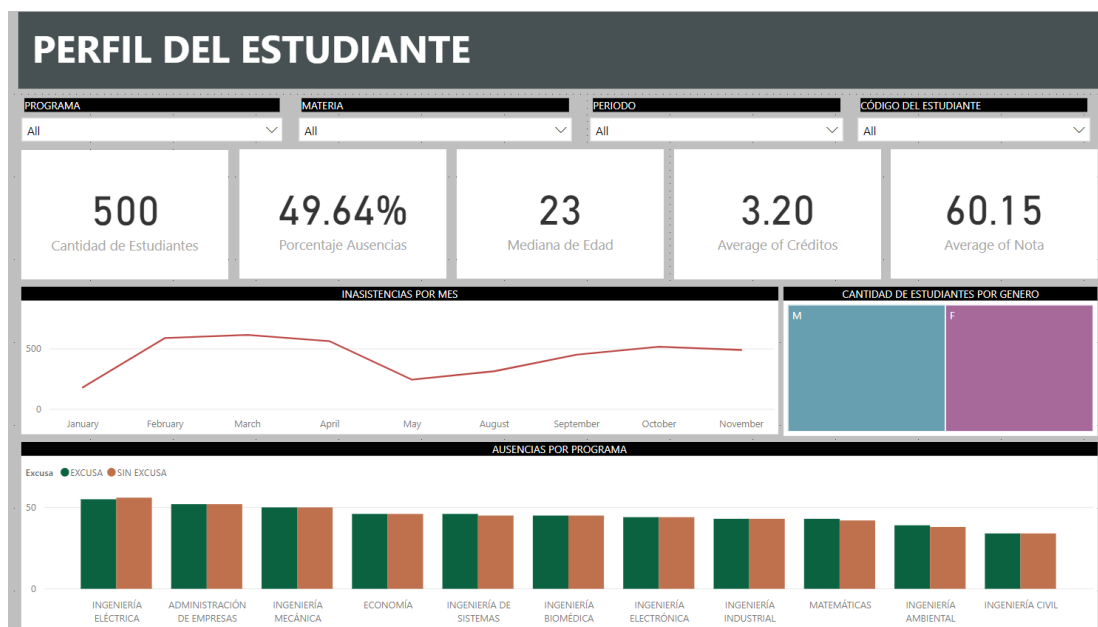


Figura 9 Dashboard del perfil de los estudiantes. Elaboración propia

En la figura 9, se muestran algunos indicadores que pueden usarse para perfilar a los estudiantes, como el porcentaje de ausencias, la mediana de la edad de los estudiantes, el género de los estudiantes, el promedio de créditos inscritos y el promedio de calificación obtenida. También se agregó información relevante como la cantidad total de estudiantes, el comportamiento de las inasistencias a clase por mes, así como la cantidad de ausencias por programa indicando para cuantas de estas los estudiantes presentaron una excusa que justificara su ausencia. El *dashboard* además

incluye filtros que le permitan a los usuarios conocer los datos mencionados anteriormente con mayor detalle, bien sea por programa, asignatura, periodo escolar o por estudiante.

El objetivo de este tipo de artefacto es que el usuario final conozca el estado actual de una situación en particular y pueda tomar decisiones basadas en información. En este caso lo que se busca es que los usuarios puedan dar un diagnóstico sobre las ausencias de los estudiantes en las clases, a través de la identificación de los programas, asignaturas y fechas en las que se presentan más ausencias esto junto con datos que den contexto a estas situaciones, de modo que se puedan implementar cambios en la institución que reduzcan los porcentajes de ausencia, por ejemplo a través de modificaciones en los horarios en los que se dicta una clase.

Es importante considerar que este diagnóstico puede dar pie para que se creen espacios en los que se discuta con los estudiantes que frecuentemente dejan de asistir a las clases los motivos por los que lo hacen y hacer seguimiento para validar qué tanto afectan las ausencias su rendimiento escolar, de modo que se identifique si existe algún servicio adicional que se le pueda prestar a los estudiantes y que haga que mejoren sus niveles de asistencia, o si se pueden mejorar los contenidos de una materia en particular.

8.4.1.5.1 Costo licencia sistema de visualización:

Power BI *Pro*: \$9.99 USD usuario/mes

8.4.2 Evaluaciones y exámenes finales

8.4.2.1 Objetivos:

- Realizar una captura biométrica rápida.
- Hacer una autenticación fuerte a cada estudiante, validando su identidad mediante la biometría y verificando sus datos (nombre o identificación) a través de un documento.
- Visualizar fraudes, suplantaciones, profesor encargado y salón requerido para tomar decisiones inmediatas si hay inconvenientes con lo anteriormente mencionado.

8.4.2.2 Sistema de autenticación sugerido:

Biometría dactilar y lector de código de barras o RFID.

En este escenario se propone utilizar dos sistemas de autenticación, para tener mayor seguridad a la hora de evaluar la identidad del estudiante, sin que el uso de dos sistemas afecte la rapidez de toma de los datos. Además de lo mencionado anteriormente en las ventajas que tiene la implementación de biometría dactilar, se eligió la captura de huellas porque es el tipo de biometría que tiene el menor porcentaje de falso rechazo y falsa aceptación; mientras que el lector de código de barras o RFID se escogió para verificar el número de identificación del estudiante.

Por medio de esta autenticación se asegurará la identidad del estudiante evitando fraudes, además de proporcionar una alerta a los tomadores de decisiones en caso de que el profesor encargado del examen o evaluación no se presente, para que puedan asignar un nuevo instructor y no se dé por perdido el tiempo y el lugar estipulado para la actividad.

8.4.2.2.1 Costo licencia sistema biométrico

\$107 USD. De 1 a 9 computadores: Fuente Anexo 6. NCheck Bio Attendance Brochure

8.4.2.3 Implementación:

En este escenario es necesario implementar el huellero y la interfaz gráfica al igual que en el escenario de asistencia a clase, agregando una aplicación móvil que maneje el profesor, que permita la lectura del carnet.

