

Universidad Católica de Santiago del Estero

Departamento Académico Rafaela



**Diseño de reportes visuales mediante la conexión a
sistemas legacy utilizando análisis de datos.**

Carrera: Ingeniería en Informática.

Tema: Anteproyecto.

Alumno: Wendy Sclerandi.

Ciudad: Rafaela.

Año: 2017.

Índice

Descripción del problema	3
Objetivos	3
Antecedentes	3
Justificación.....	3
Marco teórico.....	4
Sistema Legacy.....	4
Base de Datos	4
Inteligencia de Negocios	5
Visualización de Datos.....	5
Análisis de Datos	5
D3: Data-Driven Documents	6
Metodología para realizar el trabajo	6
Relevamiento de requerimientos	6
Análisis.....	6
Diseño.....	6
Desarrollo.....	7
Prueba	7
Alcance	7
Plan de trabajo y cronograma	7
Bibliografía	8

Descripción del problema

En la actualidad, las empresas de Rafaela necesitan contar con información en tiempo y forma para tomar decisiones que mejoren su producción y competitividad en el entorno en el cual se desenvuelven.

Para ello, la mayoría de los sistemas actuales no cuentan con funcionalidades de reportes que permitan obtener nueva información agregada sobre los datos que recopilan a diario. Esto se debe a la utilización de sistemas realizados con tecnologías obsoletas, lo que representa un problema a la hora de desarrollar este tipo de funcionalidades extras.

Mediante este proyecto, se plantea solucionar este inconveniente mediante el desarrollo de un sitio web con conexión al motor de base de datos de estos sistemas con el fin de obtener la información necesaria en tiempo y forma para que el usuario pueda realizar un análisis en formato visual y gráfico.

Objetivos

El objetivo del presente proyecto es facilitar el trabajo de los tomadores de decisiones de las empresas Rafaelinas mediante el desarrollo de un sitio web con funcionalidades de reportes gráficos.

Antecedentes

En la actualidad existen sitios web que le brindan al usuario la posibilidad de subir sus datos a la misma y elegir cómo visualizarlos entre un conjunto de gráficos.

Dichas páginas tienen como desventaja el límite en cuanto a la cantidad de datos que es posible ingresar, así como también la fuente de datos, ya que solo es posible cargar archivos de texto, no permiten conexiones a bases de datos.

Asimismo, algunas de estos sitios requieren suscripción y pago de la misma para utilizarlas.

A continuación, se listan algunos sitios web:

- <http://rawgraphs.io/>
- <https://www.tableau.com/es-es>
- <https://live.amcharts.com/>
- <https://vizzlo.com/>

Justificación

Varias empresas de Rafaela tienen antigüedad, por lo que llevan almacenado grandes cantidades de datos, los cuales no pueden aprovechar al máximo al utilizar sistemas legacy sin soporte para generación de reportes.

Este problema me llevó a pensar que las empresas están perdiendo oportunidades de mercado y crecimiento al no poder visualizar y analizar la información de forma agregada.

Asimismo, algunas empresas no tienen conciencia de la existencia de herramientas para el apoyo en la toma de decisiones.

Debido a estas razones, considero que es necesario brindar apoyo informático, ya que facilitaría el trabajo de los tomadores de decisiones y generaría beneficios en las empresas.

Marco teórico

A continuación, se exponen los principales aspectos teórico-prácticos necesarios para el desarrollo del presente proyecto.

Sistema Legacy

En el paper “Challenges in migrating legacy software systems to the cloud - an empirical study” (Mahdi Fahmideh Gholami, 2017) se define al sistema legacy o sistema heredado como un sistema informático anterior u obsoleto. En el mismo se mencionan dos estudios que definen a los sistemas heredados como “aquellos que son críticos para la misión, costosos de mantener, frágiles e inflexibles a los cambios, ejecutados en hardware obsoleto, documentación incompleta u obsoleta, y difíciles de extender e integrar con otros sistemas.”

El término sistema heredado puede hacer referencia a sistemas antiguos escritos en lenguajes de programación obsoletos, o a sistemas modernos desarrollados con últimas tecnologías, pero que actualmente no satisfacen los nuevos requisitos comerciales. (Mahdi Fahmideh Gholami, 2017)

Asimismo, en el paper “Organisational scenarios and legacy systems” (Brookea & Ramageb, 2001) se destaca la diferencia entre un sistema heredado y un software heredado. El software heredado es un software crítico para la empresa que no se puede modificar de manera eficiente sin incurrir en grandes costos (en términos de tiempo, habilidades, etc.). Suele ser software obsoleto, anticuado, difícil de mantener. En cambio, un sistema heredado es más amplio ya que abarca tanto al software como a temas de estructura organizacional, estrategia, proceso y flujo de trabajo.

Base de Datos

Citando al libro “Introducción a los sistemas de base de datos” (Date, 2001), un sistema de base de datos es un sistema computarizado para guardar registros; es decir, su finalidad general es almacenar información y permitir a los usuarios recuperar y actualizar esa información con base en peticiones.

Cabe destacar la diferencia entre dato e información. Un dato es una representación simbólica de un atributo, por lo que no tiene un valor semántico,

mientras que la información es un conjunto organizado de datos procesados. Tiene significado, relevancia y propósito. Los datos se convierten en información cuando su creador les añade significado.

