



**UNIVERSIDAD DE CASTILLA-LA MANCHA
ESCUELA SUPERIOR DE INFORMÁTICA**

GRADO EN INGENIERÍA INFORMÁTICA

TRABAJO FIN DE GRADO

**Herramienta para el análisis, seguimiento y
visualización de los precios de los productos
agrícolas**

Miguel Ángel García Quilón

Septiembre, 2017

HERRAMIENTA PARA EL ANÁLISIS, SEGUIMIENTO Y VISUALIZACIÓN DE
LOS PRECIOS DE LOS PRODUCTOS AGRÍCOLAS



UNIVERSIDAD DE CASTILLA-LA MANCHA
ESCUELA SUPERIOR DE INFORMÁTICA

Tecnologías de la Información

TECNOLOGÍA ESPECÍFICA DE
TECNOLOGÍAS DE LA INFORMACIÓN

TRABAJO FIN DE GRADO

**Herramienta para el análisis, seguimiento y
visualización de los precios de los productos
agrícolas**

Autor: Miguel Ángel García Quilón

Director: Dr. Francisco Pascual Romero Chicharro

Septiembre, 2017

Miguel Ángel García Quilón

Ciudad Real – España

E-mail: Miguelangel.garcia4@alu.uclm.es

Teléfono: 669 055 741

© 2017 Miguel Ángel García Quilón

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

Se permite la copia, distribución y/o modificación de este documento bajo los términos de la Licencia de Documentación Libre GNU, versión 1.3 o cualquier versión posterior publicada por la *Free Software Foundation*; sin secciones invariantes. Una copia de esta licencia esta incluida en el apéndice titulado «GNU Free Documentation License».

Muchos de los nombres usados por las compañías para diferenciar sus productos y servicios son reclamados como marcas registradas. Allí donde estos nombres aparezcan en este documento, y cuando el autor haya sido informado de esas marcas registradas, los nombres estarán escritos en mayúsculas o como nombres propios.

TRIBUNAL:

Presidente:

Vocal:

Secretario:

FECHA DE DEFENSA:

CALIFICACIÓN:

PRESIDENTE

VOCAL

SECRETARIO

Fdo.:

Fdo.:

Fdo.:

Resumen

Con el objetivo de ampliar el abanico de herramientas que existen a disposición de los agricultores y ganaderos, se ha planteado la creación de una plataforma web, que sea accesible de manera gratuita a cualquier persona cuya actividad pueda estar relacionada con estos sectores, pero especialmente para estos profesionales.

AgroEsi, es una plataforma web, donde agricultores y ganaderos pueden consultar los precios que alcanzan en lonja las diferentes materias primas de ambos sectores, además en la base de datos a la que tiene acceso la plataforma, existe un histórico de precios de modo que los profesionales no sólo puedan ver los últimos precios publicados, si no que también puedan ver el comportamiento o la progresión de los precios en el tiempo. En el caso de la agricultura, la plataforma también cuenta con estadísticas con datos de producciones, superficie cultivada o importaciones/exportaciones.

Por otro lado, el clima es un factor clave que puede determinar el futuro de los negocios en ambos sectores, por este motivo, la plataforma cuenta con un apartado donde se puede consultar la previsión diaria, y donde se pueden visualizar valores como la temperatura o precipitaciones medias en años anteriores.

Pero un agricultor y/o ganadero no necesita sólo saber lo que ocurre en su negocio, también necesita saber lo que sucede fuera de éste. Motivo por el cual, la plataforma dispone de un apartado donde se pueden consultar las noticias que se publican a diario y que están relacionadas con ambos sectores.

Sin embargo, esta idea sólo puede verse cumplida si tanto agricultura como ganadería empiezan a evolucionar de manera adecuada para que tecnologías como “big data” o las denominadas “open” puedan ayudar a los profesionales de este sector en la toma de decisiones. Por otro lado es necesario que toda esta tecnología empiece a ser cada vez más accesible, ya que cada vez son más los factores que pueden afectar a este tipo de negocios, de modo que los agricultores y ganaderos dispongan de herramientas de utilidad, que les puedan ayudar a conocer aún mas lo que ocurre en sus explotaciones y terrenos.

Abstract

With the aim of expanding the range of tools available to farmers and ranchers, it has been proposed to create a web platform that is accessible free of charge to anyone whose activity may be related to these sectors, but especially to these professionals.

AgroEsi, is a web platform, where farmers and ranchers can consult the prices that reach in the auctions, different materials in both sectors, in addition, in the database to which the platform has access, there is a price history so that the professionals can not only see the latest published prices, also can see the behavior or progression of prices over time. In the case of agriculture, the platform also has statistics with production data, amount of cultivated area or imports / exports.

On the other hand, climate is a key factor that can determine the future of business in both sectors, for this reason, the platform has a section where you can check the daily forecast, and where you can see values such as temperature or average precipitation in previous years.

But a farmer or rancher need not only know what is going on in his business, he also needs to know what goes on outside of it. Reason why, the platform has a section where they can consult the news that are published daily and are related to both sectors.

However, this idea can only be fulfilled if both sectors evolve properly so technologies such as “big data” or so-called “open” can help the professionals in this sector in decision making. On the other hand, it is necessary that all this technology becomes increasingly accessible, since more and more factors can affect this type of business, so the professionals have useful tools, which can help them in their business.

Agradecimientos

Después de 7 años, son muchas las personas que en un momento u otro han tenido una palabra o mensaje de ánimo, así que si tu nombre no aparece por aquí, no te preocupes que siempre te estaré agradecido.

A mis padres, nunca os podré agradecer lo suficiente tanto cariño y tanto apoyo en esta etapa que ya llega a su fin. Muchas gracias por todos los sacrificios que habéis tenido que hacer para que hoy pueda estar escribiendo estas líneas. Gracias por enseñarme el valor del trabajo, a luchar día a día y que un chico de una familia humilde también es capaz de lograr todas sus metas.

A mi hermano Alan, claro ejemplo de que el trabajo y la constancia tienen al final su recompensa, eres un referente en lo profesional y en lo personal, así como un ejemplo a seguir. Eres todo lo que un hermano pequeño puede desear, y gracias por ese fantástico sobrino que tú y Elisabeth me regalasteis hace ya un año, otro motivo más para seguir luchando por esta meta.

A todo el grupo de amigos de El Torno, que en un momento u otro os habéis preocupado por cómo me iban las cosas, sin olvidarme de Sara, pilar fundamental durante mi primer año, que tanto tiraste de mí cuando peor lo estaba pasando, y por supuesto a Irene y Mario, que durante todos estos años habéis estado siempre ahí en las “duras y en las maduras”, siempre con un mensaje de aliento en época de exámenes o si se juntaba el trabajo, y convirtiendo cada pequeño paso conseguido en un motivo de alegría y fiesta. Porque este trabajo de fin de grado también es mérito vuestro, gracias.

Y por último, y no menos importante a tí, Francisco Pascual, por toda la ayuda ofrecida y por esa respuesta amable que siempre has tenido a mis dudas y agobios cuando algo no salía. Gracias por la oportunidad que me diste de poder hacer este TFG contigo, ha sido una bonita experiencia que siempre recordaré y por la que te estoy agradecido. Creo que todo el mundo debería hacer contigo el trabajo de fin de grado, tú ya formas parte de ese bonito grupo de recuerdos que me llevo de esta etapa.