8.4.2.4 Diagrama de flujo de actividades:

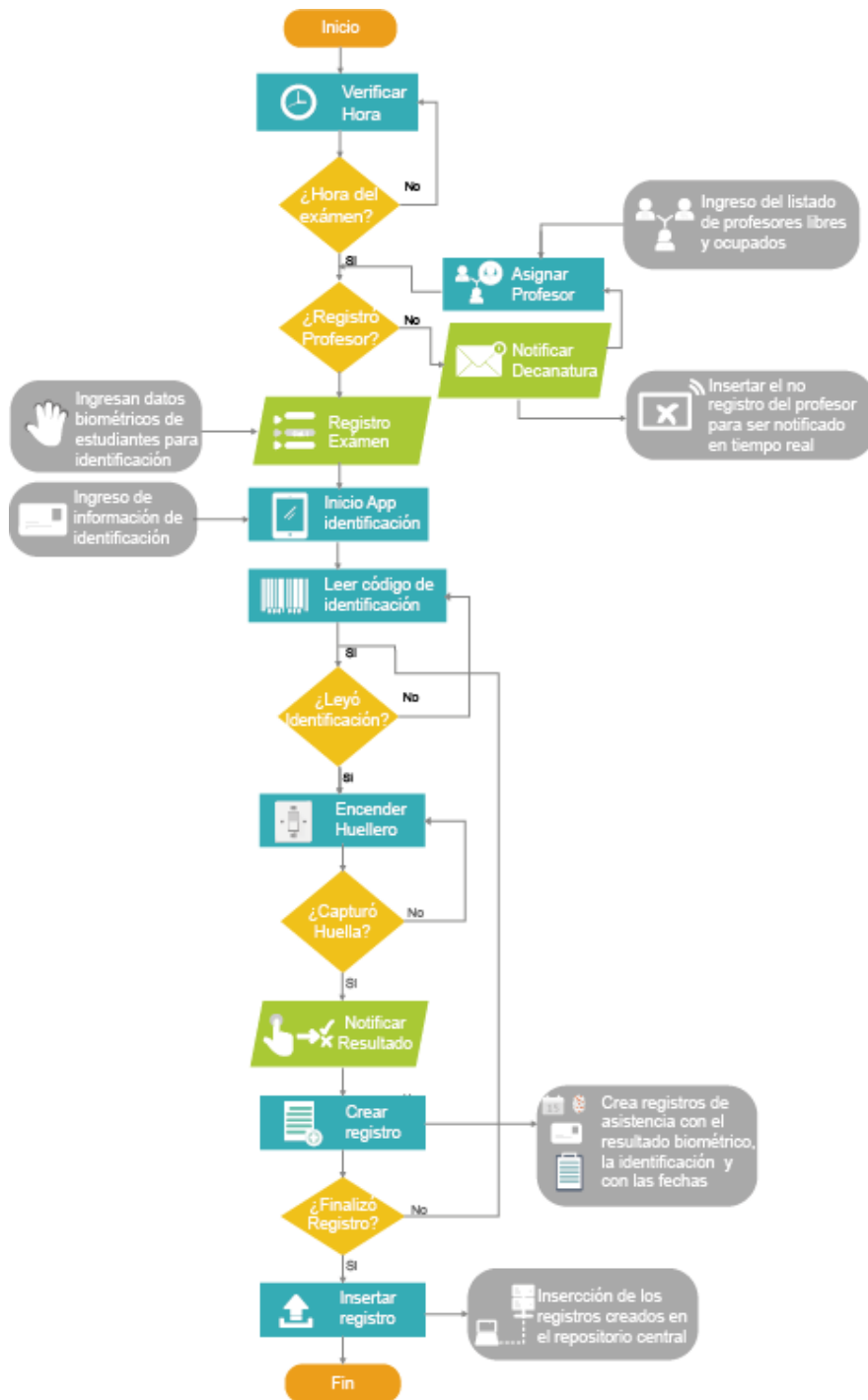


Figura 10 Diagrama de Flujo Registro Asistencia Exámenes. Elaboración propia

8.4.2.5 Análisis y explotación de los datos:

En la figura 11 se visualiza el dashboard propuesto para el escenario de asistencia a exámenes finales, en este caso se tuvo en cuenta la siguiente información: el porcentaje de asistencia, el número de asistentes al examen, la cantidad de estudiantes inscritos para la materia, la cantidad de estudiantes por programa que están activos o que cancelaron la materia, el salón en el que se lleva a cabo el examen y la capacidad que este tiene.

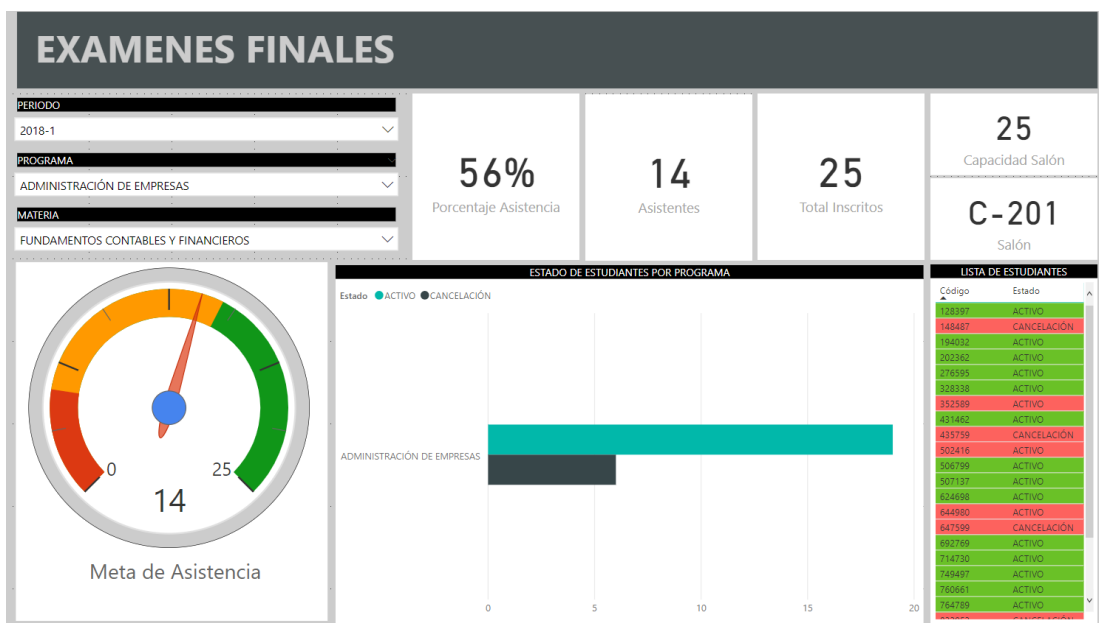


Figura 11 Dashboard de asistencia a exámenes finales. Elaboración propia

Además, se cuenta con el listado de los estudiantes inscritos, en el que resalta en color rojo aquellos estudiantes que no asistieron al examen, bien sea porque cancelaron la materia o simplemente no se presentaron.

