Tablero de Control de Gestión Educativa

Alan GABITO, Carlos ARAKAKI (Universidad de Palermo - Argentina)

Abtract:

Hoy en día se utilizan las herramientas de BI para el análisis de los datos dentro de diferentes organizaciones que desean explotar la información con diversos fines, en las instituciones educativas de nivel medio generalmente no se registran este tipo de monitoreo o de seguimiento, por lo menos de manera amplía y sistemática.

El presente proyecto es un acercamiento a los sistemas analíticos de BI. La finalidad del mismo, es acercarnos al funcionamiento de los análisis BI, para lo cual se desarrolló un Tablero de Control de indicadores estratégicos para la gestión educativa de nivel inicial, primario, secundario. Con esto se busca exponer los principales indicadores estratégicos que ayudarán, a través de la asistencia en la toma de decisiones, a mejorar los índices académicos analizados.

Palabras claves: Deserción, disciplina, rendimiento.

I NOMENCLATURA

ETL:Extract-Transform-Load.
DW: Data Warehouse.
DM: Data Mart.
BI: Business intelligence.

II INTRODUCCIÓN

La Dirección de Educación Preuniversitaria es una organización que ejerce el control funcional de las Unidades Educativas de nivel preuniversitario (7 instituciones). La misma cuenta con información variada en cuanto a la situación académica, disciplinaria y médica de los alumnos, pero esta se encuentra en diferentes formatos, administrada por diversas secretarías o dependencias dentro de las instituciones, y no se cuenta con una herramienta que permita la integración de esta información para poder llevar a cabo el seguimiento de la situación actual de los alumnos, y la posterior implementación de medidas para mejorar dicha situación. A su vez, la consolidación actual de dicha información y la generación de informes conlleva un monto de tiempo considerable, lo cual hace que la implementación de medidas y políticas de acuerdo a diferentes situaciones no generen el impacto que deberían y, en muchos casos, que no produzcan resultados beneficiosos . Es a partir de estos puntos que se establece la necesidad de mejorar a nivel institucional en estos aspectos para llevar adelante una planificación estratégica situacional.

III CASO DE ESTUDIO

Actualmente, la Direccion de Educacion Preuniversitaria inició un proceso de actualización de sistemas informáticos, partiendo de la base de sistemas individuales para cada una de las 7 instituciones que administra, con el consiguiente problema de no contar con una base de datos centralizada, lo que llevaba a diferentes inconvenientes a la hora de medir el rendimiento académico y disciplinario de los alumnos. Esto generaba serios problemas a la hora de consolidar información, realizar un análisis del estado disciplinario y un control del desempeño estudiantil, evaluar el rendimiento de los profesores y la tasa de deserción, entre otros problemas.

Con la modernización de la plataforma de gestión educativa, se consolidó la administración de dichos establecimientos en un único Sistema de Gestión, lo cual permitió subsanar el problema de consolidación de la información. A partir de esta situación, se vislumbraron una serie de nuevos inconvenientes debido a no contar con información en tiempo y formas sobre la actual tasa de deserción de los institutos y como la misma evolucionó con respecto a los años anteriores. A su vez, se necesita contar con la evolución de la misma a partir de la implementación de diferentes planes de estudio y cómo este cambio afecta a dicha tasa. Por otro lado, se planteó el problema de poder analizar la evolución del estado disciplinario de los alumnos por nivel académico (Primario, Secundario) y por clase, asi como tambien ver el grado de gravedad de las faltas disciplinarias, tendiente a identificar patrones repetitivos que requieren un cambio en las políticas institucionales para ser subsanados. Por último, tomó especial relevancia el poder medir el grado de rendimiento post pandemia de los alumnos, para así poder tomar medidas que subsanen estos inconvenientes. Para ello se busca analizar el rendimiento a nivel intelectual y físico de los alumnos, antes y después de la pandemia, teniendo en cuenta si hubo o no una disminución del mismo en el periodo pandémico y cómo evolucionó en el periodo del primer año post pandemia.