Miguel Ángel

A mi madre

Índice general

Resumen	III
Abstract	IV
Agradecimientos	V
Índice general	VII
Índice de tablas	X
Índice de figuras	XI
Índice de listados	XIII
Listado de acrónimos	XIV
1. Introducción	1
1.1. Contexto	1
1.2. Motivación	2
1.3. Estructura del documento	3
2. Objetivos	5
2.1. Objetivo general	5
2.2. Objetivos específicos técnicos	5
2.3. Objetivos específicos docentes	6
3. Antecedentes	7
3.1. Big Data y Business Intelligence	7
3.1.1. Big Data	7
3.1.2. Business Intelligence	9
3.1.3. Data Lake	11
3.2. Open Data	12

3.3.	Big Data y BI en Agricultura y Ganadería	14
3.4.	Soluciones existentes para Agricultura y Ganadería	17
3.5.	Otras iniciativas	21
4.	Metodología	23
4.1.	Programación Extrema	23
4.1.1.	Historias de Usuario	23
4.1.2.	Roles	24
4.1.3.	Prácticas	24
4.1.4.	Proceso	25
4.2.	Herramientas de ayuda	26
4.2.1.	Road Map y Vision Board	26
4.2.2.	Kanban	26
4.2.3.	Trello	28
4.3.	Marco tecnológico	29
4.3.1.	Medios Hardware	29
4.3.2.	Medios Software	29
5.	Resultados	33
5.1.	Visión General	33
5.1.1.	Vision Board	33
5.1.2.	Road Map	34
5.1.3.	Distribución del proyecto	36
5.2.	Arquitectura	37
5.3.	Versión 1: Visualización básica	41
5.3.1.	Contexto de la versión	41
5.3.2.	Desarrollo de la versión	42
5.4.	Versión 1.0.1: El Torno.	45
5.4.1.	Contexto de la versión	45
5.4.2.	Desarrollo de la versión	45
5.5.	Versión 1.0.2: Zona de los montes	49
5.5.1.	Contexto de la versión	49
5.5.2.	Desarrollo de la versión	49
5.6.	Versión 1.0.3: Ciudad Real	52
5.6.1.	Contexto de la versión	52
5.6.2.	Desarrollo de la versión	52

6. Conclusiones	55
6.1. Opinión personal	56
6.2. Trabajo futuro	57
6.2.1. Aplicación Android	57
6.2.2. Análisis Predictivo	57
6.2.3. Datos de nuevos mercados	58
6.2.4. Gestión de usuarios	58
6.2.5. Migración a Elastic Cloud	58
A. Fuentes de datos	60
B. Manual de usuario	61
B.1. Requisitos de uso	61
B.2. Página principal. Sin inicio de sesión	61
B.3. Clima	62
B.4. Noticias	63
B.5. Página principal. Con inicio de sesión	64
B.6. Agricultura	64
B.7. Ganadería	65
C. Mapeos de ElasticSearch	66
C.1. Versión 1: Visualización básica	66
C.2. Versión 1.0.1: El Torno	67
C.3. Versión 1.0.2: Zona de los montes	67
C.4. Versión 1.0.3: Ciudad Real	68
C.5. Estadísticas ES:	68
D. Historias de usuario	69
E. Servicios de actualización de datos	75
E.1. Funcionamiento	75
E.2. Implementación e instalación	75
E.3. Datos	77
E.4. Nuevo esquema de funcionamiento	77
Referencias	78

Índice de tablas

5.1. Relación de productos y códigos	44
5.2. Datos cereales	50
6.1. Objetivos del proyecto	56
A.1. Fuentes de precios y estadísticas.	60
A.2. Fuentes de noticias.	60
A.3. Fuentes de clima.	60

Índice de figuras

1.1. Esquema del proyecto	3
3.1. Aplicación VisualNacert	17
3.2. Aplicación AlcuzApp	18
3.3. Aplicación Agromapping	18
3.4. Aplicación Ganges	19
3.5. Servicios DelaVal	19
3.6. Esquema Booster	20
3.7. Aplicación Cropti	20
3.8. Mapa mundial de portales Open Data	22
4.1. Tablero kanban creado con Trello	28
5.1. Esquema aplicaciones ASP.NET MVC	38
5.2. Diagrama de secuencia MVC	38
5.3. Lista de modelos	40
5.4. Lista de vistas	40
5.5. Lista de controladores	40
5.6. Adquisición de datos	43
5.7. Lonja de Albacete	46
5.8. Lonja de Talavera	46
5.9. Portal de noticias “OVIespaña”	47
5.10. Gráficos y columna de noticias.	48
5.11. Indicadores de precios.	48
5.12. Relación de datos de cada precio.	48
5.13. Ejemplo de <i>Noticia</i>	51
5.14. Menú noticias.	51
5.15. Gráficos	54
B.1. Opciones disponibles sin iniciar sesión	61

B.2. Últimos titulares	62
B.3. Precipitaciones y temperaturas medias mensuales	62
B.4. Otras estadísticas climáticas	63
B.5. Organización de noticias	63
B.6. Ejemplo de búsqueda	64
B.7. Opciones disponibles después iniciar sesión	64
B.8. Cuadro de mandos	64
B.9. Cuadro de mandos	65
B.10. Gráfico precios Arroz.	65
B.11. Gráfico precios Trigo.	65
B.12. Indicadores ganadería	65
C.1. Mapeo ES tipo “precio”.	66
C.2. Mapeo ES tipo “noticia”.	67
C.3. Mapeo ES estadísticas	67
C.4. Mapeo ES Clima	68
C.5. Número total de registros de tipo <i>Precio</i>	68
C.6. Número total de registros de tipo <i>Noticia</i>	68
C.7. Número total de registros de tipo <i>Estadística</i>	68
E.1. Instalación del servicio.	76
E.2. Diagrama de clases (Servicio “Lonjas”).	76
E.3. Esquema final de la plataforma.	77

Índice de listados

Listado de acrónimos

ES	ElasticSearch
BI	Business Intelligence
BD	Big Data
MVC	Modelo Vista Controlador
ES	ElasticSearch
XP	Extreme Programming

Capítulo 1

Introducción

EN este capítulo se pretende dar una breve introducción del problema abordado, y de el contexto en el que se enmarca el mismo, así como la solución propuesta para resolverlo, finalmente este capítulo concluye con una sección en la que se describe la estructura del presente documento.

1.1 Contexto

Según estimaciones de la FAO, para el año 2050 la población mundial habrá crecido un 34 por ciento ¹, por otro lado, se prevé una disminución de la tierra para el cultivo, debido en gran medida, a la utilización de muchos de los espacios de cultivo para la creación de biocombustibles. Es por ello que la sociedad se enfrenta al reto de tener que aumentar la producción de alimentos en una menor superficie cultivable con menos mano de obra y graves problemas de escasez de agua.

Si en la década de los 60 se produjo la “Revolución verde”, que consiguió incrementar la productividad agrícola gracias a la mejora y selección de semillas o la evolución de los sistemas de riego, en este momento están empezando a aparecer ideas o conceptos de una “agricultura inteligente” ² o “agricultura de precisión” [Zha15], en la que el sector agrícola se apoye en las Tecnologías de la Información.