La información que brinda este *dashboard* permite saber si el espacio que se dispone para los exámenes es el adecuado para la cantidad de estudiantes, y al igual que en el escenario anterior permite abrir canales de comunicación que permita entender porque un estudiante no asiste a presentar el examen

final de la materia aun cuando en el sistema aparece activo y así crear espacios en los que se le pueda brindar ayuda a estos estudiantes.

8.4.2.5.1 Costo licencia sistema de visualización:

Power BI *Pro*: \$9.99 USD usuario/mes

8.4.3 Charlas y conferencias

8.4.3.1 Objetivos:

- Hacer un conteo de la cantidad de asistentes a una charla o conferencia.
- Dar la posibilidad de verificar los asistentes a las charlas o conferencias.

8.4.3.2 Sistema de autenticación sugerido:

Lector de código de barras o RFID.

Para este escenario se propone el RFID o lector de código de barras, ya que, al comparar los diferentes sistemas de autenticación, se determina que es el sistema más rápido para la toma de asistencia y ya que el objetivo de este escenario es el conteo y no es de vital importancia validar la identidad de la persona que va a ingresar, no se necesita que sea un sistema biométrico.

Con la implementación de este sistema, al igual que con el de asistencia a clase se tendrá un conteo real y automático de los asistentes a las conferencias o charlas, evitando la pérdida de información y el posible fraude que normalmente sucede con la asistencia manual que se realiza a través de la rotación de hojas donde se consigna el nombre del estudiante y algún dato adicional.

8.4.3.2.1 Costo licencia sistema biométrico

\$50-120 USD Neurotechnology no ofrece este sistema de identificación individualmente si no como complemento de otros sistemas biométricos, por lo que el precio es estimado.

8.4.3.3 Implementación

Para la implementación de este sistema se sugiere utilizar una aplicación móvil, que permita la lectura del carnet o documento de identificación en la entrada del auditorio o salón, permitiendo al sistema identificar al estudiante que asiste a la charla.

8.4.3.4 Diagrama de Flujo

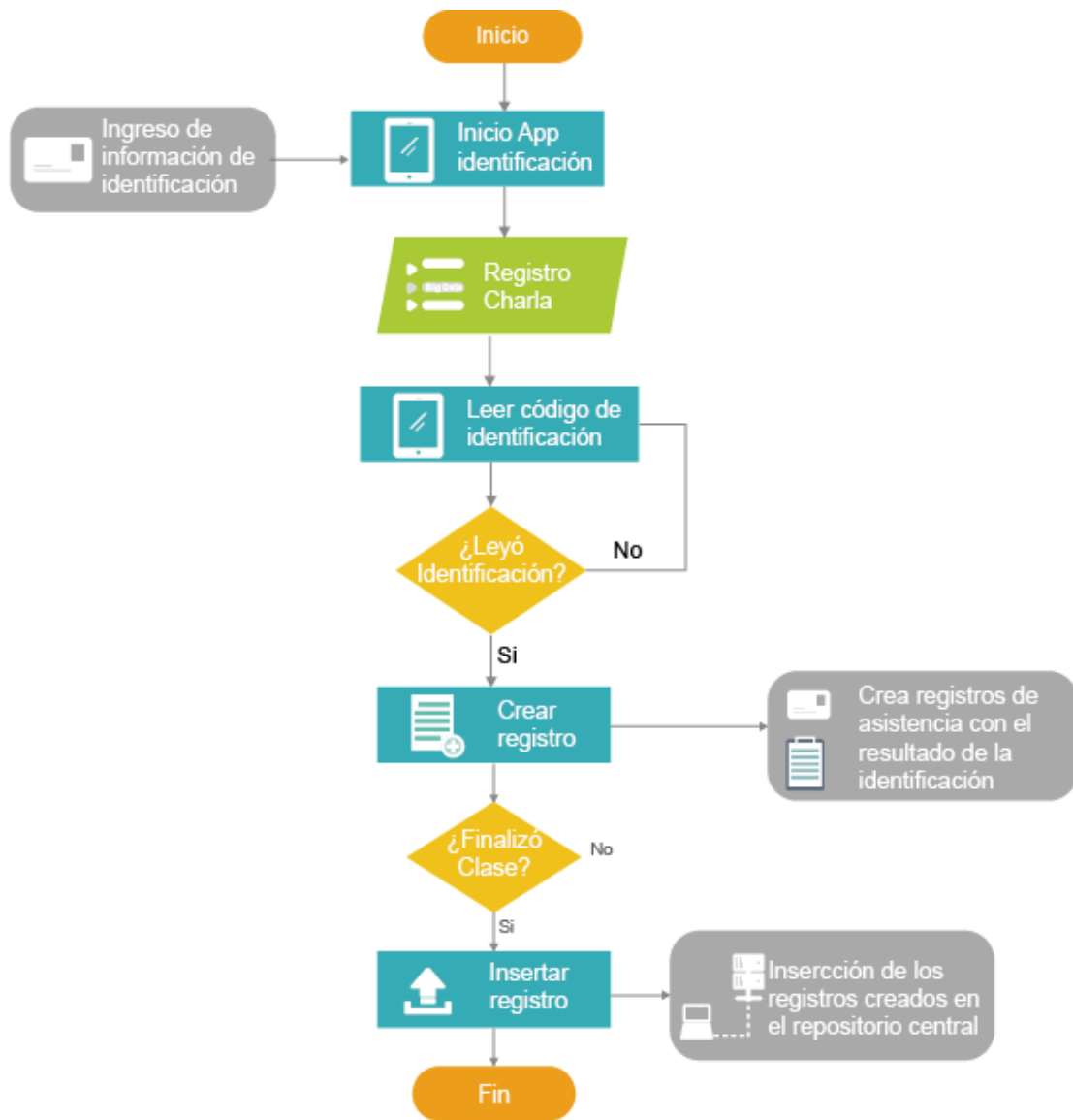


Figura 12 Diagrama de Flujo Registro Asistencia a Conferencias. Elaboración Propia

8.4.3.5 Análisis y explotación de los datos:

El *dashboard* de asistencia a eventos (ver figura 13) permite filtrar la información por evento, programa y número de sesión. Además, reúne los siguientes indicadores: Meta de sesiones, meta de asistencia, porcentaje de asistencia, cantidad de asistentes y la cantidad de asistentes al evento por

programa. Como información adicional se incluye el nombre del salón en el que se llevó a cabo el evento y la capacidad que tiene.

El indicador de meta de sesiones se incluyó debido a que las instituciones de educación superior para el cumplimiento de sus obligaciones legales cuentan con algunas charlas que se deben hacer para una cantidad específica de miembros de la comunidad educativa, por lo que generalmente se debe hacer más de una sesión para cumplir con la meta de asistencia requerida. En conclusión, la meta de sesiones hace referencia a la cantidad de veces que se requiere hacer un evento hasta cumplir con la meta de asistencia.

