

## **Evaluación y selección de herramientas de analítica visual para su implementación en una institución de educación superior**

### **Evaluation and selection of visual analytical tools for its implementation in a higher education institution**

Johnny Alexander Salazar Cardona<sup>1</sup>  
David Alberto Angarita<sup>2</sup>

Recibido: 15/02/2017 - Aceptado: 12/06/2017

Cómo citar este artículo: [#] J. Salazar y D. Angarita. “Evaluación y selección de herramientas de analítica visual para su implementación en una institución de educación superior”, *IngEam*, vol. 4, n.º 1 pp. 1- 20, 2017

#### **Resumen**

En el marco de trabajo de la Institución Universitaria EAM en Quindío - Colombia, donde se ejecutan proyectos de investigación con el fin de fomentar procesos de desarrollo en la institución y fortalecer los grupos de investigación, se ha establecido un proyecto por parte del programa de ingeniería de software titulado “Aplicación de un prototipo de analítica visual en el programa de ingeniería de Software para apoyar la toma de decisiones”, el cual busca implementar procesos de descubrimiento de conocimiento con analítica visual, a partir de open data o datos abiertos. Entre los objetivos planteados para la ejecución del proyecto, se encuentra la realización de un análisis sistémico de trabajos relacionados, tecnologías (Para analítica visual con enfoque open data) o enfoques (Metodologías disponibles en ámbitos universitarios) para la elaboración de Dashboards o tableros de mando donde la comunidad universitaria podrá realizar sus propios análisis sobre estos reportes predefinidos. Para esto se debe tener en cuenta diferentes factores, entre ellos el contexto en el que se desenvuelve la Institución Universitaria, la cual es similar a cualquier institución de educación superior en Colombia, con el fin de elegir el enfoque más adecuado para la ejecución exitosa del proyecto. Cabe destacar que lo expuesto en este artículo solo aborda el análisis de las tecnologías o herramientas disponibles y que fueron evaluadas para su implementación en un ambiente universitario.

**Palabras clave:** - Analítica visual, descubrimiento de conocimiento, dashboard, toma de decisiones, open data.

---

<sup>1</sup> Ingeniero de Sistemas y Computación – Institución Universitaria EAM. Correo electrónico: alexander9052@gmail.com

<sup>2</sup> Ingeniero de Software - Institución Universitaria EAM. Correo electrónico: davidangaritag@gmail.com

## Abstract

Within the framework of the EAM University institution in Quindío-Colombia, where research projects are carried out in order to Promote development processes in the institution and strengthen the research groups, a project has been established by the software engineering program titled "Application of a prototype of visual analytics in the Software engineering program To support decision-making ", which seeks to implement knowledge discovery processes with visual analytics, from Open data. Among the objectives proposed for the execution of the project, is the realization of a systemic analysis of related works, technologies (for visual analytics with Open data approach) and/or approaches (methodologies available in university fields ) for the elaboration of dashboards or scorecards where the university community can make their own analyses on these predefined reports. For this you should take into account different factors, including the context in which the university institution, which is similar to any institution of higher education in Colombia, in order to choose the most appropriate approach to the Successful implementation of the project. It should be noted that this article only deals with the analysis of the technologies or tools available that were evaluated for their implementation in a university environment.

**Key words:** visual analytics, Knowledge discovery, dashboard, decision making, open data.

## I. INTRODUCCIÓN

Las instituciones nacionales de educación superior deben realizar procesos constantes de investigación, ejecutando proyectos que impacten de manera positiva algún área de interés con el fin de generar impacto o conocimiento que pueda ser refinado o replicado. La Institución Universitaria EAM no es ajena a este ámbito, y desde sus programas se gestionan diferentes proyectos de investigación. Este es el caso del programa de ingeniería de software que, basado en su sublínea de investigación se ha establecido una investigación denominada “Aplicación de un prototipo de analítica visual en el programa de ingeniería de Software para apoyar la toma de decisiones” orientada a impactar positivamente el campus académico a partir de la toma de decisiones. Este se encuentra enfocado en la implementación de analítica visual en el programa de ingeniería de software con el fin de realizar análisis descriptivos sobre los datos de los estudiantes del programa. Este proyecto surgió de la necesidad continua que tiene la institución universitaria de tomar decisiones respecto a sus estudiantes, y para lograrlo, aplica procesos de reingeniería que entorpecen la libre ejecución de estos análisis, retrasando los procesos y sin obtener resultados de manera inmediata.

El proyecto hace uso de todos los conceptos de BI (Business Intelligence) o Inteligencia de Negocios con el fin de focalizar esfuerzos respecto a la elaboración de informes dinámicos predefinidos o Dashboards, con los cuales los usuarios pueden interactuar y realizar sus propios análisis. Estos se encuentran orientados a partir de unos indicadores claves de rendimiento o KPI (Key Performance Indicator) los cuales orientan la elaboración de estos informes.

Para la implementación exitosa de este proyecto, se definieron una serie de objetivos que encaminara el cumplimiento satisfactorio de este. El primer objetivo está enfocado a la comparación de las diferentes herramientas y enfoques disponibles de implementación de los Dashboards o tableros de mando públicos, a partir de una revisión sistémica sobre implementaciones similares y herramientas existentes. Luego se identificarán las diferentes fuentes de datos disponibles en la institución universitaria, seguido de la selección de la metodología KDP (Knowledge Discovery Process) para el procesamiento y análisis de los datos, y finalmente pasar a la implementación de los Dashboards o tableros de mando digitales para su libre acceso por parte de la comunidad académica. Cabe destacar que, de los objetivos anteriormente planteados solo se abordara la evaluación y selección de herramientas disponibles, siendo la metodología de aplicación de procesos de descubrimiento de conocimiento innecesario en este punto, ya que sin importar la herramienta de elaboración de los Dashboards se puede aplicar cualquier metodología que se elija en etapas posteriores del proyecto. Adicionalmente se debe aclarar el hecho que, aunque el proyecto hará uso de alguna metodología KDP, solo se aplicara analítica visual, el cual es solo un segmento en toda la aplicación de minería de datos, que a su vez es solo una etapa de todo el proceso KDP.