Hoy en día y gracias al uso de las tecnologías “big data”, se podrán tomar decisiones casi en tiempo real, a través de la captura, almacenamiento y análisis de una gran cantidad de datos procedentes de sensores localizados en el terreno, midiendo la humedad o el PH de la tierra, o por el aire, tomando imágenes aéreas. En la maquinaria agrícola, los profesionales podrán reemplazar piezas de manera proactiva, todo ello con el fin de evitar que la productividad de sus explotaciones se vea afectada, así como la mejora de éstas, eligiendo los tipos de cultivos más rentables para las características y condiciones climáticas donde se sitúan sus terrenos. En la ganadería³ también se pueden conseguir mejoras monitorizando la salud de los animales mediante sensores.

¹<http://www.fao.org/wsfs/forum2050/wsfs-forum/es/>

²<http://www.madrimasd.org/blogs/universo/2014/06/27/144648>

³<https://www.dinsa.es/work/la-tecnologia-big-data-llega-al-sector-ganadero-de-madrid/>

Pese a que todas las tecnologías base necesarias, ya están disponibles, “sólo” falta que el sector evolucione de manera adecuada, como ya lo están haciendo otros sectores, y esta evolución pasa por el uso de tecnologías “open”.

Para los agricultores sería más accesible la tecnología, y la mejora en sus explotaciones ampliaría sus márgenes económicos lo necesario para cubrir los gastos y aún así mejorar sus ingresos. Aunque para algunos expertos en el sector ⁴, todavía es demasiado pronto para empezar a hablar de Big Data en la agricultura.

1.2 Motivación

La idea clave es que los agricultores y ganaderos puedan anticiparse al futuro, mediante el uso de tecnología Big Data y el uso de analítica predictiva, debido en gran medida a que las enfermedades, plagas y las condiciones meteorológicas convierten a la agricultura y la ganadería en negocios de altísimo riesgo.

Con el fin de afrontar esta idea, lo que se plantea es crear una aplicación web que permita a cualquier persona relacionada con los sectores agrario o ganadero, la visualización, análisis y seguimiento de los precios de diferentes productos agrícolas y ganaderos. Las fuentes incluyen datos a nivel nacional como internacional, datos meteorológicos o noticias relacionadas con ambos sectores. El núcleo principal estará compuesto por varios procesos encargados de la captura, transformación y gestión de los datos contenidos en el *data lake*[Inm16] que se ha empleado durante el desarrollo.

Durante el proceso de captura, el crawler se ocupará de explorar las diferentes fuentes de datos elegidas, y descargar los datos. Después, durante el proceso de transformación de los datos, y utilizando técnicas de *web scraping*⁵, se extraerán los datos, y serán normalizados de acuerdo con el modelo de datos elegido, por último, utilizando el formato más recomendable, se almacenarán en el *data lake*.

Para la implementación del proceso de gestión de los datos, se implementará una clase *broker* que estará conectada con el *data lake* y que se encargará de interactuar con éste, así como con la aplicación web, para responder a las peticiones de datos de ésta.

⁴<http://www.parajeinnova.com/big-data-y-data-mining-en-la-agricultura/>

⁵<https://sitelabs.es/web-scraping-introduccion-y-herramientas/>

En la figura 1.1 se muestra un esquema general del proyecto.

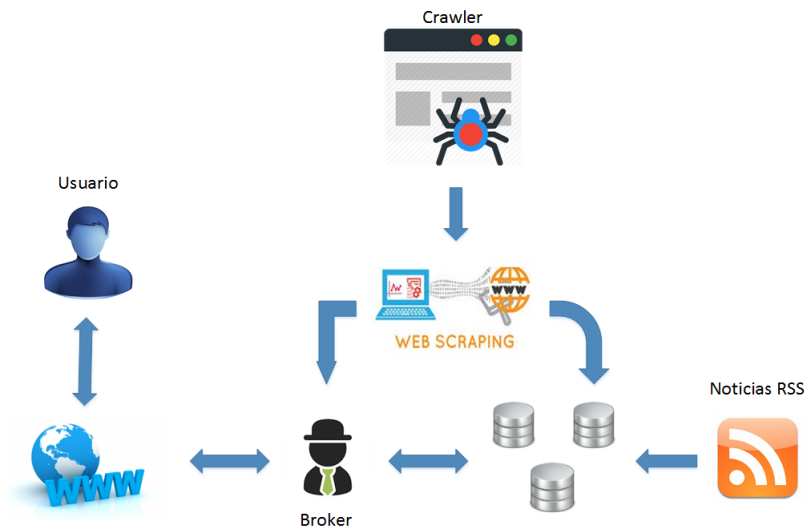


Figura 1.1: Esquema del proyecto

1.3 Estructura del documento

El presente documento ha quedado organizado en diferentes capítulos, los cuales se detallan a continuación:

Capítulo 1: Introducción En este capítulo se describe el ámbito en el que se enmarca la herramienta así como una justificación del trabajo realizado.

Capítulo 2: Objetivos En este capítulo se ofrece una lista detallada de los objetivos que se pretenden cumplir con el desarrollo de este proyecto.

Capítulo 3: Antecedentes En este capítulo se realiza una descripción del “estado del arte” existente antes de la realización del proyecto, en el que se explican conceptos clave en los que está basada la herramienta implementada así como una serie de soluciones afines que ya están en el mercado. Mencionar en este punto, que las imágenes presentes en el apartado 3.4. son capturas de las respectivas web de cada aplicación.

Capítulo 4: Metodología En este capítulo se describe el procedimiento empleado para desarrollar el trabajo, las herramientas que se han empleado para la división y planificación del trabajo y los medios hardware y software empleados durante el desarrollo del proyecto.

Capítulo 5: Resultados En este capítulo se ofrece una visión general, la arquitectura de la plataforma implementada así como las distintas versiones que se han llevado a cabo de la misma para cumplir con los objetivos propuestos.

Capítulo 6: Conclusiones En este capítulo se resumen los objetivos que se han alcanzado con el trabajo realizado, una breve opinión personal, así como una serie de propuestas para mejorar y/o añadir funcionalidades a la herramienta.

También se han incluido una serie de anexos, los cuales incluyen información complementaria a la contenida en este documento y que se explican a continuación:

Anexo A. Fuentes de datos Contiene la relación de fuentes utilizadas para la adquisición de los datos que hay presentes en la base de datos que utiliza la herramienta implementada.

Anexo B. Manual de usuario Contiene los requisitos esenciales para la ejecución de la herramienta así como una descripción de la estructura de la plataforma, de modo que el usuario se pueda mover por la aplicación fácilmente.

Anexo C. Mapeos de ElasticSearch Contiene las definiciones empleadas en la base de datos a la hora de almacenar cada tipo de dato, así como las estadísticas obtenidas una vez terminó el proceso de carga de datos.

Anexo D. Historias de usuario Contiene la lista de especificaciones iniciales de la aplicación que se ha tenido como referencia a medida que se han ido desarrollando las diferentes versiones de la herramienta.

Anexo E. Servicios de actualización de datos Contiene la especificaciones de los mecanismos utilizados para mantener actualizados los datos, la estructura de los mismos y su instalación. Por último se describe la lista de productos cuyos precios ha sido posible actualizar gracias a su ejecución.

Por último la sección **Referencias** contiene la lista de direcciones URL y referencias bibliográficas empleadas para el desarrollo del proyecto.

Objetivos

EN este capítulo se pretende ofrecer una visión clara de cuáles son los objetivos a cumplir con el desarrollo de la herramienta, de modo que se puedan comprobar las funcionalidades que ésta ofrece y así poder evaluar el resultado final obtenido tras su implementación.