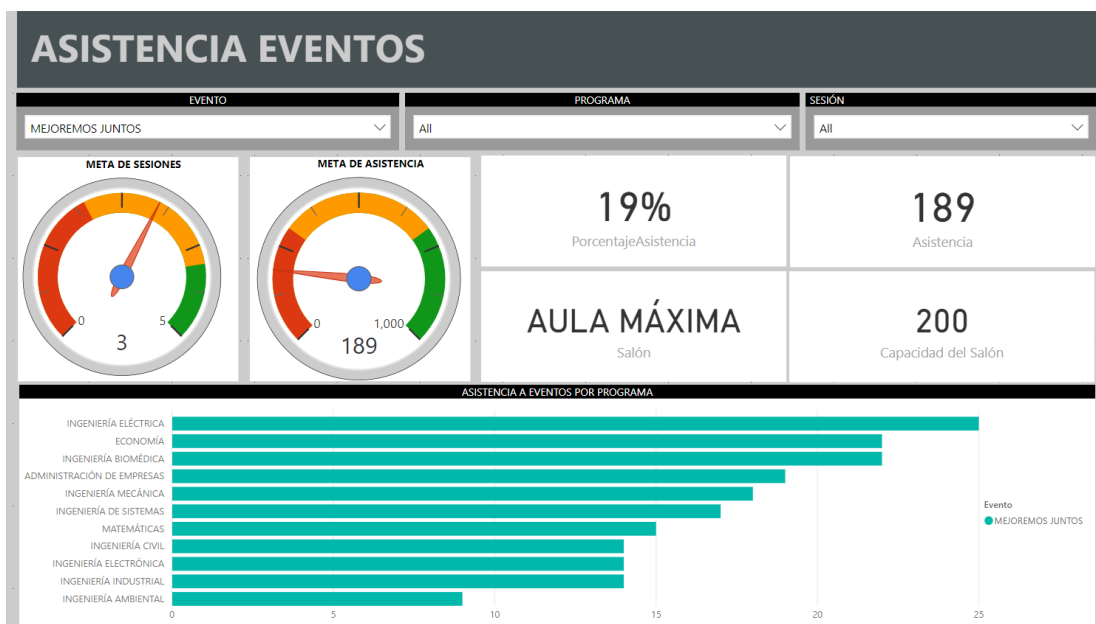


Figura 13 Dashboard de asistencia a eventos. Elaboración propia

Este *dashboard* permite conocer las preferencias que tienen los estudiantes de cada programa frente a los eventos que se hacen en la universidad, si la audiencia que asiste es el público objetivo, si se está aprovechando la capacidad de las salas para ese tipo de eventos o si a futuro podría usarse un espacio más pequeño o más grande según se requiera.

Además, para los eventos en los que se requiera realizar más de una sesión, se puede saber de manera rápida si la meta se va a cumplir con las sesiones que se tienen planeadas, si se requiere hacer más propaganda del evento para cumplir la meta o si definitivamente se va a tener que aumentar la cantidad de sesiones.

8.4.3.5.1 Costo licencia sistema de visualización:

Power BI *Pro*: \$9.99 USD usuario/mes

8.4.4 Simulacros de evacuación

8.4.4.1 Objetivos:

- Realizar una identificación de personas en los puntos de encuentro.
- Hacer un conteo rápido y aproximado de personas que hagan parte de un simulacro.
- Generar un registro de las personas aproximadas que participaron en el simulacro y evacuaron en los tiempos establecidos.

8.4.4.2 Sistema de autenticación sugerido:

Biometría facial.

Para este escenario es necesario implementar un sistema biométrico especial que requiere la identificación de varios individuos de forma simultánea. Conociendo los sistemas biométricos actuales, lo sugerido es implementar identificación facial porque tiene ciertos algoritmos que logran identificar el rostro o la silueta de una persona.

Este sistema no tiene la misma precisión que la dactilar y es mucho más costoso por el software necesario para su funcionamiento, pero se adecúa al

escenario, ya que el principal objetivo en un simulacro es hacer el conteo rápido de los estudiantes y empleados, y no identificarlos.

Adicionalmente al conteo, también se puede identificar si durante el simulacro hay personas que continúa en el edificio para notificar al tomador de decisiones.

8.4.4.2.1 Costo licencia sistema biométrico:

\$890 USD. Fuente: Anexo 7. SentiVeillance SDK Brochure

8.4.4.3 Implementación

Para el funcionamiento de este sistema es necesario el uso de las cámaras con las que cuenta la institución de educación superior en los salones y corredores y fundamentalmente en los puntos de encuentro.