El valor agregado del proyecto es el enfoque hacia Datos Abiertos u Open Data, el cual es la liberación o publicación de datos no sensibles para que cualquier interesado, en este caso para que la comunidad universitaria pueda acceder a los datos y analizarlos, brindándoles Dashboards predefinidos para su análisis. Esto permitirá a cualquier miembro poder realizar estudios sobre estos, tomar decisiones con históricos de datos y divulgar sus resultados, siendo estos de estudiantes del programa de ingeniería de software.

## II. MARCO TEÓRICO

Se realizó una búsqueda sistémica a partir de una cadena de búsqueda estructurada para filtrar al máximo detalle los resultados (ver figura 1). Esta búsqueda arrojó un total de 152 artículos científicos, con mayor concentración en algunos países específicos (ver figura 2). De estos 152 destacan un total de 18, los cuales tratan temas diversos pero son totalmente complementarios al proyecto en ejecución. Las palabras clave utilizadas para la búsqueda de los artículos científicos se podrían categorizar en dos grupos, (¿Qué se aplicó?): “Business Intelligence”, “analytics”, y (¿En dónde se aplicó?): “academy”, “university”. Las palabras de (¿Qué se aplicó?) fueron utilizadas debido a que hacen referencia a la analítica visual y no a otros tipos de análisis, filtrando así de una manera óptima los resultados. Respecto a las palabras de (¿En dónde se aplicó?), hacen referencia específicamente a las instituciones de educación superior, el cual es el objetivo puntual del proyecto.

## A. Estado del arte

Las temáticas encontradas más destacadas en el ámbito universitario y académico aplicando analítica visual se centran la comparación de herramientas para determinar cuál es la más óptima para aplicar en un ámbito universitario [1], también en el análisis de datos universitarios utilizando librerías como D3 en entornos Big Data [2], análisis de datos universitarios aplicado a (trabajos de grado, temas de investigación, materias aprobadas por alumno, materias restantes por estudiante, nivel académico por alumno, calificación media del alumno por semestre, promedio de los estudiantes por materia, nota media por profesor, rendimiento del estudiante por materia, índice académico de avance académico vs su grupo, y estudiante promedio acumulado frente a su grupo) realizada por la Corporación Unificada Nacional de Educación Superior [3], análisis de datos generados entre la interacción de los estudiantes con herramientas virtuales MOOC (Massive Open Online Courses) [4], aplicación de Learning Analytics para mejorar los procesos de aprendizaje de los estudiantes de ingeniería de software [5], aplicación de Learning Analytics para el mejoramiento de los procesos de enseñanza a los profesores y los procesos de aprendizaje de los estudiantes [6], análisis de aspectos sociales de los estudiantes de ingeniería con el fin de caracterizarlos con un framework propio [7], análisis del consumo de la energía eléctrica de un campus universitario para optimizar la administración de esta [8], análisis de diferentes sectores y áreas para estructurar la base del estudio ingenieril [9], sistematizaciones de experiencias de estudiantes frente a procesos de aprendizaje y enseñanza [10], análisis comparativo entre las universidades respecto a la industria en la aplicación de inteligencia de negocios [11], análisis comparativo entre la educación presencial vs la educación mediante el uso de las TIC [12], análisis de datos Big Data en internet de las cosas para mejorar a las personas con problemas de aprendizaje como la dislexia [13], análisis del impacto de la social media en las universidades [14], análisis predictivos para determinar la nota final de grado de un estudiante (terminará o no la carrera) [15], análisis de sentimientos en comentarios respecto al campo educativo [16], análisis de redes wifi en campus universitarios para determinar cómo responde a fallos [17] y análisis de datos MOOC para entender el comportamiento de los estudiantes en este [18].

```
((((("Document Title":Businnes Intelligence) OR "Document Title":analytics) OR  
"Abstract":Businnes Intelligence) OR "Abstract":analytics) AND (((p_Title:university) OR  
"Document Title":academy) OR "Abstract":university) OR "Abstract":academy))))
```

**Figura 1 Cadena de búsqueda ejecutada.**



**Figura 2. Puntos con mayor investigación en el área**

### III. METODOLOGÍA

Para la implementación de cualquier proyecto de Inteligencia de negocios con Analítica visual se cuentan con una gran variedad de herramientas y enfoques de implementación, como herramientas gratuitas, pagas o API de desarrollo para realizar implementaciones a la medida que se ajusten a las necesidades de análisis de datos de una determinada entidad. A continuación se describirán las características de las herramientas más utilizadas respecto a analítica visual se refiere, iniciando con herramientas de terceros y luego se evaluarán las bibliotecas disponibles para desarrollos a la medida:

#### A. Herramientas de terceros

**Qlik Sense:** Se enfoca en la implementación ágil y dinámica de generación de Dashboards, enlazando todos los datos que se encuentren en el entorno de trabajo, con indicadores claves de rendimiento y todos los elementos que involucran un proceso de implementación de un proyecto de inteligencia de negocios, pero el producto generado no es de acceso libre por un tercero y no permite crear repositorios centrados de resultados en una página web externa, al menos no en su versión gratuita. Sin importar esto, posee una capacidad incremental y un motor propio de datos que permite procesos de descubrimiento de conocimiento, incluyendo un lenguaje de comandos y una API para la creación y definición a la medida de nuevos elementos de visualización con múltiples conectores de datos, por lo que deja de ser solo una herramienta de visualización, sino una plataforma completa de inteligencia de negocios.

Esta herramienta en su línea Sense ofrece dos versiones: Una de escritorio la cual es gratuita y permite acceder a múltiples fuentes de datos, realizar reportes y aplicar analítica visual pero no posee capacidad de análisis estadístico y predicción avanzada. En su versión en la nube que permite hacer los mismos reportes, solo que con un almacenamiento limitado. Esta última aunque es web, solo permite compartir el enlace del reporte generado con una cantidad limitada de usuarios, pero no permite embeber el código de este para ser mostrada en un repositorio externo.