2.1 Objetivo general

El objetivo principal del proyecto es la implementación de una aplicación web que permita la visualización, el análisis y el seguimiento de los precios de diferentes productos agrícolas y ganaderos, cuyos datos estén previamente almacenados, y los cuales van a ser gestionados por la aplicación web según las necesidades del usuario que la utilice. A continuación se enumeran divididos entre técnicos y docentes los objetivos específicos del proyecto.

Los objetivos técnicos que se describen a continuación tratan de enumerar las funcionalidades que la herramienta ofrece así como una breve descripción de cómo éstas, van a ser implementadas.

2.2 Objetivos específicos técnicos

- Se implementará un proceso de crawling encargado de obtener sólo referencias concretas a datos, procedentes de las diferentes fuentes de datos.
- Debido a los distintos tipos de datos, se implementará un proceso de adquisición de datos, dependiendo de la naturaleza de los mismos, incluyendo noticias RSS.
- Para una mejor gestión de las operaciones con la base de datos, se implementará una clase *broker*, encargada entre otras operaciones, de almacenar los datos en el *data lake* de acuerdo con los objetivos anteriores.
- Con el fin de disponer de una herramienta que permita a los usuarios visualizar, analizar y realizar el seguimiento de la cotización de los productos agrícolas, se implementará una aplicación web usando el patrón MVC.

2.3 Objetivos específicos docentes

Los objetivos docentes descritos a continuación tratan de describir los conocimientos que han sido necesarios adquirir, para poder cumplir los objetivos anteriormente expuestos.

- Aprender algunas de las técnicas existentes para la búsqueda y adquisición de datos procedentes de sitios web, así como técnicas para el manejo de noticias RSS.
- Aprender y comprender algunas de las tecnologías Big Data que se utilizan hoy en día.
- Aprender y poner en práctica patrones de diseño para el desarrollo de aplicaciones web.
- Comprender técnicas que sirven para extraer conocimiento de los datos.

Capítulo 3

Antecedentes

EN este capítulo se pretenden explicar algunos conceptos clave para dar al lector una idea del marco en el se sitúa la herramienta así como una visión general del “estado del arte”, mostrando algunas herramientas que ya existen en la actualidad, así como iniciativas relacionadas, y que tratan de ayudar a los profesionales de la ganadería y la agricultura en sus actividades diarias.

3.1 Big Data y Business Intelligence

El tratamiento de grandes volúmenes de datos, “Big Data”, está creciendo rápidamente a medida que las empresas están dedicando cada vez más recursos para poder aprovecharlos, así como datos externos procedentes de redes sociales y otras fuentes externas.

La integración de análisis avanzados para Big Data con sistemas de Business Intelligence (BI) es un paso fundamental para poder obtener beneficios [PG16]. Mientras que el análisis proporciona la perspectiva más compleja de los datos, los sistemas de BI (Business Intelligence) ofrecen a los usuarios una perspectiva más estructurada y sencilla de los mismos. La clave reside en que los sistemas de BI (Business Intelligence) mediante cuadros de mando, herramientas de visualización y o generación de informes, resultan vitales para que los resultados del análisis de datos resulten útiles¹.

A continuación se explican estos conceptos más en detalle.

3.1.1 Big Data

El término Big Data² se refiere al conjunto de técnicas encargado del tratamiento de grandes cantidades de datos, cuyo propósito es la extracción del conocimiento presente en los mismos, para que sirva en la toma de decisiones y en las actividades de un negocio.

Cuando se habla de este concepto, también se suele utilizar la definición de las tres V, que representan el gran Volumen de datos que debe ser capaz de tratar, la Velocidad con la que puede procesar esos datos, y la Variedad de formas que pueden tomar los mismos. En ocasiones se añade una cuarta V, la del Valor obtenido gracias al conocimiento extraído.

¹<https://goo.gl/hcBof2>

²<http://www.iic.uam.es/big-data/>

Los proyectos dedicados a la aplicación de éstas técnicas se componen de tres elementos clave:

- **Almacenamiento:** Los recursos hardware y software almacenan los datos de manera distribuida y redundante, facilitando su acceso y disponibilidad.
- **Procesamiento:** Las herramientas de procesamiento de datos proporcionan la base para poder trabajar estas cantidades de datos.
- **Análisis:** Los algoritmos y métodos analizan los datos, permitiendo la extracción de conocimiento útil para el negocio.

Para la extracción de este conocimiento, los análisis que se aplican se basan en técnicas matemáticas[JA13], como la estadística, además, también se realizan análisis procedentes de campos como la minería de datos³ o el aprendizaje automático⁴. Gracias a estos procedimientos es posible utilizar cualquier tipo de dato; desde registros hasta bases de datos, pasando por archivos de contenido multimedia o la actividad de un usuario en una aplicación web.

El análisis del que se habla en el párrafo anterior, se conoce con el término de “Big Data Analytics” y aporta beneficios tales como velocidad y eficiencia. Hace unos años, una empresa habría tenido que reunir la información para poder ejecutar los análisis pertinentes y así obtener conocimiento de esos datos para la toma de decisiones, en la actualidad, las empresas pueden obtener ese conocimiento de sus datos, de inmediato. Tener la capacidad de trabajar de una manera más rápida y ágil proporciona a las empresas una ventaja competitiva.

La importancia de este concepto se debe a que ayuda a las empresas a aprovechar sus datos e identificar nuevas oportunidades, lo que conduce a su vez a movimientos de negocio más inteligentes, operaciones más eficientes y mayores beneficios. La realización de este tipo de análisis de datos tiene los siguientes beneficios⁵:

- **Reducción de costes:** Las tecnologías Big Data, como Hadoop o el análisis de datos basado en la nube, aportan importantes ventajas en términos de costes cuando se trata de almacenar grandes cantidades de datos, además de identificar maneras más eficientes de hacer negocios.
- **Toma de decisiones mejorada:** Gracias a la capacidad de analizar nuevas fuentes de datos, las empresas pueden conocer mejor su entorno de manera inmediata lo que implica una mejora en la toma de decisiones.
- **Nuevos productos y servicios:** Con la capacidad de medir las necesidades y satisfacción de los clientes, viene la capacidad de poder dar a los mismos lo que quieren.

³<https://goo.gl/ITIKoU>

⁴<http://ingenierobeta.com/que-es-machine-learning/>

⁵<https://goo.gl/BeUBII>

Según el nivel de valor de la información obtenida pueden distinguirse tres clases de analítica [Cor16]:

- **Analítica descriptiva:** Consiste en recopilar y almacenar los datos que se generan en una organización, en el tiempo, con el objetivo de crear un histórico de datos, de forma que puedan ayudar a la comprensión del estado actual y pasado del negocio. La analítica descriptiva explica el funcionamiento de un negocio en el tiempo.
- **Analítica predictiva** [Sie13]: Funciona sobre la analítica descriptiva y utiliza modelos estadísticos para añadir información que la organización desconocía. De esta forma la analítica predictiva describe el funcionamiento futuro de un negocio.
- **Analítica prescriptiva:** Explora los niveles anteriores para indicar qué acciones de negocio producirán los mejores resultados. Este tipo de analítica es capaz de ofrecer recomendaciones de manera automática sobre el momento más óptimo para realizar mantenimientos u otras operaciones cuantificables. La analítica prescriptiva describe qué se puede hacer para optimizar un negocio.

En cuanto a las ventajas que ofrecen las soluciones Big Data:

- Mayor conocimiento del negocio, a través de la analítica descriptiva.
- Mejores capacidades organizativas y de planificación del negocio mediante el uso de analítica predictiva.
- Mejora de los márgenes de beneficio al emplear técnicas de analítica prescriptiva que recomienden acciones basadas en datos sólidos acerca de los costes y beneficios que se esperan.