IV MARCO TEÓRICO

ETL[1]: Extraer, transformar y cargar (ETL) es una canalización de datos utilizada para recopilar datos de varias fuentes. Luego transforma los datos de acuerdo con las reglas comerciales y carga los datos en un almacén de datos de destino. El trabajo de transformación en ETL se lleva a cabo en un motor especializado y, a menudo, implica el uso de tablas provisionales para almacenar temporalmente los datos a medida que se transforman y, en última instancia, se cargan en su destino, según Fig 1.

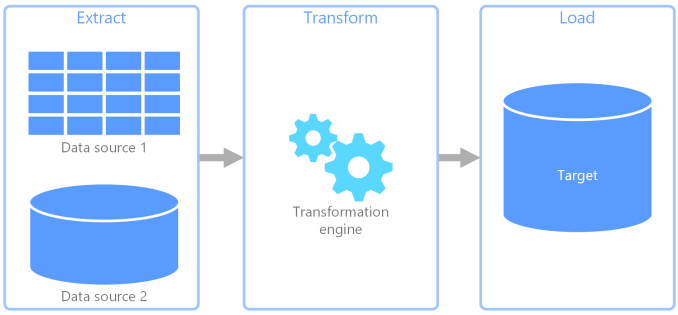


Fig. 1. La transformación de datos que tiene lugar suele implicar varias operaciones, como filtrado, clasificación, agregación, unión de datos, limpieza de datos, duplicación y validación de datos.

DW[3]: Un almacenamiento de datos es un repositorio centralizado que almacena datos estructurados (tablas de base de datos, hojas de Excel) y datos semiestructurados (archivos XML, páginas web) para la creación de informes y el análisis. Los datos fluyen desde una variedad de orígenes, como sistemas de punto de venta, aplicaciones empresariales y bases de datos relacionales y, normalmente, se limpian y normalizan antes de llegar al almacén. Dado que un almacenamiento de datos puede almacenar grandes cantidades de información, proporciona a los usuarios acceso fácil a una gran cantidad de datos históricos que se pueden usar para minería de datos, visualización de datos y otras formas de creación de informes de inteligencia empresarial.

DM[2]: Los datamarts ofrecen una capa semántica que se genera y se sincroniza de forma automática con el contenido de las tablas de datamart, su estructura y los datos subyacentes, todo ello en un conjunto de datos que se genera automáticamente. Esta generación y sincronización automáticas le permiten describir mejor el dominio de los datos con elementos como jerarquías, nombres descriptivos y descripciones, así como establecer un formato específico de la configuración regional o los requisitos empresariales. Con los datamarts también puede crear medidas y métricas estandarizadas para los informes. Power BI (y otras herramientas de cliente) pueden crear objetos visuales y proporcionar resultados para estos cálculos en función de los datos en contexto. Según Fig. 2.

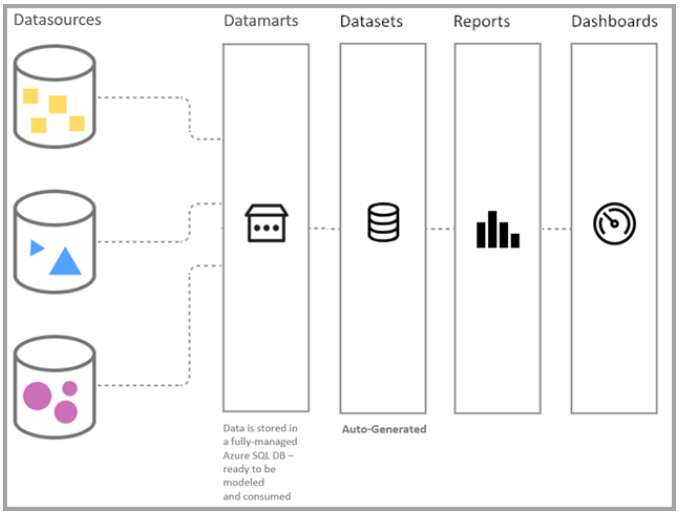


Fig. 2. Los conjuntos de datos generados automáticamente son distintos de los conjuntos de datos tradicionales de Power BI

BI[4]: Business Intelligence es un tipo de software que se alimenta de datos de negocios y presenta reportes, paneles, tablas y gráficos de forma amigable para el usuario.