Qlik Sense permite cargar fácilmente archivos XLS y CSV, pero también permite conexiones a bases de datos remotas y locales, a partir de ODBC y con soporte para Big Data con Apache Hadoop. Una vez los datos son cargados se comprimen y se pasan a memoria para un mejor rendimiento. Cuando los datos se encuentran listos la herramienta

permite incluir gráficas, tablas, imágenes y aplicar filtros, integrando una o muchas visualizaciones [19].

**Qlik Sense Empresarial:** Qlik Sense en su versión paga posee una funcionalidad y mecánica similar, solo que con elementos extendidos. Este es ejecutado directamente en un servidor, y se accede a él a través de un navegador web, ofreciendo escalabilidad, una arquitectura distribuida, gobernanza de datos, seguridad y soporte para dispositivos móviles. Su arquitectura en la versión empresarial es sofisticada, contando con alrededor de una docena de componentes modulares, que se pueden establecer en diferentes servidores para ser totalmente escalables.

**Qlik Sense para desarrollo y administración:** Esta herramienta ofrece facilidades para el desarrollo personalizado de reportes, brindando un entorno de trabajo similar a un “Drag and Drop” o “arrastrar y soltar”. Además incorpora diferentes API que facilitan aún más el proceso, volviéndolo extensible

**Tableau Public y Desktop:** Es un software especializado en facilidad de implementación orientada al consumo en la web es decir, en construir *Dashboards* y compartirlos de manera inmediata para que pueda ser accedido por terceros para su consulta. Esto aunque beneficioso, ha conllevado a que la construcción de este software se quede corto en procesos de escalabilidad y robustez cuando ya no se desee limitar al uso de analítica visual, como en la aplicación de procesos de descubrimiento predictivos.

Este posee un bajo soporte para desarrolladores que deseen realizar sus propios reportes personalizados y los puedan integrar al ecosistema de Tableau, pero funciona como un gran complemento para una plataforma de BI. Esto quiere decir que esta plataforma se queda corta en muchas características de descubrimiento de conocimiento fuera de analítica visual, por lo que no es considerado como una plataforma BI, pero en el segmento que se enfoca de analítica visual realiza un buen trabajo.

Los aspectos que más destacan en esta plataforma, es el soporte de visualización de datos en entornos móviles y su integración con el lenguaje de análisis R. Este paradójicamente contradice el entorno amigable y de fácil uso de Tableau, ya que este lenguaje es complejo pero ofrece todas las características de análisis que esta plataforma no ofrece de forma nativa. Se encuentra enfocado a usuarios avanzados y programadores, brindando análisis estadístico y predictivo. Adicionalmente brinda soporte para datos geográficos y cubos OLAP (de una manera limitada).

Tableau posee diferentes versiones: Una versión gratuita de escritorio llamada Tableau Desktop, versión web llamada Tableau Public y una versión paga llamada Tableau Server. La versión Desktop ofrece un entorno autónomo, permitiendo realizar análisis visuales personalizados y graficas tabulares, sea con repositorios de datos en el ecosistema, o fuentes externas de datos como archivos XLS, CSV, bases de datos locales y remotos. Adicionalmente permite a los usuarios manipular los metadatos y publicarlos para otros trabajos, cargar datos en memoria y comprimirlos para un trabajo eficiente. En su versión



de pago, el ecosistema actúa como una instalación central para el acceso a datos, entregando visualizaciones, reforzando la seguridad y gestión de acceso a los usuarios.

**Power BI:** Es una herramienta de Microsoft que soporta la creación y manipulación de tablas dinámicas, visualización, centralización y preparación de datos (Reducción de dimensionalidades, conversión de tipos de datos, agrupaciones y eliminación de registros anómalos). Actualmente cuenta con una versión de escritorio llamada Power BI Desktop y una versión en la nube, con el fin de compartir en la web, paneles y gráficos, como también la posibilidad de embeberlos en páginas web de terceros. A través de sus diferentes versiones, esta herramienta ha logrado integrarse para extraer datos de diferentes fuentes, como R, Datazen, Hadoop, Azure Machine Learning, soporte móvil entre muchas otras.

Su interfaz es muy intuitiva, debido a que se enfoca en un arrastrar y soltar elementos visuales como informes gráficos (Arboles, Graficas de tortas y donas, Mapas) y tablas, enlazados a un conjunto de datos previamente cargados. Adicionalmente dichas graficas se encuentran disponibles en el repositorio abierto D3, con lo cual se pueden generar gráficas y reportes a la medida. Los Dashboards generados pueden alimentarse con datos en tiempo real, siendo alimentados por una *API REST* o un servicio *Azure*.

**SpagoBI:** Hace parte de una iniciativa Open Source y consta de 15 componentes diferentes, permitiendo generación de informes, OLAP, minería de datos, generación de gráficos, cuadros de mando y ETL [20]. El servidor SpagoBI incluye un panel interactivo y la integración con otras herramientas de informes. Este software permite obtener datos desde Oracle, MySQL y PostgreSQL y permite desarrollo para Java Enterprise y SOA. Además posee una fuerte comunidad de desarrollo y soporte móvil [1].

**Spotfire:** Este ecosistema es una competencia directa a Tableau, ya que posee características similares a este, pero brinda como valor agregado funcionalidades de descubrimiento de conocimiento como minería de texto. Ofrece capacidades de visualización y exploración similares. Este software no cuenta con una versión gratuita permanente como si lo hace Tableau (Solo tiene una versión gratuita de 30 días) [21].

**Yellowfin:** Es una plataforma que aborda múltiples necesidades de una organización, pero no brinda de una manera intuitiva la generación de cuadros de mando en comparación a Power BI, Tableau y Qlik Sense. Sin embargo, para una persona que posea los suficientes conocimientos técnicos encontrara todos los elementos que cualquier herramienta BI brinda como gráficos, cuadros de mando, soporte móvil, mapas, gran soporte de datos, análisis avanzados, base de datos en memoria para alto rendimiento, metadatos, y una interfaz web para que el acceso está disponible en todos los dispositivos con un navegador, seguridad, escalabilidad y gobernabilidad.