Inteligencia de Negocios

La inteligencia de negocios es una herramienta bajo la cual diferentes tipos de organizaciones pueden soportar la toma de decisiones. Esto se logra mediante el uso de metodologías, aplicaciones y tecnologías que permiten reunir, depurar, transformar datos, y aplicar en ellos técnicas analíticas de extracción de conocimiento, permitiendo a las empresas establecer sus fortalezas y debilidades. (Gomez, 2010)

Visualización de Datos

La visualización de datos es un término que describe cualquier esfuerzo para ayudar a las personas a comprender el significado de los datos colocándolos en un contexto visual. Los patrones, las tendencias y las correlaciones que pueden pasar desapercibidos en los datos basados en texto pueden exponerse y reconocerse más fácilmente con un software de visualización de datos. (Rouse, Search Business Analytics)

Por lo tanto, la visualización de datos posibilita la visualización gráfica de información abstracta con dos propósitos: creación de sentido y comunicación para el usuario. (Few)

Las herramientas de visualización de datos se conectan a las bases de datos de las empresas para extraer los datos y así poder aplicar algún tipo de gráfico a los mismos. Usualmente, estas herramientas le permiten al usuario seleccionar la mejor manera de presentar los datos mediante un componente de tablero. (Few)

Análisis de Datos

El Análisis de Datos es la ciencia que examina datos en bruto con el propósito de sacar conclusiones sobre la información. El análisis de datos es usado en varias industrias para permitir que las compañías y las organizaciones tomen mejores decisiones empresariales y también es usado en las ciencias para verificar o reprobando modelos o teorías existentes. (Rouse, Search Datacenter en español)

En el mismo sitio se menciona que es necesario distinguir entre la extracción de datos y el análisis de datos. Los extractores de datos clasifican inmensos conjuntos de datos usando software sofisticado para identificar patrones no descubiertos y establecer relaciones escondidas. En cambio, el análisis de datos se centra en la inferencia, el proceso de derivar una conclusión basándose solamente en lo que conoce el investigador.

D3: Data-Driven Documents

Esta herramienta se utilizará para visualizar gráficamente los datos recolectados de la empresa.

Como se menciona en el sitio web de D3 (D3 Data-Driven Documents), D3.js es una biblioteca de JavaScript para manipular documentos basados en datos. D3 ayuda a dar vida a los datos usando HTML, SVG y CSS. El énfasis de D3 en los estándares web permite utilizar todas las capacidades de los navegadores modernos sin vincularse a un marco propietario, combinando poderosos componentes de visualización y un enfoque basado en datos para la manipulación de DOM.

D3 permite vincular datos arbitrarios con un Modelo de Objetos de Documento (DOM) y luego aplicar transformaciones basadas en datos al documento.

Además, D3 ofrece una flexibilidad extraordinaria, exponiendo todas las capacidades de los estándares web como HTML, SVG y CSS. Presenta una sobrecarga mínima, es extremadamente rápido y admite grandes conjuntos de datos y comportamientos dinámicos para la interacción y la animación. El estilo funcional de D3 permite la reutilización del código a través de una colección diversa de módulos oficiales y desarrollados por la comunidad.

Metodología para realizar el trabajo

Para llevar a cabo este proyecto se realizarán las actividades descritas posteriormente. Cabe aclarar que las etapas detalladas suelen ser iterativas ya que el proceso de creación del software es un trabajo incremental.

Relevamiento de requerimientos

En esta primera etapa se realizarán entrevistas con la empresa a fin de relevar las necesidades de la misma. Será necesario constatar los tipos de datos que maneja la empresa y cuáles son de relevancia a la hora de tomar decisiones empresariales.

Análisis

Se analizará lo relevado con el objetivo de corroborar que los requisitos cumplan con las necesidades reales de la empresa para luego definir un modelo de desarrollo a seguir para la creación del sitio web.

Diseño

A partir de los requerimientos analizados, se realizarán modelos de diseño que guíen el proceso de creación del software, detallando las tareas y funcionalidades a desarrollar.

Desarrollo

En esta etapa se programarán todas las funcionalidades modeladas en la etapa de diseño.

Prueba

Se evalúan los requisitos descritos en la etapa de análisis en contraste al software desarrollado, para validar y verificar que se haya cumplido con los mismo y que responda a las necesidades de la empresa.

Alcance

- Recolectar datos de la empresa.
- Analizar los datos recolectados.
- Planificar el desarrollo del sitio web.
- Desarrollar un sitio web para generación de reportes.
- Testear el sitio web.

Plan de trabajo y cronograma

Actividades	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5	Mes 6	Mes 7
Investigación Bibliográfica	X						
Definición del Marco Teórico	X						
Definición del Marco Metodológico	X						
Recolección y Procesamiento de los Datos		X					
Análisis de datos		X					
Planificación para el desarrollo			X				
Desarrollo del sitio web			X	X	X		
Testeo del sitio web						X	
Presentación del proyecto							X

Bibliografía

- Brookea, C., & Ramageb, M. (2001). Organisational scenarios and legacy systems. *International Journal of Information Management*.
- D3 Data-Driven Documents. (s.f.). Obtenido de <https://d3js.org/>
- Date, C. J. (2001). *Introducción a los sistemas de Base de Datos*. Pearson Educación.
- Few, S. (s.f.). *Interaccion Desing Foundation*. Obtenido de <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception>
- Friedman, V. (2008). *Smashing Magazine*. Obtenido de <https://www.smashingmagazine.com/2008/01/monday-inspiration-data-visualization-and-infographics/>
- Gomez, A. A. (2010). *Inteligencia de Negocios: Estado del Arte*. Obtenido de <http://www.redalyc.org/html/849/84917316060/>
- Mahdi Fahmideh Gholami, F. D. (2017). Challenges in migrating legacy software systems to the cloud - an empirical study. *Information Systems*.
- Rouse, M. (s.f.). *Search Business Analytics*. Obtenido de <http://searchbusinessanalytics.techtarget.com/definition/data-visualization>
- Rouse, M. (s.f.). *Search Datacenter en español*. Obtenido de <http://searchdatacenter.techtarget.com/es/definicion/Analisis-de-Datos>