3.1.2 Business Intelligence

Business Intelligence⁶, es un término que se refiere a la capacidad de conseguir información útil para un negocio, a partir de los datos que éste posee con el objetivo de poder conocer en detalle todos los aspectos relacionados con el mismo y así, poder optimizar la toma de decisiones por parte del personal encargado.

En relación con las tecnologías de la información, es posible definirlo como el conjunto de técnicas y herramientas que permiten la adquisición, depuración y transformación de los datos de una empresa, tanto internos como externos, de forma que sirvan de apoyo en la toma de decisiones sobre el negocio.

Las herramientas que se emplean, están basadas en el uso de sistemas de información inteligentes los cuales están formados con datos de distintos orígenes, como aspectos económicos, datos sobre la producción de la empresa, etc.

⁶<https://goo.gl/DGDCfg>

Mediante las herramientas y técnicas ETL⁷ (Del inglés “Extract, Transform, Load”), o ETC (equivalente en castellano: “extracción, transformación y carga”), se extraen los datos de distintas fuentes, se depuran y preparan (homogeneización de los datos), para luego cargarlos en un almacén de datos.

La inteligencia de negocio es un elemento clave para cualquier empresa, ya que les permite obtener información valiosa, que puede ayudar a solucionar los problemas que sus negocios puedan tener, además la posibilidad de aprovechar las nuevas oportunidades que puedan surgir en los mercados donde participan, tener un mejor control de su economía, o encontrar puntos débiles en sus procesos de negocio, lo que se denomina ventaja competitiva⁸.

El éxito de este tipo de software, depende de los beneficios que una empresa o negocio puedan obtener con su uso[CD10]; si se consigue aumentar los márgenes de beneficios, así como la reducción de gastos y consumo de recursos, y por otro lado, las decisiones llevadas a cabo mejoran las actividades de la empresa, entonces el software utilizado se mantendrá en la misma por mucho tiempo, en caso contrario, será cambiado por otro que consiga proporcionar unos mejores resultados.

La inteligencia de negocio, tal como se entiende en la actualidad, es la evolución de los primeros sistemas de apoyo empleados en la década de 1960 y que se fueron desarrollando a mediados de 1980.

Todas las herramientas y metodologías tienen algunas características comunes, las cuales se describen a continuación:

- **Accesibilidad:** Todo gira entorno a los datos, por este motivo, las herramientas y técnicas empleadas tienen que asegurar el acceso a los mismos por parte de los usuarios, sea cual sea el origen de los datos.
- **Apoyo:** No sólo se trata de conseguir una buena presentación de la información, también es necesario que los usuarios de este tipo de aplicaciones puedan seleccionar y modificar aquellos datos que necesiten.
- **Usuario:** Independientemente de los conocimientos técnicos de los usuarios o de sus capacidades, este tipo de herramientas deben estar diseñadas de forma que la información mostrada sea comprensible y lo suficientemente clara.

⁷<https://goo.gl/bHg6EF>

⁸<https://goo.gl/SUh1WX>

3.1.3 Data Lake

Los datos suelen almacenarse en lo que se denomina un “Data Lake” [Res14], es decir, un repositorio de almacenamiento que contiene una gran cantidad de datos en su formato original, incluyendo datos estructurados, semi-estructurados y no estructurados. Ya no se almacena sólo el dato “clásico”, ahora también se tienen en cuenta otros, provenientes del mundo digital, como tweets o los datos procedentes de sensores.

No hay que confundirlo con un Datawarehouse, las principales motivaciones de esta nueva arquitectura de almacenamiento de datos son las siguientes⁹:

- **La información es poder:** Necesidad de recoger y analizar todos los datos que se generan en el entorno de las empresas, con el fin de conocer y mejorar la experiencia de sus clientes. Si en el pasado bastaba con sistemas de BI, ahora, en un mundo donde conviven paginas web, distintos tipos de aplicaciones conectadas a Internet y redes sociales, sólo es posible realizar un análisis eficiente de los datos procedentes de estos canales, con los que interacciona un usuario, si éstos, se encuentran almacenados de acuerdo con la idea en la que se basa el funcionamiento de un Data Lake.
- **Usabilidad:** El valor de estos sistemas de almacenamiento viene proporcionado por su capacidad para poder ser consultados y analizados, no porque tenga una gran cantidad de datos almacenados, quiere decir que un Data Lake resulta más útil. La variedad en los datos almacenados permite que se puedan aplicar distintos tipos de análisis de datos, y sea posible, cumplir diferentes tipos de objetivos clave para un negocio.
- **Capacidad y agilidad:** La idea de que puedan convivir datos de diferentes fuentes y en distintos formatos, asegura la escalabilidad de estos sistemas, pudiendo incorporar nuevas fuentes de datos en cualquier momento, ya que el límite de capacidad de estos sistemas, lo pone la tecnología que les da soporte.
- **Diversidad:** En el Data Lake conviven diferentes modelos de datos, por lo que diferentes modelos de análisis pueden ser desplegados para analizar las relaciones existentes.

⁹<http://www.divisadero.es/el-monstruo-del-data-lake/>

3.2 Open Data

Los datos abiertos son datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, siempre y cuando, se identifique su origen y sean compartidos en el mismo formato en el que fueron adquiridos.

Las principales características de este tipo de datos son las siguientes:

- **Disponibilidad y acceso:** La información debe estar disponible en su totalidad y sin coste de reproducción, preferiblemente accesible desde Internet.
- **Reutilización y redistribución:** Los datos deben ser proporcionados de forma que puedan ser reutilizados y redistribuidos, con la posibilidad de integrarlos con otros conjuntos de datos.
- **Participación universal:** Cualquier individuo debe poder utilizar, reutilizar y redistribuir la información. No debe haber restricción alguna en términos de esfuerzo, personas o grupos. No se permiten restricciones del tipo; “uso comercial” o para ciertos sectores económicos.
- **Actualización de los datos:** No sirve sólo que los datos estén disponibles para cualquier individuo si no que es de vital importancia que los mismos estén actualizados.

La importancia del concepto de “abierto” reside en una idea clave: **interoperabilidad**.

Esta idea se refiere a la capacidad que tienen distintos elementos para trabajar juntos. Esta característica es esencial para construir sistemas complejos y grandes.

La esencia de los datos (o códigos) compartidos es que una parte del material pueda ser integrado con otros elementos. Esta idea es absolutamente fundamental para entender los principales beneficios prácticos de la apertura: el exponencial aumento de la capacidad para integrar distintas bases de datos o conjuntos de datos, con el objetivo de ofrecer unos productos mejores.

Por otro lado, es necesario definir los tipos de datos que pueden ser abiertos y cuáles no. La clave reside en que cuando se ponen datos a disposición del público, éstos no pueden contener información personal, es decir, datos sobre individuos físicos. Del mismo modo, también se puede restringir la publicación de datos gubernamentales por motivos de seguridad nacional¹⁰.

El “Open Data” tiene la capacidad de ser una parte importante de la solución a los problemas globales de producción de alimentos y pobreza nutricional. Puede suponer una verdadera diferencia para las economías en desarrollo, donde la gente todavía no tiene acceso a algunas de las infraestructuras básicas que puedan suponer una manera para una mejor participación en la actividad económica¹¹.