8.4.4.4 Diagrama de flujo de actividades:

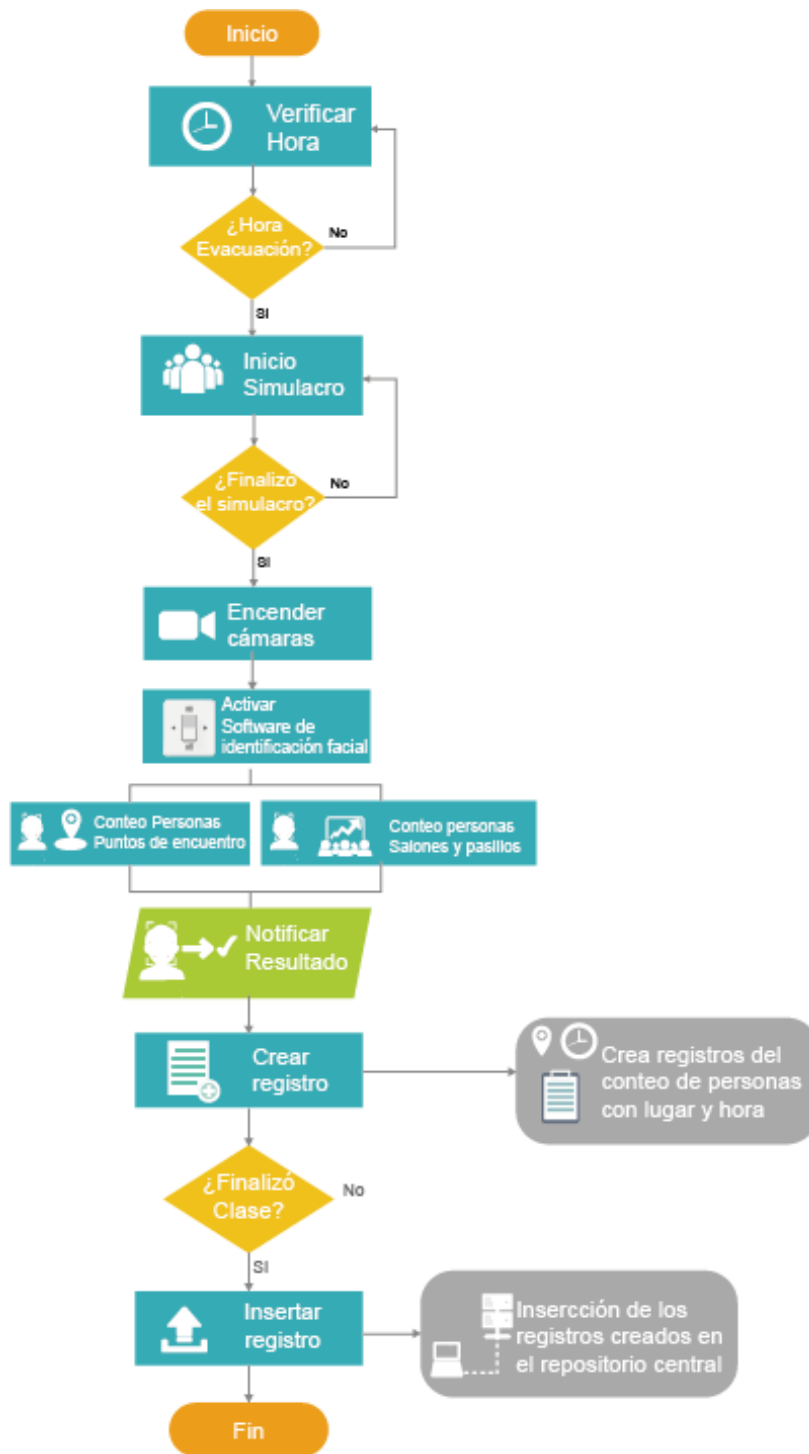


Figura 14 Diagrama de Flujo Registro Conteo Simulacros. Elaboración propia

8.4.4.5 Análisis y explotación de los datos:

El *dashboard* de la figura 15 permite conocer el tiempo en minutos que tomó la evacuación de cada uno de los edificios, así como la cantidad de personas que participaron del simulacro. Los indicadores propuestos para este escenario fueron: Porcentaje de personas evacuadas, tiempo promedio de evacuación y meta de evacuación.

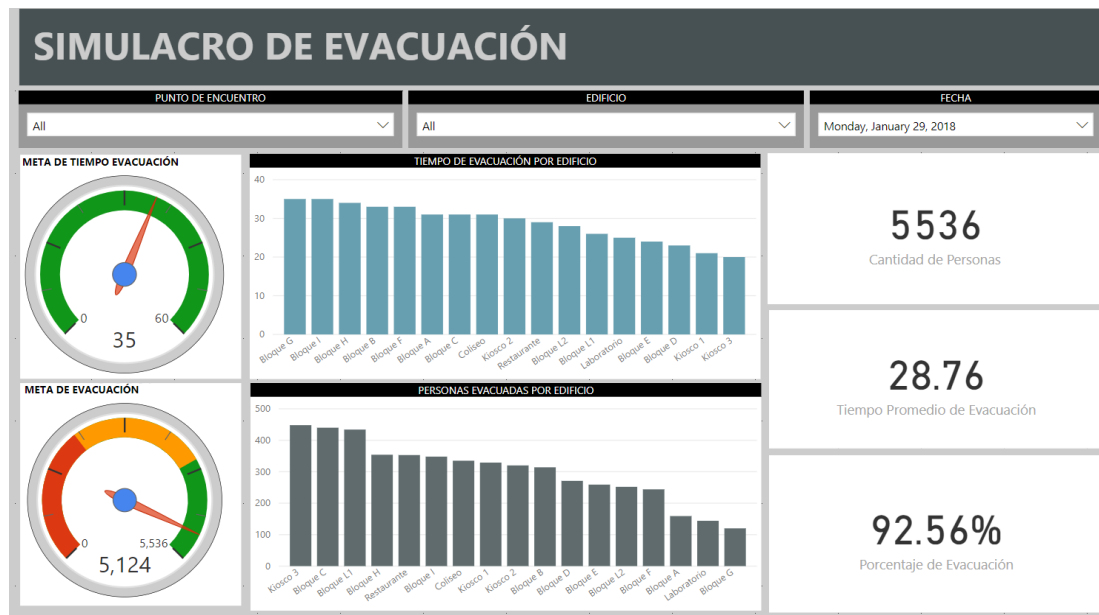


Figura 15 Dashboard de participación en los simulacros de evacuación. Elaboración propia

El uso de la información publicada en el *dashboard* de participación en los simulacros de evacuación permite conocer los resultados del simulacro de manera ágil, y se puede identificar en cuál de los edificios se está tomando más tiempo realizar la evacuación para posteriormente hacer una revisión sobre si se siguen las rutas de evacuación y si están diseñadas de forma óptima para tardar el menor tiempo posible en salir del edificio.

Aunque no se incluyó en el *dashboard* de la figura 15 un dato que es muy importante conocer en este escenario es la meta y la cantidad de personas que llegaron a determinado punto de encuentro, esto debido a que, si no se

alcanza la meta, se tendrían que hacer campañas y eventos que le recuerden al personal de la comunidad educativa cuáles son los puntos de encuentro y dependiendo del edificio en el que se encuentren a cuál deben llegar.

8.4.4.5.1 Costo licencia sistema de visualización:

Power BI *Pro*: \$9.99 USD usuario/mes

9 Prueba Piloto

Para validar la aplicabilidad de este sistema en un ambiente real, se realizó una prueba piloto en una compañía, en donde se quería medir la asistencia a los cursos de certificación que se llevan a cabo en las instalaciones de la compañía, con una periodicidad trimestral y tienen una duración de entre 5 y 10 sesiones, con una intensidad horaria que va de 4 a 2 horas respectivamente.

Los directivos de esta compañía requerían una herramienta que les permitiera visualizar el desempeño de los empleados en los cursos, para tomar decisiones relacionadas con los planes de carrera de cada empleado y evaluar la efectividad de los cursos para decidir si se continuaba apoyando la financiación del curso y el examen de certificación o se reducía el apoyo únicamente al financiamiento del examen de certificación.

El equipo estableció que era válido realizar las pruebas en esta compañía, considerando que los procesos como: asistencia a cursos, ejecución de planes de carrera, entrega de certificados de asistencia a cursos, entre otros, también son llevados a cabo en instituciones de educación superior. Además, se evidenció que se podía implementar la arquitectura propuesta para el sistema, que fue detallada en este documento y que consta de tres etapas: recolección de datos, ETL y visualización de los datos.