MODELO ESTRELLA[5]: el Modelo Estrella es un tipo de Modelo de Base de Datos relacional que consta de una sola tabla de hechos central rodeada de tablas de dimensiones.

MODELO COPO DE NIEVE[6]: El modelo copo de nieve consta de una tabla de hechos que está conectada a muchas tablas de dimensiones, que pueden estar conectadas a otras tablas de dimensiones a través de una relación de muchos a uno. Las tablas de un modelo copo de nieve generalmente se normalizan en el tercer formulario de normalización. Cada tabla de dimensiones representa exactamente un nivel en una jerarquía.

MODELO CONSTELACIÓN[7]: Un modelo de constelación es una combinación de un esquema de estrella y un esquema de copo de nieve. Los modelos de constelación son modelos de copo de nieve en los que sólo algunas de las tablas de dimensiones se han desnormalizado. El objetivo de los modelos de constelación es aprovechar las ventajas de los modelos de estrella y de copo de nieve. Las jerarquías de los modelos de estrella están desnormalizadas, mientras que las jerarquías de los modelos de copo de nieve están normalizadas. Los modelos de constelación están normalizados para eliminar las redundancias de las dimensiones. Para normalizar el esquema, las jerarquías dimensionales compartidas se colocan en outriggers.

POWER BI[8]: es un servicio de análisis de datos de Microsoft orientado a proporcionar visualizaciones interactivas y capacidades de inteligencia empresarial con una interfaz lo suficientemente simple como para que los usuarios finales puedan crear por sí mismos sus propios informes y

paneles.Entre sus características se destacan la interoperabilidad nativa con Active Directory y su gran biblioteca de conectores de datos.

TABLEAU[9]: Es una plataforma de visualización de datos interactivos, enfocados en BI. Esta herramienta consulta bases de datos relacionales, cubos de procesamiento analítico en línea, bases de datos en la nube y hojas de cálculo para generar visualizaciones de datos de tipo gráfico. El software también puede extraer, almacenar y recuperar datos de un motor de datos en memoria

SQL SERVER[10]: Es un sistema de gestión de base de datos relacional, desarrollado por la empresa Microsoft. El lenguaje de desarrollo utilizado por este es Transact-SQL (TSQL), una implementación del estándar ANSI del lenguaje SQL, utilizado para manipular y recuperar datos, crear tablas y definir relaciones entre ellas. Entre sus características más notables se encuentran el soporte de transacciones y procedimientos almacenados, poseer un entorno gráfico de administración y que permite administrar información de otros servidores de datos. A su vez, posee varios servicios que se ejecutan en la memoria del servidor por parte del sistema del mismo, aprovechando los recursos con los que cuenta y pudiendo programar tareas que se ejecuten sin la necesidad del control de un cliente. Su uso y funcionalidades están limitadas a partir de cierto nivel por las licencias con las que se cuentan del mismo y por las diferentes ediciones del motor.

V Desarrollo

A. Objetivo del producto

Requerimientos Funcionales:

1) La solución se ajusta al modelo establecido.
2) La solución contiene una dimensión temporal de forma mensual.
3) Las cuantías de tipo numérico tienen al menos dos decimales.
4) La solución mantiene, en todo momento, el tipo de formato de los datos de entrada y salida.
5) La solución permite visualizar la tasa de deserción de los institutos.
6) La solución permite visualizar la evolución del estado disciplinario de los alumnos por nivel académico (Primario, Secundario).
7) La solución permite analizar el rendimiento a nivel intelectual y físico de los alumnos, antes y después de la pandemia.
8) La solución permite generar informes que contendrán gráficos, que cambiarán con los filtros, para facilitar la comprensión de los datos.