¹⁰<http://opendata.cloudbcn.cat/MULTI/es/what-is-open-data>

¹¹<http://www.bbc.com/news/science-environment-37353865>

En los países desarrollados, la manera de interactuar de los diferentes elementos que participan en la cadena económica, puede diferir, pero los efectos que se persiguen son los mismos. Esto se consigue, fundamentalmente, permitiendo que la información pueda ser liberada por parte de aquellos que tienen el conocimiento, de manera que aquellos que lo necesitan lo puedan utilizar. Si por ejemplo, un centro de investigación o empresa dispone de datos sobre plagas y enfermedades de las plantas, pueden ayudar a los agricultores a entender las enfermedades a las que se enfrentan y cómo adaptar sus métodos de cultivo para reducir el impacto de la enfermedad. Un mejor uso de los datos puede ayudar a los agricultores a encontrar mercados para obtener el mejor beneficio a sus productos, al mismo tiempo que encuentran las mejores variedades de semillas para las condiciones del suelo del que disponen.

En el Reino Unido se está tratando que los propios datos gubernamentales sean abiertos. En un esfuerzo conjunto, más del 40 % de todos los datos abiertos del gobierno del Reino Unido ahora provienen de DEFRA¹² (Department for Environment, Food and Rural Affairs). Hasta el momento esto supone más de 12.000 conjuntos de datos principalmente sobre la alimentación y la agricultura o que pueden ser relevantes en un contexto relacionado con la agricultura.

El caso de España, es un sector que, en el año 2013, tuvo un volumen de negocio entre 450 y 500 millones y empleó entre 4.200 y 4.700 personas¹³; que alberga un potencial de 200.000 millones de euros en Europa si se combina con el big data y puede generar entre 3 y 5 billones de dólares anuales en EE.UU.

Es un hecho la posibilidad que existe de poder obtener beneficios económicos con este tipo de datos, y que además, sirven como base para crear soluciones que pueden ayudar a sectores económicos tales como la agricultura, la alimentación o los servicios sanitarios. Que además permiten crear empleos y ofrecen soluciones útiles para los ciudadanos de acuerdo con sus necesidades¹⁴.

Sin embargo, los datos abiertos no son suficientes por sí mismos. Tenemos que asegurarnos de que hay herramientas a disposición de la población, de forma que puedan resolver sus necesidades.

A pesar de las prometedoras posibilidades que esta idea ofrece, hay que tomárselo con cautela. Los datos personales, son el principal riesgo. Todo el mundo deja atrás su rastro en Internet u otros medios digitales, sin pararse a pensar de que toda esa actividad puede ser averiguada con distintas herramientas. Es por esto, que hay que hacer un mejor uso de los datos, asegurándonos de que la gente que los utiliza sepa qué datos hay sobre ellos, cómo se usan y quién los está viendo.

¹²<https://goo.gl/GnrtSy>

¹³<http://datos.gob.es/es/node/2347>

¹⁴<http://datos.gob.es/es/node/2347>

Por lo que es responsabilidad de los profesionales que los manejan, eliminar cualquier información personal de los conjuntos de datos. A veces esto es fácil, pero a veces es extremadamente difícil¹⁵.

En el contexto agrícola, los datos abiertos tienen el potencial de permitir el intercambio de información y permitir así una cultura de mejora continua.

3.3 Big Data y BI en Agricultura y Ganadería

Con frecuencia se observa como en los medios de comunicación y en la opinión pública, tanto la agricultura como la ganadería suelen ser los sectores económicos que menos importancia reciben. Puede resultar fácil pensar que las herramientas de BI y las tecnologías BD sólo sirven para los negocios financieros, sin embargo todas las empresas de cualquier sector económico se pueden beneficiar con su uso, incluidas las pertenecientes a estos dos sectores.

Pese a tener un peso importante en el desarrollo de cualquier tipo de sociedad y con unas necesidades, que desde el punto de vista tecnológico, no difieren mucho de las de cualquier otro sector, son poco conocidas por parte de la población¹⁶.

La agricultura y la producción de alimentos deben evolucionar de manera considerable en el futuro, con ayuda de la tecnología, ya que hasta la fecha, hay muy pocos casos donde se haya utilizado la tecnología con el fin de obtener una mejora en las actividades o en las producciones obtenidas de los cultivos. En este momento, parece que se está produciendo un punto de inflexión, en el que cada vez son más las iniciativas que se dan a conocer, y que tratan de obtener los mismos resultados que ya se obtienen en sectores como la industria, el transporte o el comercio¹⁷.

Debido a la rapidez con la que avanzan tanto Internet como la tecnología en general, aunque sea difícil de asimilar, no es extraño pensar que en poco tiempo, será posible la implantación de nuevos modelos de producción de alimentos, basados en la aplicación de distintos tipos de tecnologías y que seguramente cambien los hábitos de alimentación de la población.

Una de las iniciativas más conocidas de la aplicación de la tecnología a la producción de alimentos es la Open Agriculture Initiative¹⁸, desarrollada en el MIT, tiene como objetivo, mejorar las actividades que se llevan a cabo en el campo, para que el trabajo de agricultores y ganaderos sea más eficiente y que los alimentos producidos y consumidos a diario, sean de una mejor calidad. Tanto el hardware como el software empleado, es “open”.

¹⁵<https://governmentscienceandengineering.blog.gov.uk/2016/10/05/the-opportunities-and-challenges-of-open-data-for-agriculture/>

¹⁶<https://blog.es.logicalis.com/analytics/soluciones-big-data-analytics-para-el-sector-agricola>

¹⁷<http://agriculturers.com/aplicando-la-tecnologia-a-la-agricultura-podremos-salvar-el-mundo/>

¹⁸<https://www.media.mit.edu/groups/open-agriculture-openag/overview/>

En segundo lugar, cada vez son más los casos donde se producen inversiones para el desarrollo de herramientas de análisis de datos y de inteligencia de negocio, tanto para agricultura como para ganadería¹⁹.

En un futuro muy cercano, las herramientas de análisis predictivo pueden desempeñar un papel fundamental a la hora de desarrollar el potencial de este sector, según afirman un gran número de expertos en la materia.

Sin embargo, los retos a los que se enfrentan las actividades diarias de estos sectores, no son fáciles e incluso difíciles de resolver²⁰:

- El factor **climático** es el más influyente para negocios de este tipo, por ello, es necesario la obtención de unas predicciones meteorológicas fiables de acuerdo con la actividad desarrollada y las características del terreno.
- Optimizar los **costes** derivados de las actividades diarias.
- Tratar de reducir el impacto que puedan tener plagas o enfermedades, así como los costes asociados.
- Conocer el comportamiento de los **mercados** y las actividades de la **competencia** para tratar de ofrecer una producción lo más realista posible de acuerdo con la demanda existente.
- Adaptar el modelo y las estrategias de negocio, de forma que las actividades que se llevan a cabo de manera diaria, sean lo más **sostenible** posible.

Como con cualquier negocio, hay un objetivo común, aumentar los ingresos y reducir los gastos. Gracias a este tipo de soluciones, las actividades que se realizan a diario, acaban siendo más eficaces y resulta más sencillo alcanzar los objetivos propuestos.

Además, se reduce la aparición de errores, gracias a que este tipo de herramientas, se encarga de averiguar qué aspectos del negocio funcionan correctamente y cuáles no.