El proceso de control de asistencia de la compañía era manual, por lo que al iniciar las sesiones los asistentes debían firmar un formato, el cual era entregado al día siguiente al director del área, quien posteriormente hacía la solicitud de transcripción de los formatos a hojas de cálculo, con las que debía establecer el porcentaje de asistencia y el cumplimiento de la meta de cada empleado. Este proceso, al igual que en los escenarios de asistencia a clase, a exámenes finales y a charlas y eventos, descritos en el documento,

presentaban problemas de pérdida de información, sensibilidad a inyección de errores, demoras en la entrega de la información a los tomadores de decisión debido a reprocesos y ausencias injustificadas.

Para resolver estos problemas, el equipo implemento la arquitectura del sistema propuesto en el proyecto, de la siguiente manera:

1. Determino como fuentes de datos: las bases de datos con información de los empleados, el reporte de notas de los exámenes de certificación, el sistema de solicitud de permisos y para el control de asistencia la base de datos del sistema de biometría dactilar de la compañía.
2. Desarrollo un ETL, para hacer la extracción de los datos que se requerían para la construcción de los reportes, y la transformación y carga de estos a un repositorio central.
3. Finalmente se utilizó la herramienta Power BI, para hacer la visualización de los datos, como se evidencia en la figura 16

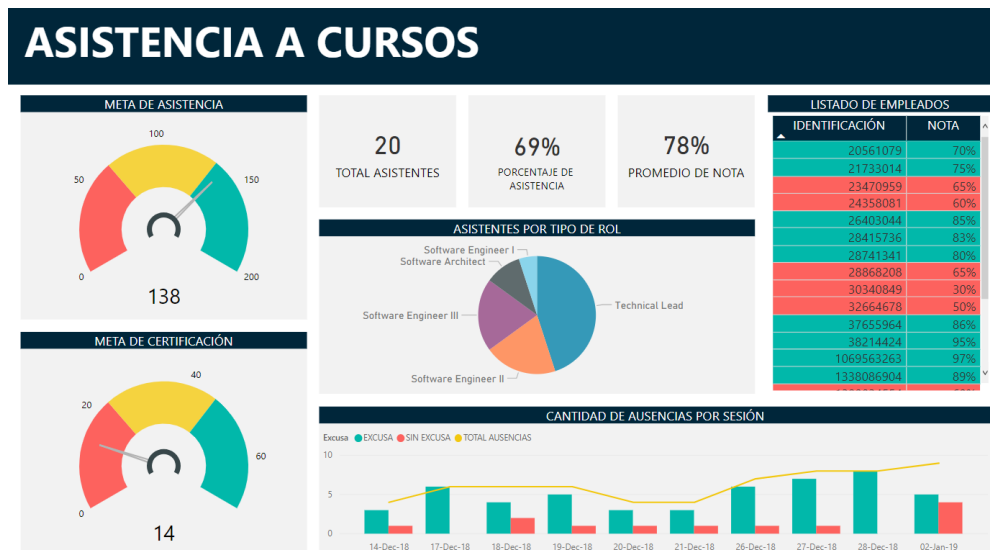


Figura 16 Dashboard del piloto de asistencia a cursos. Elaboración propia

Al finalizar el piloto, se obtuvieron los siguientes resultados:

- Para un grupo de 20 personas el proceso de control de asistencia manual que en promedio tardaba entre 20 y 25 minutos paso a aproximadamente 4 minutos, a través del uso del sistema biométrico.
- Se eliminó el reproceso y la dependencia de terceros para la digitalización de los datos
- Se brindo a los directores la capacidad de conocer el comportamiento de asistencia a los cursos a diario a través de la visualización del *dashboard*.
- Al encontrar que existe una relación fuerte entre la nota y el porcentaje de asistencia al curso, los directores decidieron continuar apoyando la financiación del curso y el examen de certificación.
- Se abrieron canales de comunicación con los empleados que permiten mejorar los cursos y esclarecer las causas de las inasistencias.

Los resultados obtenidos de la prueba piloto comprueban los beneficios que tiene para una compañía implementar el sistema propuesto en el proyecto, considerando que no solo optimiza el proceso del control de asistencia, sino que brinda a los tomadores de decisiones la capacidad de conocer lo que está sucediendo y qué lo está causando, de modo que pueda tomar mejores decisiones.

10 Costo – Beneficio

ESCENARIO	COSTO	BENEFICIO
Asistencia a clases	Equipos biométricos \$300 USD cada uno Power BI Pro: \$9.99 USD usuario/mes	<ul style="list-style-type: none"> - Disminuye el riesgo de deserción escolar, a través del conocimiento de la población escolar y de las causas de inasistencia junto con la periodicidad con la que ocurren. - Proporciona información para mejorar la oferta académica. - Mejoras en la comunicación con los estudiantes a través de - Mejoras en los sistemas de soporte al estudiante, dependiendo del problema que se detecte con un estudiante en particular redirigirlo al área de la institución que le puede prestar ayuda. - Visualización grupal e individual del comportamiento de asistencia a clases.
Asistencia a exámenes	Equipos biométricos \$382 USD cada uno. Power BI Pro: \$9.99 USD usuario/mes	<ul style="list-style-type: none"> - Disminuye el riesgo de suplantación, que en caso de presentarse podría acarrear altos costos jurídicos y reputacionales para la institución. - Disminución del costo de reprogramar exámenes, al permitir tomar acciones inmediatas en caso de que: <ul style="list-style-type: none"> - Los profesores no puedan presentarse la hora y el lugar estipulado. - El salón se encuentre ocupado para tomar acciones inmediatas.
Asistencia a charlas y eventos	Sistema de autenticación: \$90 USD Power BI Pro: \$9.99 USD usuario/mes	<ul style="list-style-type: none"> - Optimización del uso de los recursos de la institución, al conocer los intereses de los estudiantes y los horarios adecuados para programar los eventos. - Para los eventos en los que por ley la institución requiere que se cumpla una cantidad de sesiones y de personal, la herramienta brinda agilidad para conocer los resultados y establecer estrategias que permitan cumplir las metas más rápido
Participación simulacros de evacuación	Costo de licencia del Sistema Biométrico \$890 USD Power BI Pro: \$9.99 USD usuario/mes	<ul style="list-style-type: none"> - Agilidad para la optimización de los procedimientos de evacuación al permitir la identificación rápida de eventos como: <ul style="list-style-type: none"> - Tiempos altos en evacuación de una ubicación específica. - Inconsistencias entre personas registradas al ingreso y participantes del simulacro - Afluencia de personas

Tabla 3 Costo-Beneficio del proyecto. Elaboración propia

11 Conclusiones

Este proyecto evaluó las diferentes opciones de sistemas de visualización de datos y sistemas biométricos, para cuatro escenarios en los que una institución de educación superior requería recolectar de manera automática y ágil datos sobre los participantes de las actividades para posteriormente hacer el análisis a través de la visualización de estos datos.