B. Requerimientos No Funcionales:

- Rendimiento:

  Se podrá acceder a la solución a través de un navegador web mediante cualquier dispositivo, tanto en Tablet como en celulares.

  La solución deberá ser intuitiva (que cualquiera que acceda entienda la información que allí se ve y pueda interpretar la misma) para los usuarios, estos estarán comprendido en el rango de los 18 a los 65 años de edad promedio.

- Escalabilidad:

  En primera instancia, se contará con la información de los 7 institutos que administra la Dirección de Educación Preuniversitaria, pudiendo incorporar otras instituciones dependientes de la órbita educativa superior o de otras reparticiones.

  También se contempla la incorporación de otros orígenes de datos a los procesos ETL para la adquisición de información para alimentar a los modelos definidos.

- Integridad de información:

  Se asegura la calidad de los datos obtenidos mediante los procesos ETL generados que conforman los diferentes modelos que integran la solución.

- Seguridad:

  Los usuarios serán gestionados por un administrador, que tendrá los privilegios máximos en la solución, y estos tendrán únicamente acceso de lectura a la solución.

- Usabilidad:

  La solución abrirá el menú de informes en un tiempo máximo de 5 segundos.

  Un usuario experimentado realizará una solicitud de visualización de informe en un tiempo promedio de 30 segundos y máximo de 1 minuto, en el 95 % de las veces.

  Un usuario nuevo realizará una solicitud de visualización de informe en un tiempo promedio de 1 minuto y máximo de 2 minutos, en el 95 % de las veces.

- Disponibilidad:

  La solución debe tener una disponibilidad del 99% de las veces en que un usuario intente acceder.

  La tasa de tiempos de falla del sistema no podrá ser mayor al 0,5% del tiempo de operación total.

- Privacidad de la información:

La solución deberá mantener disociados los datos exhibidos. evitando vincular los resultados con un alumno en particular.

C. *Tecnología*

- Data Warehouse:. La tecnología utilizada para dar soporte al datawarehouse es SQL Server 2019. Las estructuras de los diferentes modelos como la información que los compone está almacenada en los servidores de la Dirección de Educación Preuniversitaria. Microsoft SQL Server permite la gestión de base de datos mediante la aplicación SQL Server Management Studio.
- ETL/Visualización: Para la gestión de los proceso ETL y el armado de los tableros se utilizara PowerBI. Esta herramienta cuenta con diferentes tipos de gráficos que permiten la visualización de información y el entrecruzamiento de la misma de forma sencilla. A su vez, permite la realización de procesos ETL mediante herramientas integradas consumiendo un gran número de orígenes de datos.

D. *Experiencia*

Durante la etapa inicial de la solución, se tomó en cuenta la existencia de un sistema de gestión previo, el cual cuenta con una base de datos SQL Server 2019. En base a esto, se realizó el análisis de la estructura de la misma y se identificaron las diferentes tablas que son relevantes en la construcción de la solución para la problemática planteada.

A partir de la identificación del origen de datos, se construyó el modelo de datos necesario, optando por un modelo estrella, el cual permite mantener una cierta jerarquización de los datos a la vez que se construyen las diferentes tablas de hechos que contengan las métricas que correspondan a cada una de las problemáticas planteadas.

Se configuraron diferentes procedimientos almacenados en la base de datos que van a ir poblando los diferentes modelos construidos, adicionando los registros tanto a las tablas de hechos, como así también a las diferentes tablas de dimensiones, permitiendo generar un registro histórico de hechos para ser analizados mediante el tablero. Dichos procedimientos se configuran para que el motor los ejecute automáticamente en un horario donde no afecte el funcionamiento de la base de datos del sistema de gestión.