**Sisense:** Es una plataforma de datos flexible, fácil de usar y con altos niveles de rendimiento. Un elemento diferenciador de este ecosistema es que mientras la mayoría de las plataformas de BI contemporáneas utilizan una base de datos de procesamiento en memoria, este utiliza una tecnología denominada In-Chip, la cual hace uso de manera

nativa de la cache disponible en los procesadores de última generación, la cual típicamente es decenas de veces más rápida que una memoria RAM. Pero no solo este cambio de enfoque mejora el rendimiento, también supera la limitación de que al cargar los datos en memoria RAM, los datos no pueden superar el tamaño de esta, permitiendo solo procesar cierta cantidad de datos simultáneamente. Este software no cuenta con una versión gratuita permanente como si lo hacen algunas otras plataformas.

**Pentaho Community:** Cuenta con múltiples funcionalidades como generación de informes, integración de datos, análisis ROLAP y minería de datos [22].

**Zoho Reports:** Es un servicio de informes y de inteligencia de negocios en línea que permite analizar fácilmente datos empresariales, crear informes y cuadros de mando para tomar decisiones. El servicio gratuito se limita a 2 usuarios y reducen drásticamente su funcionalidad.

**ClicData:** Es una plataforma de escritorio basada en la nube que ofrece una interfaz gráfica fácil de usar tipo “drag and drop”, incorporando una amplia gama de gráficos, salidas, tablas dinámicas, cálculos y muchas otras características. Es esencialmente compatible con datos CSV y Excel.

**Dashzen:** Permite la creación de Dashboards públicos y privados en la nube (Los paneles de control privados se pueden compartir con personas específicas). Permite integrar diferentes gadgets, algunos de los cuales están conectados con varias fuentes de datos (salesforce, Twitter, StackExchange), pero tiene una funcionalidad limitada.

**Slemma:** Es un software cloud computing que permite a desarrolladores y no desarrolladores hacer análisis visuales. Permite conectarse directamente a las bases de datos más populares, almacenamiento en la nube y servicios en la nube evitando así un Data Warehouse local. Posee una interfaz intuitiva de arrastrar y soltar, como también la ejecución de sentencias SQL para obtener los datos necesarios para el análisis. Los paneles creados pueden ser accedidos a través de la web, permitiendo acceder a los datos en tiempo real por dimensiones y medidas. La plataforma también admite características de transformación de datos como agregación, cálculos, filtrado, clasificación y pedidos de datos. Su versión gratuita solo soporta archivos de Excel, CSV e integración con Google Drive, Dropbox y OneDrive.

**IBM Watson analytics:** Es un servicio web que permite de una manera muy sencilla aplicar procesos de descubrimiento de conocimiento, permitiendo cargarlos, explorarlos y automatizar análisis predictivo y descriptivo sobre estos. Se puede contar con una versión de prueba de 30 días con funcionalidades limitadas, pero en su versión completa permite acceso a mayores recursos computacionales, carga de datos mayores y análisis sobre redes sociales. Actualmente la institución tiene un convenio con IBM, por lo que podría gestionarse la adquisición de este software para análisis de datos [23, 24]. Esta herramienta



no permite generar Desbordadas en la nube para embeberlos en repositorios centralizados, pero es una poderosa herramienta para descubrimiento de conocimiento.

## **B. API y tecnologías open source**

**D3:** Es una librería JavaScript para la manipulación de archivos con datos estructurados, ayudando a la visualización de estos usando HTML, SVG y CSS. El enfoque de D3 es basado en los estándares web para ser soportado en todos los navegadores modernos sin utilizar ningún framework propietario o tecnología adicional, combinando poderosos componentes de visualización para la gestión de datos [25-27]. La aplicación de esta librería se enfoca en la visualización de datos para la gestión del conocimiento y la toma de decisiones a partir de esta.

**C3:** Librería JavaScript que genera gráficos basados en D3 de una manera muy sencilla, encapsulando el código necesario para construir el grafico completo, evitando así escribir código innecesario. Permite generar clases y estilos personalizados, extendiendo de esta manera la estructura por defecto que ofrece C3 [28].

**Google Charts:** Proporciona una forma fácil de visualizar datos en sitios web, desde gráficos de líneas sencillas hasta mapas de árbol jerárquicos complejos. Este es utilizado normalmente con JavaScript cargado en la página web, procesando los datos e incrustándolos en un segmento <DIV> [29, 30].

**ReportServer:** Proporciona un entorno de creación de informes de escritorio de una manera muy flexible. Es compatible con los motores de informes de Eclipse Birt, Jasper Reports Y SAP Crystal Reports, además de sus propias herramientas de generación de informes orientadas a trabajo colaborativo. Su interfaz gráfica está orientada a la web y soporta herramientas de administración, funciones como selección de columnas, filtrado, clasificación, agrupación, subtotales, cálculos entre otros. El resultado final puede ser desde un informe hasta una aplicación interactiva en HTML5.

**BIRT:** Es un proyecto de software de código abierto que proporciona la plataforma tecnológica BIRT para crear visualizaciones de datos e informes que se puedan integrar en aplicaciones de escritorio y web, sobre tecnología JAVA EE. Este es un proyecto de alto nivel de la fundación eclipse, patrocinado por Actuate, IBM e Innovent Solutions.

**DataVision:** Es una herramienta de generación de informes Open Source similar a Crystal Reports. Estos son diseñados en una interfaz gráfica de arrastrar y soltar, y pueden ser ejecutados y consultados desde la misma aplicación o como HTML, XML, PDF, Excel, LaTeX2e, DocBook o archivos de texto delimitados. Se encuentra escrito en JAVA, permitiendo conectarse a cualquier base de datos a través de JDBC [31].

**dmyReports:** Es un generador de informes dinámicos para MySQL, permitiendo generarlos en forma de anuncios en tiempo real utilizando interfaces basadas en la web, seleccionando tablas, campos y condiciones establecidas para generar los informes.