Considerando que las fuentes de datos son la base del sistema, durante la evaluación y selección de los sistemas biométricos que apoyan la recolección de datos del estudiante, se encontró que además de los requerimientos iniciales que se tenían como: los costos, la facilidad de implementación y la precisión, era necesario identificar los requerimientos de cada escenario. Por lo anterior, para cada escenario el equipo hizo la identificación de las condiciones de entorno y estableció objetivos desde el punto de vista de la recolección de los datos, encontrando que:

- Si se requiere precisión en la detección de la identidad de los estudiantes, lo recomendable es usar un sistema de biometría dactilar.
- Si el proceso no requiere precisión para detectar la identidad, pero requiere velocidad para identificar formas humanas se sugiere el uso de biometría facial.
- Para complementar los procesos biométricos, se recomienda el uso de un lector de código de barras o RFID.

Además se encontró que, la implementación de los dos primeros niveles de la escala de valor de análisis de información en el proyecto, análisis descriptivo y diagnóstico, que se ejecutó a través de la recolección, análisis y visualización de los datos para cada uno de los escenarios descritos en el proyecto, permitió establecer la capacidad que tiene el sistema de brindar a los tomadores de decisiones la información necesaria para determinar ¿qué está sucediendo en

cada escenario? Y ¿por qué? Esto se logró a través de la caracterización de los estudiantes y la visualización de su comportamiento en los diferentes escenarios, para apoyar la toma de decisiones que permitan mejorar las actividades descritas en cada escenario.

Es importante resaltar que debido a que los esfuerzos del equipo estuvieron enfocados en detallar los recursos necesarios para los sistemas de visualización de datos y de biometría que integran el sistema, y que el desarrollo del proyecto se hizo bajo el supuesto de que las instituciones cuentan con la infraestructura necesaria para la operación de sus sistemas, los costos de implementación del proyecto pueden variar considerablemente debido a factores como:

- Necesidad de ejecutar procesos de calidad de datos.
- Desarrollos requeridos para el ETL y la aplicación móvil para efectuar la lectura de códigos de barras o RFID.
- Infraestructura y licenciamiento requerido para el repositorio de datos del sistema, sujeto a las políticas de IT de la institución.

Por otra parte, con el desarrollo del proyecto se evidenció que la implementación de un sistema tecnológico no sólo evita reprocesos, sino que optimiza el proceso en términos de tiempo. Además, en el caso del proyecto el uso de la información biométrica como una fuente de datos y apoyo a la toma de asistencia, brinda un valor agregado en términos de seguridad, ya que evita fraudes en procesos que la institución considera sensibles.

Finalmente, con la ejecución del piloto se evidenció que este proyecto se puede implementar en ambientes diferentes a instituciones educativas, en procesos que necesiten validación de identidad, para tomar decisiones específicas, basado en la información recolectada de los sistemas actuales de la empresa y de los sistemas biométricos.

12 Trabajos Futuros

Al finalizar el proyecto, se encontraron los siguientes puntos que pueden desarrollarse en trabajos futuros:

- Agregar los tipos de análisis predictivo y analítico de modo que se le aporte al sistema la capacidad de predecir eventos y generar estrategias para evitar o hacer que un evento en particular suceda, a partir de los datos recolectados por el sistema.
- En relación al dashboard propuesto para el escenario de ausencias académicas se sugiere que se complemente el perfil del estudiante con datos como: el estado civil, si tiene personas que dependan de él, la forma de financiamiento que usa para costear sus estudios, su estado laboral o incluso su comportamiento basado en su actividad en redes sociales, de manera que se tenga mayor contexto sobre su situación actual y se pueda establecer la causa por la cual se presentan inasistencias y tomar acciones para evitar la deserción de estos alumnos.
- Evaluar los requisitos necesarios para que este proyecto pueda desarrollarse en la asistencia de clases y evaluaciones virtuales, donde es de vital importancia comprobar la identidad del estudiante para evitar suplantaciones o fraudes y mejorar la dinámica académica de este tipo de formación educativa que cada vez tiene más acogida.

13 Bibliografía

- Abas, M. A., Tuck, T. B., & Dahlui, M. (2014). Attendance management system (AMS) with fast track analysis. In *Proceeding - 2014 International Conference on Computer, Control, Informatics and Its Applications: "New Challenges and Opportunities in Big Data", IC3INA 2014*.
<https://doi.org/10.1109/IC3INA.2014.7042597>
- Adeniji, V. O., Scott, M. S., & Phumzile, N. (2016). Development of an Online Biometric-enabled Class Attendance Register System. In *2016 IST-Africa Conference, IST-Africa 2016*.
<https://doi.org/10.1109/ISTAFRICA.2016.7530647>
- Benyo, B., Sódor, B., Doktor, T., & Fördös, G. (2012). Student attendance monitoring at the university using NFC. In *Wireless Telecommunications Symposium*. <https://doi.org/10.1109/WTS.2012.6266137>
- Bowen, E., Price, T., Lloyd, S., & Thomas, S. (2005). Improving the quantity and quality of attendance data to enhance student retention. *Journal of Further and Higher Education*.
<https://doi.org/10.1080/03098770500353714>
- C.O, A., A.O, A., O.O, O., & E.O, I. (2013). Fingerprint-Based Attendance Management System. *Journal of Computer Sciences and Applications*.
<https://doi.org/10.12691/jcsa-1-5-4>
- Curto Díaz, J., & Conesa Caralt, J. (2010). *Introducción al Business Intelligence* (Primera). Barcelona: Editorial UOC, S.L.
- DataCentric. (2017). ¿Qué es el Data Visualization o cómo interpretar grandes cantidades de datos? Retrieved September 15, 2018, from <https://www.datacentric.es/blog/geomarketing/data-visualization-analisis-datos/>

Duró, V. E. (2014). Evaluación de Sistemas de Reconocimiento Biométrico. *Departamento de Electrónica y Automática.*

Esteves Wild, S. (2017). Power BI Pro y Power BI Premium licenciamiento Q&A. Retrieved December 14, 2018, from <https://blogs.technet.microsoft.com/latampts/2017/11/17/power-bi-pro-y-power-bi-premium-licenciamiento-qa/>

