

TFG - Health IT: YouTubeCrawlerTool

PEC 3: Informe de seguimiento

Año académico 2017/18-2

Javier Sánchez Mendoza

20-5-2018

TFG - Health IT: YouTubeCrawlerTool : PEC 3 – Informe de seguimiento

1. Avance del proyecto	2
1.1. Grado de cumplimiento de los objetivos	2
1.2. Justificación de cambios	3
2. Relación de actividades realizadas	4
2.1. Actividades previstas en el plan de trabajo	4
2.2. Actividades no previstas y realizadas o programadas	4
3. Relación de desviaciones en la temporización, acciones de mitigación y actualización del cronograma	5
4. Listado de resultados parciales obtenidos hasta el momento	6

1. Avance del proyecto

1.1. Grado de cumplimiento de los objetivos

A continuación se facilita la relación de objetivos definidos para el proyecto y su grado de cumplimiento al termino de la segunda fase de desarrollo:

- *Investigar que funcionalidades aportan las API públicas ofrecidas por YouTube y analizar como se pueden utilizar para la obtención de la información requerida.*
 - Se ha analizado la API de YouTube y se ha estudiado conjuntamente con el cliente que información va a ser recolectada por la aplicación. Este objetivo se realizo durante la primera fase de desarrollo.
- *Determinar como almacenar y acceder de forma eficiente a la gran cantidad de información que se obtendrá.*
 - Se ha tomado la decisión de utilizar una base de datos NoSQL enfocada a grafos (Neo4j) para poder acceder a la información de forma eficiente. Se ha diseñado un diagrama UML para representar el esquema invariante de la información y se ha adaptado al concepto de nodos y relaciones de las base de datos de grafos. Se realizo una prueba de concepto para estudiar la viabilidad de utilizar Neo4j como base de datos en el proyecto con resultado favorable. La prueba de concepto realizada fue utilizada como base para la implementación de la aplicación durante la segunda fase de desarrollo. Este objetivo se considera realizado.
- *Permitir la recolección de información según criterios de búsqueda proporcionados por el usuario final.*
 - Conjuntamente con el cliente se han determinado los criterios de búsqueda a utilizar. Se ha desarrollado una prueba de concepto en la cual a sido posible recolectar información de videos y canales utilizando la API de YouTube. La prueba de concepto realizada fue utilizada como base para la implementación de la aplicación durante la segunda fase de desarrollo. Este objetivo se considera realizado.
- *Habilitar la gestión, visualización y exportación de datos obtenidos en distintos procesos de extracción para su posterior análisis en herramientas especializadas.*
 - Se ha diseñado una interfaz de usuario para proporcionar estas funcionalidades y se ha diseñado la arquitectura de la aplicación para hacerlas viables. En la segunda fase de desarrollo se ha adoptado y implementado la arquitectura propuesta en la primera fase de desarrollo y la interfaz de usuario necesarias para habilitar estas funcionalidades al usuario final. Este objetivo se considera realizado.
- *Ofrecer herramientas de visualización para el análisis y comprensión de los datos obtenidos.*
 - Se ha decidido el uso de un grafo para la visualización de los datos obtenidos y se ha determinado cuales serán sus componentes. Se ha realizado una prueba de concepto en la cual ha sido posible visualizar los datos obtenidos a través de la API de YouTube

en un grafo. La prueba de concepto realizada fue utilizada como base para la implementación final de la aplicación y se consiguió visualizar los datos obtenidos en un grafo acorde a las especificaciones del cliente. Este objetivo se considera realizado.

- *Proporcionar una interfaz de usuario usable que permita realizar las acciones requeridas por el usuario final.*
 - Se ha realizado una propuesta de interfaz de usuario que ha sido aprobada por la cliente. En la segunda fase de desarrollo se implementó la interfaz gráfica propuesta con la aprobación final de la cliente. Este objetivo se considera realizado.

1.2. Justificación de cambios

Durante la realización de la segunda fase de desarrollo se realizaron cambios que afectaron al alcance del proyecto o su implementación, la relación de cambios es la siguiente:

- **Información estadística:** En la propuesta de interfaz de usuario presentada en la primera fase de desarrollo en algunas secciones de la aplicación se incluía información estadística sobre el estado de la aplicación y de la información recolectada (porcentaje de videos categorizados, distribución de categorías, etc.). Debido al ajustado tiempo disponible para la implementación de esta funcionalidad y a la baja prioridad dada por el cliente en comparación a otras funcionalidades, finalmente esta característica no fue incluida dentro del piloto de la aplicación.
- **Opción de favoritos:** Después de realizar la presentación del piloto de la aplicación a la cliente, debido a la gran cantidad de videos que pueden ser recolectados por la aplicación y para poder facilitar su identificación, se considero oportuno añadir la funcionalidad de favoritos para poder marcar los videos que la cliente considera oportunos como favoritos para poder acceder a ellos de forma directa. Además se diseño y añadió un nuevo filtro en la sección de “Analysis” para poder visualizar en el grafo resultante solo los videos marcados como favoritos.