Para el diseño del tablero se consideraron dos herramientas, PowerBI y Tableau. Mientras que Tableau puede operar con una mayor cantidad de datos en comparación con Power BI, su interfaz es más difícil de utilizar por usuarios con poca experiencia, ya que está pensado para ser usado por usuarios con un gran conocimiento del análisis de datos. En este sentido, Power BI se destaca por ser mucho más intuitivo, además de poder conectarse con una gran cantidad de orígenes de datos, lo cual le da una gran flexibilidad para el desarrollo de la solución. En cuanto a licencias para cada uno, Tableau requiere una inversión mayor que su alternativa, lo cual hace que Power BI sea nuestra elección para el desarrollo, teniendo en cuenta que además contamos con una licencia para el mismo en el servidor donde generamos nuestro DW.



Fig. 3. Diagrama de base de datos.

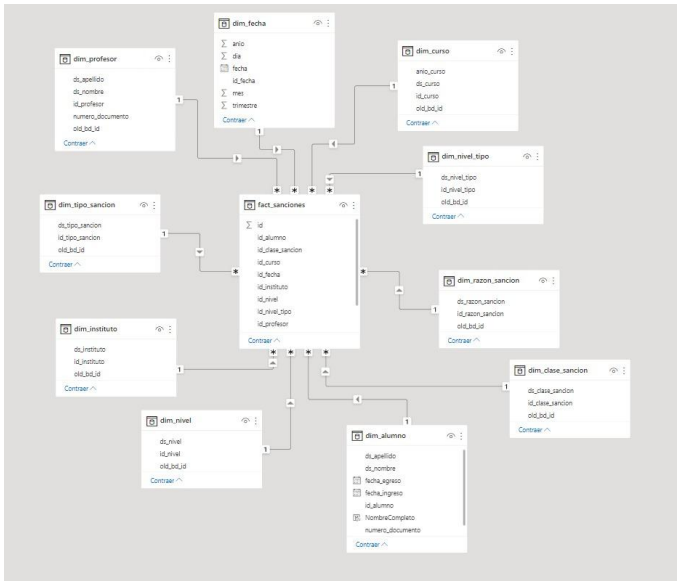


Fig. 4. Modelo estrella de sanciones.

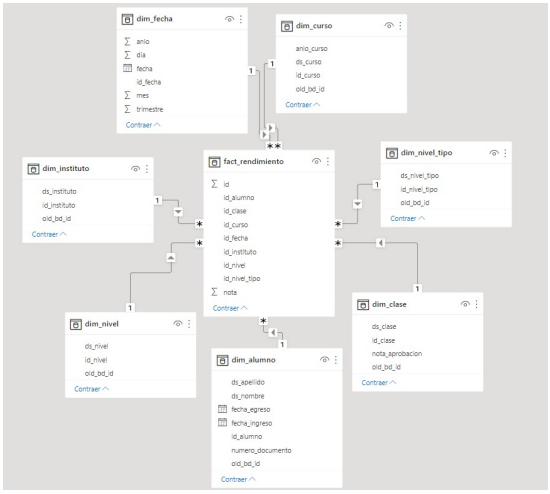


Fig.5.Modelo Estrella de Rendimiento Académico.

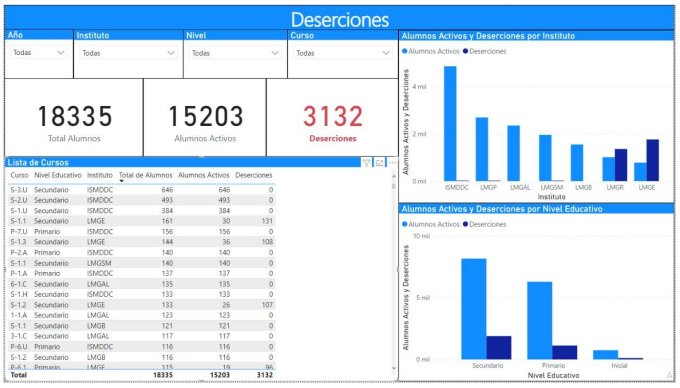


Fig.8. Vista de deserciones.