Eze Peter, U., Joe-Uzuegbu, C. K. A., Uzoechi, L., & Opara, F. K. (2013). Biometric-based attendance system with remote real-time monitoring for tertiary institutions in developing countries. In *2nd International Conference on Emerging and Sustainable Technologies for Power and ICT in a Developing Society, IEEE NIGERCON 2013 - Proceedings* (pp. 1–8). <https://doi.org/10.1109/NIGERCON.2013.6715633>

Gartner. (2018). Gartner Magic Quadrant. Retrieved December 6, 2018, from <https://www.gartner.com/en/research/methodologies/magic-quadrants-research>

Gobierno de Canadá. (n.d.). Visit Canada. Retrieved November 9, 2018, from <http://www.cic.gc.ca/english/visit/biometrics.asp>

Howson, C., Sallam, R., Richardson, J., Tapadinhas, J., Idoine, C., & Woodward, A. (2018). *Magic Quadrant for Analytics and Business Intelligence Platforms.*

IDEMIA. (n.d.). Biometrics. Retrieved November 12, 2018, from <https://www.morpho.com/en/biometrics>

Iniciativa Aporta. (2016). *Visualización de datos: Definición, tecnologías y herramientas.* Retrieved from http://datos.gob.es/sites/default/files/doc/file/informe_herramientas_visualizacion.pdf

- Jain, A. K., Ross, A., & Prabhakar, S. (2004). An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*. <https://doi.org/10.1109/TCSVT.2003.818349>
- Jha, A. (2007). Class Room Attendance System Using Facial Recognition System. *The International Journal of Mathematics Science, Technology and Management*.
- Khatun, A., Haque, A. K. M. F., Ahmed, S., & Rahman, M. M. (2015). Design and implementation of iris recognition based attendance management system. In *2nd International Conference on Electrical Engineering and Information and Communication Technology, iCEEICT 2015*. <https://doi.org/10.1109/ICEEICT.2015.7307458>
- Levshankova, C., Hiron, D., Kirton, J. A., Knighting, K., & Jinks, A. M. (2018). Student nurse non-attendance in relation to academic performance and progression. *Nurse Education Today*. <https://doi.org/10.1016/j.nedt.2017.10.006>
- Loshin, D. (2013). *Business Intelligence: The Savvy Manager's Guide*. Morgan Kaufmann. <https://doi.org/10.1111/jace.13474>
- Lukas, S., Mitra, A. R., Desanti, R. I., & Krisnadi, D. (2016). Student attendance system in classroom using face recognition technique. In *2016 International Conference on Information and Communication Technology Convergence (ICTC)* (pp. 1032–1035). <https://doi.org/10.1109/ICTC.2016.7763360>
- Mohd Ali, S., Gupta, N., Lenka, R., & Krishna Nayak, G. (2016). Big Data Visualization: Tools and Challenges.
- Mshelia, D. E., Alkali, H. A., Isuwa, S., Dibal, E., & Onogwu, C. (2017). An RFID and Fingerprint Automated Attendance System. *IOSR Journal of*

Computer Engineering (IOSR-JCE), 19(4), 75–84.
<https://doi.org/10.9790/0661-1904067584>

Muir, J. (2009). Student Attendance: Is It Important, and What Do Students Think? *Transactions*. <https://doi.org/10.11120/tran.2009.06020050>

Rosales Cruz, A. (2001). *Clasificación de huellas digitales mediante minucias*. Retrieved from https://ccc.inaoep.mx/~esucar/Clases-mgp/Proyectos/reporte_modelos_huellas.pdf

Said, M. A. M., Misran, M. H., Othman, M. A., Ismail, M. M., Sulaiman, H. A., Salleh, A., & Yusop, N. (2014). Biometric attendance. In *ISTMET 2014 - 1st International Symposium on Technology Management and Emerging Technologies, Proceedings* (pp. 258–263). <https://doi.org/10.1109/ISTMET.2014.6936516>

Simon, D., García Gomar, M., Sánchez Asenjo, M., L Sánchez Bote, J., & Ortega García, J. (2018). ESQUEMA COMPLETO DE IDENTIFICACIÓN Y VERIFICACIÓN DE PATRONES BIOMÉTRICOS DE HUELLAS DACTILARES.

Singh, N., Agrawal, A., & Khan, P. R. (2018). Voice Biometric: A Technology for Voice Based Authentication. *Advanced Science, Engineering and Medicine*, 10. <https://doi.org/10.1166/ asem.2018.2219>

Sinnexus. (n.d.-a). Datamart. Retrieved October 5, 2018, from https://www.sinnexus.com/business_intelligence/datamart.aspx

Sinnexus. (n.d.-b). Datawarehouse. Retrieved October 5, 2018, from https://www.sinnexus.com/business_intelligence/datawarehouse.aspx

Sulaiman, A. A., Bakar, M. S. A., Noor, M. Z. H., & Abdullah, S. A. C. (2015). Easy Access Attendance Management System (EAMS). In *ICEED 2014 -*

2014 IEEE 6th Conference on Engineering Education.
<https://doi.org/10.1109/ICEED.2014.7194697>

Technovelgy. (n.d.). Biometric Match. Retrieved November 12, 2018, from
<http://www.technovelgy.com/ct/Technology-Article.asp?ArtNum=82>

Trabelsi, Z., & Shuaib, K. (2011). IMPLEMENTATION OF AN EFFECTIVE AND SECURE BIOMETRICS-BASED STUDENT ATTENDANCE SYSTEM. *International Journal of Computers and Applications*.
<https://doi.org/10.2316/Journal.202.2011.2.202-2928>

Vazquez, H., Chang, L., Rizo, D., & Morales-González, A. (2012). Evaluación de la calidad de las imágenes de rostros utilizadas para la identificación de las personas. *Computación y Sistemas*, 16, 147–165.

Wang, T. (2016). Big Data Needs Thick Data. *Ethnography Matters*.

Yang, S., Song, Y., Ren, H., & Huang, X. (2016). An automated student attendance tracking system based on voiceprint and location. In *ICCSE 2016 - 11th International Conference on Computer Science and Education* (pp. 214–219). <https://doi.org/10.1109/ICCSE.2016.7581583>

14 Anexos

Anexo 1. Cronograma

Anexo 2. VeriFinger SDK Brochure

Anexo 3. Face Verifaction SDK Brochure

Anexo 4. VeriEye SDK Brochure,

Anexo 5. VeriSpeak SDK Brochure

Anexo 6. NCheck Bio Attendance Brochure

Anexo 7. SentiVeillance SDK Brochure