En cuanto a la industria agrícola en particular, hay una variedad de usos y necesidades de soluciones de BI, tales como:

- **Pronóstico de los datos:** Esta es una parte muy importante de dirigir un negocio agrícola. Cuándo se cultivan y cosechan cultivos, o cuándo se venden los animales; los trabajos deben realizarse en determinados momentos y de manera eficiente. Al identificar exactamente cuándo es el mejor momento para obtener la mayor producción posible, puede conseguir aumentar sus ingresos de manera significativa.

¹⁹<http://www.elmundo.es/economia/2015/07/29/55b9175622601dca0c8b4593.html>
2021

- **Impacto ambiental:** Los pronósticos no sirven sólo para obtener un aumento de los beneficios. También pueden mostrar el impacto ambiental que puede ocasionar determinadas prácticas en sus negocios. De esta forma ayuda a las empresas a mejorar su gestión de emisiones y buscar cambios en las políticas.
- **Ventaja competitiva:** La industria agrícola es un mercado competitivo a menudo en el que numerosas compañías compiten por las ventas de un determinado producto. La implementación de soluciones de BI puede ayudar a un negocio a progresar y conocer mejor el mercado en el que participa.
- **Toma de decisiones:** Es vital que las decisiones se tomen correctamente cuando se trata de un negocio, haciendo énfasis en los recursos de los que dispone, proporcionando a los animales, agua, comida y medicinas. No hay lugar para los errores, que es lo que el software de BI debe hacer; proporcionando datos y análisis precisos que erradican cualquier suposición humana que conduzca a resultados desfavorables.
- **Malgasto de recursos:** Los pronósticos en base a los datos permiten a los profesionales conocer qué aspectos de sus negocios funcionan correctamente y cuáles necesitan atención. También habrá una reducción en cualquier sobreproducción que pueda ocurrir cuando las compañías no saben cuál es la demanda real del producto que ofrecen.

3.4 Soluciones existentes para Agricultura y Ganadería

En cuanto a soluciones existentes para este sector se van a comentar algunas:

VisualNacert²², figura 3.1, es una aplicación basada en mapas que aporta al usuario una cantidad ingente de información. La plataforma combina datos climáticos, del suelo o datos de identificación geográfica con datos de los agricultores. Para el agricultor supone un importante ahorro de costes. Los modelos que utiliza la aplicación se personalizan según lo que necesite la explotación, y permiten que los productores controlen en tiempo real, y a través de cualquier dispositivo con conexión a Internet, el estado actual y las necesidades futuras de su terreno. La plataforma, está pensada para empresas medianas y grandes, especialmente para cooperativas, aunque acabará llegando a pequeños agricultores.



Figura 3.1: Aplicación VisualNacert

AlcuzApp²³, figura 3.2, es una aplicación dirigida al sector olivarero, cuyo objetivo es optimizar y acelerar la comunicación entre las almazaras y sus socios. Esta herramienta dota al agricultor de herramientas para la gestión y el control de su producción durante toda la campaña. Entre las utilidades que ofrece la aplicación encontramos; poder conocer la producción media por árbol ya sea en fincas, polígonos o parcelas. También ahorra al agricultor múltiples desplazamientos a las almazaras para solicitar repeticiones de análisis; disponer de toda la información de su cosecha de manera detallada en la palma de su mano; y estar permanentemente informado de qué ocurre en su almazara. En el caso de las almazaras permite: un ahorro de tiempo en gestionar reclamaciones de rendimientos; firma de documentos de IVA, entregas a cuenta, información sobre tratamientos de plagas, etc.

²²<https://visualnacert.com/>

²³<http://www.alcuzapp.com/>

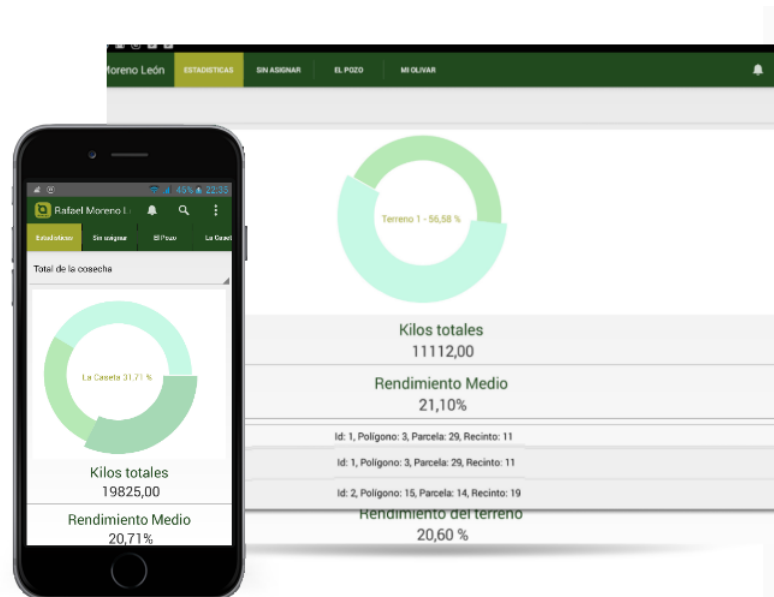


Figura 3.2: Aplicación AlcuzApp

Para algunos expertos, la agricultura basada en macrodatos permite crear “un nuevo ecosistema de trabajo que permite recolectar, almacenar y procesar grandes volúmenes de información, de diferentes orígenes y, a partir de aquí, desarrollar aplicaciones capaces de responder a nuestras necesidades”²⁴.

Agromapping²⁵, figura 3.3, es otro caso de uso de mapas. Esta compañía ofrece mapas donde ver el estado de salud de los cultivos y su variabilidad espacial en la parcela. Esta información permite implementar técnicas de manejo específicas para cada zona.

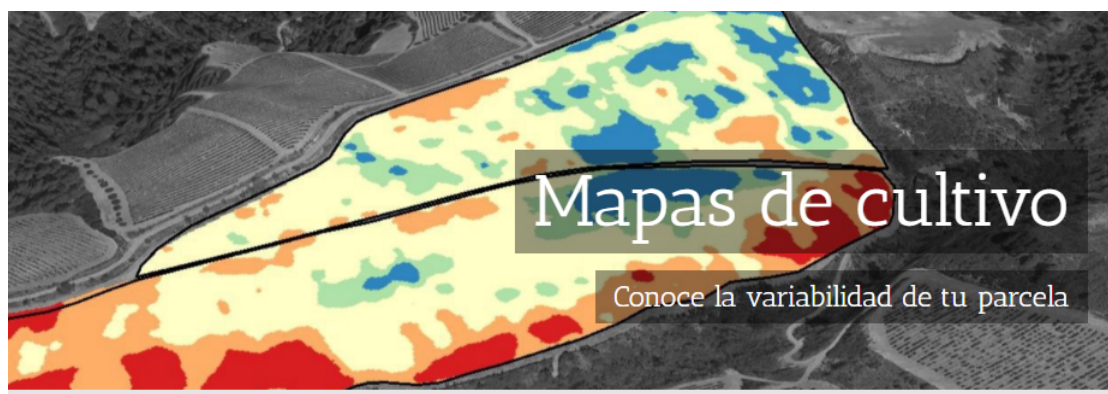


Figura 3.3: Aplicación Agromapping

En definitiva, Big Data y las nuevas herramientas de análisis predictivo abren un horizonte de optimismo para el sector. Un futuro que se escribirá en clave TIC²⁶.