**FyiReporting:** Es un sistema de gráficos basado en RDL (Report Definition Language). Las salidas de los reportes generados pueden ser mostradas como HTML, PDF, Excel, RTF, XML y .NET Control. Adicionalmente existe un asistente WYSIWYG que permite crear los informes sin conocimientos en RDL.

**IcCube:** Es una solución gratuita de BI que incluye informes y análisis. El informe está diseñado para usuarios empresariales y facilita la exploración de datos OLAP a través de gráficos y cuadros de mando. Los análisis incluyen una capa ETL y soporte para análisis multidimensional y MDX +. También permite realizar conexiones a otras herramientas de informes como Excel, R y Java.

**Jaspersoft Community:** Permite la creación de informes y gráficos, imágenes, tablas de referencias cruzadas y sub informes para diseños de informes sofisticados. La visualización interactiva es lograda en navegador con clasificación, filtrado y formateo de vistas instantáneas de informes. Un repositorio centralizado proporciona infraestructura para la generación de informes y almacena perfiles de usuario, informes, cuadros de mando y vistas analíticas [32].

**JMagallanes:** Es una aplicación de código abierto para manejo OLAP y Dynamic Reports escrita en Java / J2EE. Combina informes estáticos (basados en JasperReports), tablas dinámicas Swing para análisis OLAP y gráficos (basado en JFreeChart). Puede alimentarse de variadas fuentes de datos como SQL, Excel, XML y otros. Además produce muchas salidas como PDF, XML y otros archivos específicos.

**ReportMax:** Es una herramienta de informes gratuita para el framework .NET, permitiendo exportar informes a PDF, Excel o CSV. Permite construir informes simples o complejos con campos de datos, imágenes, informes secundarios, gráficos, campos calculados y más. Está integrado para ser manejado con WinForms y WebForms. Fácilmente personalizable y configurable para conectarse a cualquier fuente de datos.

**Seal Report:** Es un proyecto Open Source para Microsoft .NET totalmente escrito en C #. Soporta fuentes SQL dinámicas, tablas pivote nativas, gráficos HTML 5 y gráficos de Microsoft. Los reportes dan como resultado un HTML usando el motor Razor.

**Highcharts:** Es una librería escrita en JavaScript puro, que ofrece una forma fácil de añadir gráficos interactivos en un sitio o aplicación web. Este posee una baja curva de aprendizaje para los nuevos desarrolladores. Actualmente es open pero no comercial [33].

## IV. RESULTADOS

### A. Herramientas de terceros

Finalmente se compararon las diferentes herramientas previamente expuestas, buscando filtrar las que de una u otra manera, no cumplen las condiciones mínimas según el contexto de la EAM y el enfoque del proyecto (ver tabla 1). Los requerimientos mínimos que debe cumplir una herramienta para poder ser utilizada satisfactoriamente en la ejecución del proyecto son:

- La herramienta debe ser gratuita, debido a que la universidad no cuenta con un monto inicial de inversión para una herramienta de analítica.
- La herramienta debe permitir realizar publicaciones de los Dashboards en la Web
- Se desea que permita embeber código web, ya que se busca centralizar los reportes en un solo portal web creado a la medida.
- La herramienta debe ser lo más transparente posible, funcionando bajo HTML5 y JavaScript, con el fin de que pueda ser integrado con otras tecnologías.
- Debe permitir realizar conexión a múltiples bases de datos
- Debe ser totalmente transparente en su instalación es decir, que no requiera ningún componente adicional para poder funcionar. Esto se requiere debido a que no hay ningún tipo de garantía en el servidor donde vaya a desplegarse.

Tabla 1 Comparación entre herramientas de analítica visual

Herramienta	Fortaleza para el proyecto	Deficiencia frente al proyecto	Estado
Qlik Sense	El mejor ecosistema para analítica visual local de una manera rápida e intuitiva, además permite conectarse a bases de datos y es gratuito	No permite embeber los reportes generados en un portal web	No cumple los requerimientos mínimos
Qlik Sense Empresarial	El mejor ecosistema para analítica visual local de una manera rápida e intuitiva, además permite conectarse a bases de datos	No permite embeber los reportes generados en un portal web y no es gratuito	No cumple los requerimientos mínimos

<b>Qlik Sense</b> para desarrollo y administración	Excelente ecosistema para desarrollos a la medida, y permite conectarse a bases de datos. Además permite realizar exportaciones web	Solo permite trabajar con ciertas tecnologías web y no es gratuito	No cumple los requerimientos mínimos
<b>Tableau Public y Desktop</b>	Aunque no se considera un ecosistema BI, es excelente en el campo de analítica visual web y permite conexión a bases de datos	No ofrece KDP, pero no es relevante para el proyecto en curso.	Aceptable para el proyecto
<b>Tableau Server</b>	Aunque no se considera un ecosistema BI, es excelente en el campo de analítica visual web y permite conexión a bases de datos y Data Warehouse	No ofrece KDP, pero no es relevante para el proyecto en curso. Además no es gratuito	No cumple los requerimientos mínimos
<b>Power BI</b>	Excelente ecosistema para analítica visual, permite exportarlo en la web y embeber código. Adicionalmente es gratuito	En su versión gratuita no ofrece procesos de KDP, aunque no es relevante para el proyecto	Aceptable para el proyecto
<b>SpagoBI</b>	Permite conexión a bases de datos y aplicación de procesos KDP	Su uso no es tan intuitivo como Tableau o Power BI, y necesita de software pre – instalado para funcionar como lo es open office.	No cumple los requerimientos mínimos
<b>Spotfire</b>	Ofrece servicios muy similares a Tableau. Adicionalmente minería de texto	Es de pago	No cumple los requerimientos mínimos
<b>Yellowfin</b>	Buena herramienta de analítica visual	La curva de aprendizaje es muy grande, debido a la complejidad que	No cumple los requerimientos mínimos