2. Relación de actividades realizadas

2.1. Actividades previstas en el plan de trabajo

De las actividades previstas en el plan de trabajo para la segunda fase de desarrollo se llevaron a cabo todas las actividades planificadas. Su relación es la siguiente:

- Implementación de la aplicación
 - Modelo de datos
 - Capa de acceso a datos
 - Capa de presentación
 - Definición y ejecución de pruebas
- Actualización servidor de explotación
- Redacción de la memoria

2.2. Actividades no previstas y realizadas o programadas

Durante la realización de la segunda fase de desarrollo se realizaron cambios que afectaron al alcance de la implementación de la aplicación. Estos cambios, que se detallan en la sección 1.2 de este documento, al tratarse de nuevas funcionalidades a añadir en la aplicación pudieron ser absorbidas por las actividades previstas previamente para la implementación de la aplicación, por lo que no fue necesario realizar modificaciones de las tareas previstas inicialmente.

3. Relación de desviaciones en la temporización, acciones de mitigación y actualización del cronograma

Debido a que todas las tareas planificadas para la segunda fase de desarrollo han podido realizarse, no se ha producido ninguna desviación en la temporización del proyecto y no es necesario realizar ninguna acción de mitigación o actualizar el cronograma.

Siendo así, las tareas a realizar a continuación son las que figurando en el diagrama de Gantt proporcionado al término de la primera fase de desarrollo:

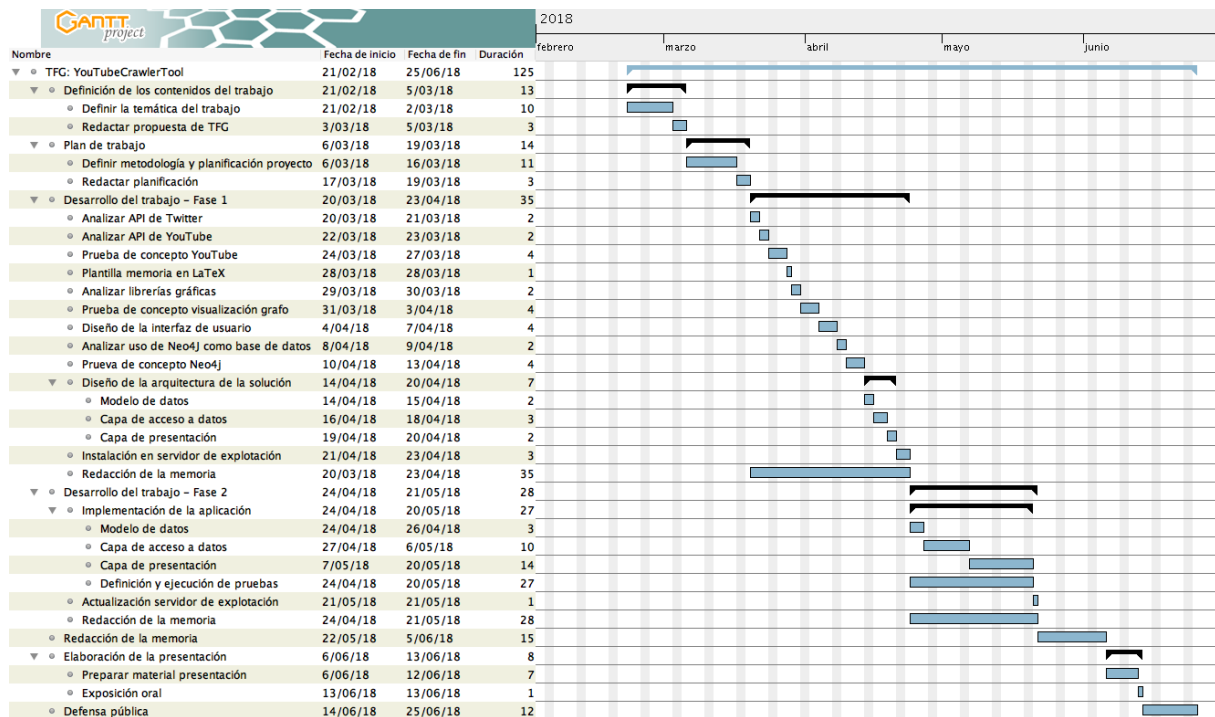


Figure 1 Listado de tareas y diagrama de Gantt

4. Listado de resultados parciales obtenidos hasta el momento

A continuación se detalla el listado de los resultados parciales obtenidos durante la segunda fase de desarrollo del proyecto:

- Implementación de la aplicación:
<https://github.com/jsanchezmend/TFGAntivacunas/tree/master/YouTubeCrawlerTool>
- Instalación en servidor de explotación:
Se ha realizado la instalación de la aplicación en un servidor el cual los stakeholders del proyecto pueden tener acceso mediante autenticación previa.
- Informe de test de integración realizados:
<https://github.com/jsanchezmend/TFGAntivacunas/tree/master/Pruebas%20de%20integraci%C3%B3n>
- Memoria del trabajo (inacabada):
https://github.com/jsanchezmend/TFGAntivacunas/blob/master/Memoria/TFG_JavierSanchezMendoza.pdf