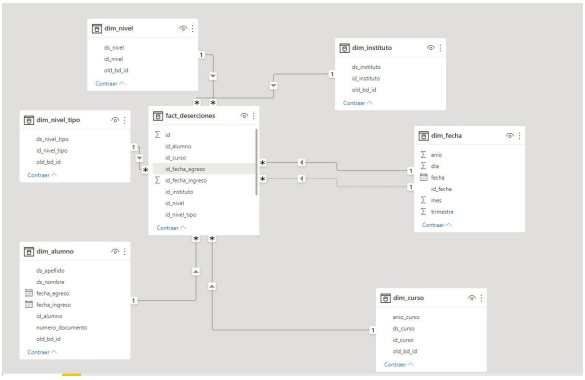


Fig. 6.Modelo estrella de deserciones.

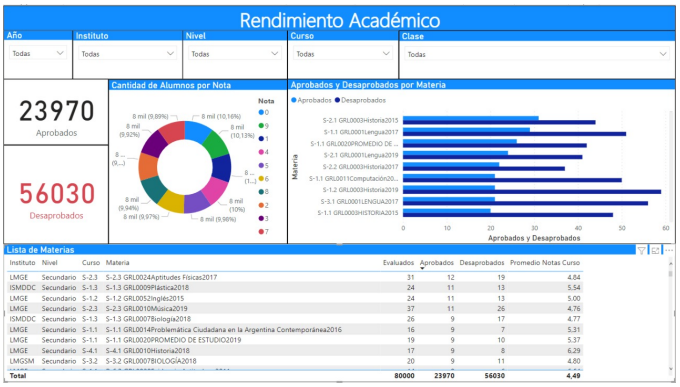


Fig.8.Vista de Rendimiento Académico.

Diseño del tablero:



Fig.7. Vista de estado disciplinario.

E. ARQUITECTURA DE LA SOLUCIÓN

La arquitectura de una Solución de BI presenta una serie de pasos que pueden llegar a omitirse dependiendo de la persona que la construya.

Estas soluciones parten del análisis de todos los sistemas y sus bases de datos, tanto internas o externas, las cuales son utilizadas por la organización a analizar. En nuestro caso partimos del sistema de gestión académica existente. Luego llevamos a cabo las fases de extracción, transformación y carga de datos, en donde contestamos 3 preguntas bases:

1. ¿Cuál es la información requerida para gestionar y tomar decisiones?

2. ¿Cuál debe ser el formato y comprensión de los datos a utilizar?

3. ¿De dónde proceden esos datos y cuál es la disponibilidad y periodicidad requerida?.

Esta etapa suele apoyarse en la conformación de diferentes DM en los cuales se almacenaran los diferentes modelos definidos, y cuyo

principal objetivo consiste en evitar la saturación de servidores funcionales y el acceso rápido y adecuado de los datos en el formato requerido.

La información resultante ya unificada, depurada y consolidada, la guardamos en un almacén de datos corporativo, esta se caracteriza por poseer la estructura óptima para el análisis de los datos, para lo cual utilizamos Power Bi para realizar el análisis de la información a través del tablero de control.

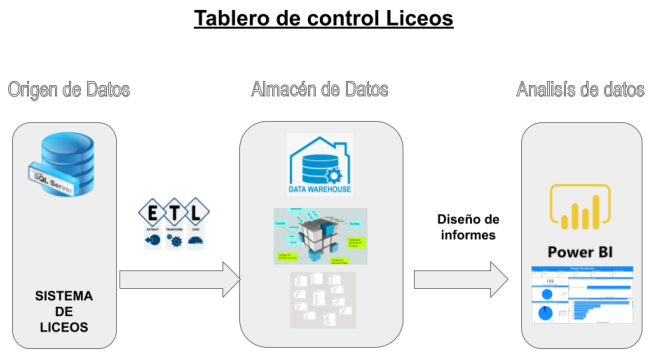


Fig. 9. Arquitectura de la solución

VI Conclusión

En el análisis de situación inicial se encontró que la institución no tenía conocimiento de estrategias de inteligencia del negocio. Las encuestas realizadas permitieron que el personal del mismo reconociera la importancia de este tema, para la medición del desempeño académico.