		tiene para la generación de cualquier reporte	
<b>Sisense</b>	Excelente rendimiento al procesar los datos en memoria	Es de pago	No cumple los requerimientos mínimos
<b>Pentaho Community:</b>	Permite aplicar procesos KDP y analítica visual	Solo permite compartir elementos en la misma plataforma, no de manera web.	No cumple los requerimientos mínimos
<b>Zoho Reports</b>	Permite aplicar analítica visual en la web	Su versión gratuita es muy limitada.	No cumple los requerimientos mínimos
<b>ClicData</b>	Permite realizar procesos de analítica visual de una manera muy sencilla e intuitiva	No tiene soporte web y no tiene conexión a base de datos	No cumple los requerimientos mínimos
<b>Dashzen</b>	Permite realizar reportes básicos web	Tiene una funcionalidad muy limitada	No cumple los requerimientos mínimos
<b>Slemma</b>	Permite aplicar procesos de analítica visual orientados a la web	Su versión gratuita no soporta conexión a base de datos, lo cual es necesario a mediano plazo	No cumple los requerimientos mínimos
<b>Watson Analytics</b>	Permite aplicar procesos de descubrimiento de conocimiento de una manera intuitiva	Solo dispone de una versión gratuita de 30 días. Además no permite publicar los resultados embebidos en una página web	No cumple los requerimientos mínimos

A partir de esta comparación, las únicas herramientas que cumplen con los requerimientos establecidos son: Tableau Public y Power BI. Aunque Qlik View Sense no cumple un requerimiento fundamental el cual es embeber código para un repositorio centralizado web, se realizara una comparación también con este, debido a que como plataforma local es la herramienta que más sobresaliente frente a las demás, y puede fortalecer algunos procesos de análisis (ver tabla 2).

Tabla 2 Comparación Tableau Public, Power BI y Qlik View Sense

	<b>Qlik View Sense</b>	<b>Power BI</b>	<b>Tableau</b>
<b>Pre-procesamiento de datos</b>	Si (Bueno)	Si (Bueno)	Si (Pobre)
<b>Medidas</b>	Sí	Sí	Sí
<b>Apertura publica de datos en la web</b>	No	Sí	Sí
<b>Almacenamiento gratuito en la nube</b>	Sí	Sí	Sí
<b>Documentación</b>	Bueno	Bueno	Pobre
<b>Carga incremental de datos</b>	Sí	Sí	Próximamente
<b>Tiempo de implementación</b>	Rápido	Normal	Normal
<b>Soporte para español</b>	Sí	Sí	No
<b>Visualización</b>	Sí (Buena)	Sí (Buena)	Sí (Buena)
<b>Filtros dinámicos</b>	Sí (Automáticos)	Sí (Semi automático)	Sí (Manual)
<b>Conexión de datos</b>	Sí (Bueno)	Sí (Bueno)	Sí (Bueno)

## B. API de desarrollo y soluciones open source

Finalmente se compararon las diferentes *API* expuestas anteriormente, buscando filtrar las que de una u otra manera, no cumplen las condiciones mínimas según el contexto de la EAM y el enfoque del proyecto (ver tabla 3). Los requerimientos mínimos que debe cumplir una herramienta para ser poder ser utilizada satisfactoriamente en la ejecución del proyecto son:

- La *API* debe ser gratuita, debido a que la universidad no cuenta con un monto inicial de inversión para una herramienta de analítica.
- La *API* debe ser lo más transparente posible, funcionando bajo HTML5 y JavaScript, con el fin de que pueda ser integrado con otras tecnologías dado el caso que se requiera.
- Debe ser totalmente transparente en su uso, es decir, que no requiera ningún componente o framework adicional para poder funcionar.
- Debe ser soportado por cualquier tecnología, no debe estar ligada a una o varias.
- Debe tener una curva de aprendizaje muy baja, con el fin de aplicar proyectos a una alta velocidad.
- Debe soportar múltiples conexiones a bases de datos, dato el caso que se encargue de esto.



**Tabla 3 Comparación entre API y Open Source**

Herramienta	Fortaleza para el proyecto	Deficiencia frente al proyecto	Estado
D3	Soporte web transparente para cualquier tecnología.	No está enfocado totalmente a BI, si no a gestión del conocimiento, aunque ofrece muchas visualizaciones BI	Aceptable para el proyecto
C3	Soporte web transparente para cualquier tecnología.	Es un sub framework de D3, por lo que tocar su base para modificaciones específicas puede ser más complejo.	Aceptable para el proyecto
<b>Google Charts</b>	Soporte web transparente para cualquier tecnología	Herramienta menos intuitiva y con menos soporte que D3 o C3	Aceptable para el proyecto
<b>ReportServer</b>	Es orientada a la web generando los elementos en HTML5	Solo soporta JAVA y .NET	No cumple los requerimientos mínimos
<b>BIRT</b>	Intuitivo y fácil de usar	Solo tiene soporte para JAVA de escritorio	No cumple los requerimientos mínimos
<b>DataVision</b>	Orientado a la web	Curva de aprendizaje muy alta	No cumple los requerimientos mínimos
<b>dmyReports</b>	Permita visualizar datos en tiempo real	Solo soporta base de datos MySQL	No cumple los requerimientos mínimos
<b>FyiReporting</b>	Orientado a la web	Curva de aprendizaje muy alta	No cumple los requerimientos mínimos
<b>icCube</b>	Permite la aplicación de procesos de descubrimiento de conocimiento	Curva de aprendizaje muy alta y bajo soporte web	No cumple los requerimientos mínimos
<b>Jaspersoft Community</b>	Gran variedad y posibilidades de informes	Curva de aprendizaje y no es totalmente transparente para cualquier tecnología.	No cumple los requerimientos mínimos

<b>JMagallanes</b>	Gran variedad y de posibilidades informes	Solo funciona bajo la tecnología .NET	No cumple los requerimientos mínimos
<b>ReportMax</b>	Gran variedad y de posibilidades informes	Solo funciona bajo la tecnología .NET	No cumple los requerimientos mínimos
<b>Seal Report</b>	Gran variedad y de posibilidades informes	Solo funciona bajo la tecnología .NET	No cumple los requerimientos mínimos
<b>Highcharts</b>	Soporte web transparente para cualquier tecnología.	Es open pero no comercial	Aceptable para el proyecto

A partir de esta comparación, las únicas herramientas open source o *API* que cumplen con los requerimientos establecidos son: C3, D3 y Google Charts. Se realizara una comparación entre estas tres, analizando múltiples facetas para determinar cuál es la más adecuada para una posible ejecución del proyecto (ver tabla 4).

**Tabla 4. Comparación D3, C3, Google Charts, HighCharts**

	<b>D3</b>	<b>C3</b>	<b>Google Charts</b>	<b>Highcharts</b>
<b>Ejemplos y documentación en línea</b>	Sí (Excelente)	Sí (Bueno)	Sí (Sobresaliente)	Sí (Bueno)
<b>Requiere de componentes adicionales</b>	No	Si (Este es un sub framework de D3, pero es soportado por todas las tecnologías)	No	No
<b>Comunidad activa para consultas</b>	Sí (Excelente)	Si (Sobresaliente)	Sí (Bueno)	Sí (Bueno)
<b>Curva de aprendizaje</b>	Aunque es fácil de usar, es más complejo que C3 y Google Charts	Sí (Excelente)	Sí (Sobresaliente)	Aunque es fácil de usar, es más complejo que C3 y Google Charts

<b>Flexibilidad en la manipulación de datos</b>	Sí (Excelente)	Sí (Excelente)	Sí (Bueno)	Sí (Bueno)
<b>Ofrece interactividad</b>	Sí (Bueno)	Sí (Bueno)	Sí (Limitado)	Sí (Bueno)
<b>Compatible con navegadores no modernos</b>	Sí (Limitado)	No	No	Sí (Limitado)
<b>Ligero</b>	Sí (Bueno)	No	Sí (Excelente)	Sí (Bueno)

## V. CONCLUSIONES

Se han identificado 2 posibles enfoques de implementación del proyecto para una institución de educación superior: Utilizando herramientas de terceros donde se pueden obtener resultados en un corto tiempo, con interactividad y cruces de datos de manera abierta a toda la comunidad universitaria, pero se está sujeto al alcance y limitaciones de la herramienta que se esté utilizando. El otro posible enfoque es desarrollos a la medida a partir de API existentes, en el cual los resultados se pueden ver a mediano o largo plazo, pero ofrece mayores posibilidades respecto a necesidades específicas. En este momento, las herramientas de terceros analizadas satisfacen las necesidades de análisis de datos que tiene la Institución Universitaria EAM, siendo estas en conjunto una herramienta poderosa para la implementación del proyecto. Por otra parte las API de desarrollo deben ser implementadas en la institución pero, aprovechándolas para la elaboración de trabajos de grado o proyectos de investigación en semilleros del programa de ingeniería de software, fortaleciendo los conocimientos de estas librerías en los estudiantes y aprovechando esta mano de obra para construcción de Dashboards a la medida, creando un marco de trabajo y estandarizando el proceso, creando equipos de desarrollo interno conformado por los estudiantes fortaleciendo así la sublínea existente de Ingeniería de Software (ver figura 3).

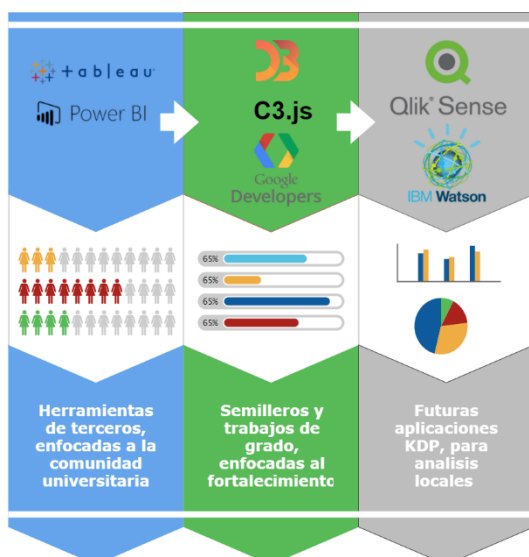


Tableau aunque no posee características de escalabilidad robustas como sus competidores, pero ofrece una gran facilidad en la publicación de dashboards de una manera casi inmediata, lo cual satisface las necesidades de lo que está buscando la institución. Es importante resaltar que en términos generales Qlik Sense es mejor que Tableau, pero para el objetivo del proyecto de investigación y las necesidades de la EAM, el cual es publicar y liberar dichos reportes a la comunidad universitaria, Qlik Sense no brinda esta posibilidad, mientras que Tableau sí. Otra opción es formar un equipo de desarrollo en

Figura 3 Trazabilidad establecida de herramientas.

API especializadas de visualización como D3, C3 o Google Charts, pero este tipo de enfoques económicamente hablando es costoso, por lo que se propone es destinar semilleros de investigación que utilicen dichos repositorios para validar prácticas como auxiliares de investigación y elaboración de trabajos de grado en el nivel profesional.

Basándose en que Tableau se queda corto para ser considerada una herramienta de inteligencia de negocios, pero si es una herramienta sumamente poderosa para visualización de datos, esta debe ser tomada como herramienta de apoyo con otro tipo de herramientas, como lo es Power BI. Se debe tener presente que, a mediano plazo, si la institución desea aplicar procesos de descubrimiento de conocimiento, puede utilizar Watson Analytics y Qlik Sense ya que, aunque no cumplieron los requerimientos mínimos en la ejecución del proyecto, ofrecen resultados a corto plazo respecto a análisis de datos local y además, puede ser adquirida la licencia de Watson con relativa facilidad, debido a los convenios existentes. Los resultados que se obtuvieron al analizar las herramientas analizadas fueron similares a investigaciones existentes como la de S Gounder, et al. [1], siendo las herramientas Tableau y Power BI las herramientas seleccionadas para la publicación de reportes en el portal para resultados inmediatos y las API de desarrollo D3, C3 y Google Charts para la focalización de proyectos de investigación en semilleros y trabajos de grado para formar los cimientos de futuros desarrollos a la medida.