Dadas las limitaciones en el análisis de la información de la gestión Académica, y centrándose en el problema de la deserción a nivel de los estudiantes, el bajo rendimiento escolar y la elevada incidencia de sanciones, se logró crear una propuesta de solución adecuada a los requerimientos de información que los directivos y jefes de docencia necesitan para identificar los riesgos de deserción, distintos patrones de conducta y evolución académica, y tomar decisiones para enfrentar dicho problema con una base de fundamentos clara y concisa.

Las herramientas tecnológicas de BI, como las utilizadas en el presente trabajo pueden ser utilizadas por cualquier institución de educación media, con el fin de monitorear el desempeño académico de los estudiantes. La definición de indicadores clave para la institución a nivel académico, permite tomar decisiones a tiempo e implementar estrategias de mejora con el fin de poder monitorear el desempeño académico de los estudiantes.

VII Líneas futuras de investigación

Se sugiere tomar esta investigación como prueba piloto para otras instituciones a nivel nacional, con el fin de realizar el monitoreo académico por medio de estrategias de BI.

Con el fin de usar efectivamente la herramienta tecnológica de análisis de información, es necesario capacitar a los docentes y empleados administrativos en el uso de la herramienta.

Se pueden aplicar en los colegios Primarios y Secundarios para monitorear el desempeño académico obtenido, con el fin de mejorar el rendimiento académico.

VIII BIBLIOGRAFÍA

[1] ETL:

https://docs.microsoft.com/en-us/azure/architecture/data-guide/relational-data/etl

[2] DM:

https://docs.microsoft.com/es-es/power-bi/transform-model/datamarts/datamarts-understand

[3] DW:

https://es.wikipedia.org/wiki/Miner%C3%ADa_de_datos

[4] BI:

https://www.ibm.com/ar-es/topics/business-intelligence

[5] Modelo Estrella:

https://www.ibm.com/docs/es/ida/9.1.2?topics=schemas-star

[6] Modelo Copo de Nieve:

https://www.ibm.com/docs/es/ida/9.1.2?topic=schemas-snowflake

[7] Modelo Constelación:

https://www.ibm.com/docs/es/ida/9.1.2?topic=schemas-starflake

[8] Power BI:

https://es.wikipedia.org/wiki/Power_BI
https://powerbi.microsoft.com/es-es/what-is-power-bi/

[9] Tableau:

https://en.wikipedia.org/wiki/Tableau_Software
https://www.tableau.com/es-es/why-tableau/what-is-tableau

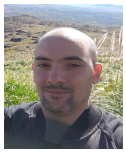
[10] Sql Server:

https://es.wikipedia.org/wiki/Microsoft_SQL_Server
https://www.microsoft.com/es-ES/sql-server

IX. BIOGRAFÍA DE LOS AUTORES



Carlos Eduardo Arakaki. (Leg.:99459) nacido el 13 de septiembre de 1985 en la localidad de Pilar, provincia de Buenos Aires, Argentina. Tiene el título de Técnico Superior en Informática . Actual estudiante de Licenciatura en Tecnología de la Información en la Universidad de Palermo.



Alan Gabito. (Leg.:119519) nacido el 18 de abril de 1989 en la Ciudad Autónoma de Buenos Aires, Argentina. Tiene el título de Técnico Superior en Desarrollo de Software . Actual estudiante de Licenciatura en Tecnología de la Información en la Universidad de Palermo.