### Referencias bibliográficas

- [1] M. S Gounder, V. Vasudevan Iyer, and A. Al Mazyad, "A Survey on Business Intelligence tools for University Dashboard development," *IEEE*, 2016.
- [2] T. Soklakova, A. Ziarmand, and S. Osadchyieva, "Big Data Visualization in Smart Cyber University " *IEEE*, 2016.
- [3] Y. V. Nieto and V. García, "Academic Decision Making Model for Higher Education Institutions using Learning Analytics," *IEEE*, pp. 27 - 32, 2016.
- [4] Á. Del Blanco, Á. Serrano, M. Freire, I. Martínez, and B. Fernández, "E-Learning Standards and Learning Analytics " *IEEE*, pp. 1255 - 1261, 2013.
- [5] N. Pratheesh and T. Devi, "Influence of Learning Analytics in Software Engineering Education," *IEEE*, pp. 712 - 716, 2013.
- [6] Q. Zhou, X. Han, J. Yang, and J. Cheng, "Design and Implementation of Learning Analytics System for Teachers and Learners Based on the Specified LMS," *IEEE*, pp. 79 - 82, 2014.
- [7] J.-L. Mondisa, J. Millunchick, C. Davis, and D. Koch, "The University of Michigan's M-STEM Academies Program: Examining the Social Community of Future Engineers " *IEEE*, 2016.
- [8] B. Diong, G. Zheng, and M. Ginn, "Establishing the Foundation for Energy Management on University Campuses via Data Analytics " *IEEE*, 2015.
- [9] A. Friesel and D. Cojocar, "Identifying how PELARS-project can support the development of new curriculum structures in engineering education," *IEEE*, pp. 219 - 223, 2015.
- [10] A. Gewerc , F. Fraga, A. Rodríguez, A. Ferreira, and V. Rodés, "E-portfolios and social networks in university education " *IEEE*, 2016.

- [11] I. Guitart, J. Conesa, and J. Casas, "A Preliminary Study about the Analytic Maturity of Educational Organizations " *IEEE*, pp. 345 - 350, 2016.
- [12] E. Khoo, J. Scott, M. Peter, and H. Round, "Evaluating Flipped Classrooms with respect to Threshold Concepts Learning in Undergraduate Engineering," *IEEE*, 2015.
- [13] L. Lenz, T. Meisen, A. Pomp, and S. Jeschke, "How will The Internet of Things and Big Data Analytics impact the Education of Learning-Disabled Students?," *IEEE*, 2016.
- [14] B. M. Waseemah and A. S. J. Harouni, "Evaluating the Strategic Role of Social Media Analytics to Gain Business Intelligence in Higher Education Institutions " *IEEE*, 2016.
- [15] K. Mouri, A. Shimada, F. Okubo, and H. Ogata, "Bayesian Network for predicting students' final grade using e-book Logs in University Education," *IEEE*, pp. 85 - 89, 2016.
- [16] G. I. Nitin, G. Swapna, and V. Shankararaman, "Analyzing Educational Comments for Topics and Sentiments: A Text Analytics Approach," *IEEE*, 2015.
- [17] H. Park, H. Gebre-Amlak, B.-Y. Choi, S. Song, and D. Wolfenbarger, "Understanding University Campus Network Reliability Characteristics using a Big Data Analytics Tool," *IEEE*, pp. 107 - 110, 2015.
- [18] H. Qu and Q. Chen, "Visual Analytics for MOOC Data," *IEEE*, pp. 69 - 75, 2015.
- [19] E. Nicklas, T. Vanja, B. Markus, A. o. Anders , and R. Per, "Navigating Information Overload Caused by Automated Testing – A Clustering Approach in Multi-Branch Development," *IEEE*, 2015.
- [20] A. Gioia, G. Cazzin, and E. Damiani, "SpagoBI: a distinctive approach in Open Source Business Intelligence," *IEEE*, pp. 592 - 595, 2008.
- [21] N. Tandon, D. Cleverdon, and B. Hinshaw, "Integrated web-based architecture for correlative engineering data analysis and decision support," *IEEE*, pp. 273 - 279, 2003.
- [22] A. Marinheiro and J. Bernardino, "Experimental Evaluation of Open Source Business Intelligence Suites using OpenBRR," *IEEE*, pp. 810 - 817, 2015.
- [23] A. A. Almutairi and S. A. El Rahman, "The impact of IBM cloud solutions on students in Saudi Arabia," *IEEE*, pp. 115 - 118, 2016.
- [24] S. S. Murtaza, P. Lak, and A. Bener, "How to Effectively Train IBM Watson: Classroom Experience," *IEEE*, pp. 1663 -1670, 2016.
- [25] M. Bostock. (2017, Mayo). *D3 Data-Driven Documents*. Available: <https://d3js.org/>
- [26] L. Chen and H. Zhou, "Research and application of dynamic and interactive data visualization based on D3," *IEEE*, pp. 150 - 155, 2016.
- [27] J. Harper and M. Agrawala, "Converting Basic D3 Charts into Reusable Style Templates," *IEEE*, pp. 1 - 11, 2017.
- [28] M. Tanaka. (2014, Mayo). *C3.js D3-based reusable chart library*. Available: <http://c3js.org/>
- [29] Google. (2017, Abril). *Using Google Charts*. Available: <https://developers.google.com/chart/interactive/docs/>
- [30] J.-G. Lee, Y.-H. Kim, and I.-K. Lim, "Implementation of u-RPMS Using Google Chart in Hybrid Application for Visualization of Patient's Biometric Information," *IEEE*, pp. 6 - 9, 2013.

- [31] S. D. Shaye, "A test data collection system for uniform data analysis," pp. 242 - 251, 2002.
- [32] A. Marinho and J. Bernardino, "Analysis of open source Business Intelligence suites," *IEEE*, pp. 19-22, 2013.
- [33] O. ElTayeb, D. John, and P. Patel, "Comparative case study between D3 & Highcharts on Lustre metadata visualization," *IEEE*, 2013.