²⁴<http://www.expansion.com/economia-digital/innovacion/2016/05/04/5729adca268e3e5a7c8b45d0.html>

²⁵<http://www.agromapping.com/>

²⁶<https://blog.es.logicalis.com/analytics/soluciones-big-data-analytics-para-el-sector-agricola>

El sector ganadero va un paso por detrás que el agrícola en cuanto a digitalización, pero empiezan a aparecer soluciones tecnológicas que mejoran la productividad del ganado.

Además de soluciones como **Ganges**²⁷, figura 3.4, una plataforma de gestión de vacas cárnicas para pequeños ganaderos, la ganadería intensiva cuenta con otras como **iSagri**²⁸, que permite tratar los principales problemas de las explotaciones ganaderas (porcino, vacuno, ovino y caprino). Esta aplicación pone al día al ganadero de sus obligaciones administrativas y legales, permite realizar un seguimiento individual de los animales, y ayuda en la gestión de costes y recursos.

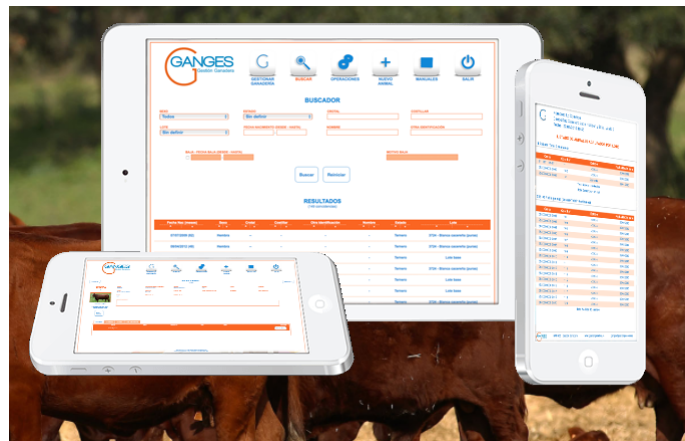


Figura 3.4: Aplicación Ganges

Otras, como **DelaVal**²⁹, figura 3.5, están enfocadas en vacas lecheras. Esta multinacional ofrece al ganadero sistemas informáticos para controlar el ordeño, collares inteligentes para las vacas e, incluso, programas avanzados que analizan al ganado a diario y detectan cuándo necesitan un cuidado especial.



Figura 3.5: Servicios DelaVal

²⁷<http://www.gestionganadera.es/>

²⁸<http://www.isagri.es/Inicio/Inicio.aspx>

²⁹<http://www.delaval.es/>

Veamos a continuación algunas de las startups dedicadas a la agrotecnología más interesantes en España y América Latina:

Booster³⁰ (Argentina), figura 3.6, es un proyecto que tiene como objetivo la aplicación de Inteligencia Artificial en la mejora de las predicciones climáticas en la agricultura en un futuro determinado. Los pronósticos del tiempo proporcionan información crítica sobre el clima futuro. Existen diversos métodos disponibles en la predicción del tiempo, desde la simple observación del cielo a los modelos matemáticos computarizados de alta complejidad. La probabilidad de que las condiciones climáticas del día en consideración coincidirán con el mismo día en el año anterior es muy baja. Pero la probabilidad de que coincidirá en un lapso de quince días del año anterior es muy alta. La intención es crear una API pública con la que cualquiera podrá poner a prueba sus datos. Los resultados actuales están arrojando una predicción climática de hasta un 90 % de precisión media.



Figura 3.6: Esquema Booster

Cropti³¹, figura 3.7, es una aplicación destinada a ayudar a técnicos de Cooperativas y agricultores a gestionar el cuaderno de explotación de sus cultivos. La herramienta tiene una utilidad muy importante para los agricultores ya que permite rellenar y descargar fácilmente el modelo oficial del MAGRAMA, permite la exportación de los datos y ordenada por campañas, cultivos o parcelas, con el objetivo de cumplir con la Normativa que aplica a las explotaciones agrícolas con ayudas PAC.



Figura 3.7: Aplicación Cropti

³⁰<https://loogic.com/booster/>

³¹<https://cropti.com/>

3.5 Otras iniciativas

Pero no sólo pueden proporcionar soluciones al sector primario tecnologías como Big Data o Business Intelligence, la aplicación de las ideas del movimiento Open Source aplicadas a la producción de alimentos es una de las que más impacto puede tener en los próximos años, como la iniciativa Open Agriculture del MIT comentada anteriormente.

Un ejemplo muy llamativo de esta situación es la experiencia compartida por el granjero canadiense Matt Reimer³² que, utilizando software libre, con los conocimientos adquiridos en un curso gratuito sobre Python del MIT, ha modificado el software utilizado para controlar drones, lo ha adaptado a su tractor y ha logrado que éste pueda funcionar de forma autónoma. Se ha convertido en un pionero adaptando el software libre a la agricultura.

Ante la necesidad de contar con maquinaria para cosechar sin conductor y viendo que la que existe en el mercado es muy cara y no puede ser controlada desde un ordenador convencional o smartphone, Matt Reimer se puso manos a la obra y ha marcado el camino para que otros muchos agricultores aprovechen la tecnología para darle un impulso a su trabajo y sus negocios.

También se conocen iniciativas para la aplicación de los ya mencionados, datos abiertos, con el fin de que puedan resultar también de utilidad en las actividades del sector primario.

La primera de ellas es el caso del centro de investigación económico de Estados Unidos³³, que tiene como objetivo mejorar e informar acerca de la toma de decisiones tanto pública como privada, para que los profesionales cuyas actividades están relacionadas con la agricultura, el medio rural o la producción de alimentos puedan anticiparse a las tendencias de los mercados o situaciones de emergencia que puedan surgir.

La actividad de este centro cubre los siguientes sectores económicos y políticos:

- **Economía agrícola:** Datos acerca del comportamiento del rendimiento agrícola, así como el tamaño y concentración de las diferentes explotaciones tanto ganaderas como agrícolas, análisis de mercados y proyecciones sobre la oferta, la demanda y los precios de los productos básicos.
- **Alimentación y nutrición:** Programas de seguridad alimentaria, resultados sanitarios, acceso a los alimentos y proximidad de las tiendas, ventas al por menor y comercialización y precio de otros alimentos.
- **Mercados e intercambios comerciales:** Datos acerca de mercados nacionales e internacionales, estado del sector agroalimentario de Estados Unidos en mercados cada vez más globalizados.

³²<https://hackaday.io/project/10697-autonomous-tractor>

³³<https://www.ers.usda.gov/>

- **Recursos y entorno:** Repercusiones económicas de los programas alternativos de conservación, eficacia de las políticas diseñadas para proteger el medioambiente, desafíos del cambio climático y la escasez de agua, así como el aumento de la competitividad agrícola a través de la tecnología.
- **Economía Rural:** Inversiones en las comunidades rurales y la capacidad de éstas para prosperar en un mercado mundial en constante cambio, así como los motores del funcionamiento de la economía rural.

Otro caso es el de Reino Unido³⁴, en el que el propio gobierno está poniendo a disposición pública cada vez mayor cantidad de datos, haciendo que sean accesibles a cualquier persona para que puedan servir de ayuda en la toma de decisiones de sus negocios o puedan servir de ayuda a la hora de implementar nuevas políticas por parte del gobierno basadas en estos datos.

La naturaleza de estos datos va desde los proporcionados por los diferentes departamentos del gobierno central hasta los que son procedentes de otras autoridades públicas o locales.

A nivel mundial también existen otras iniciativas como “Open Data Inception”³⁵, un proyecto creado por la empresa OpenDataSoft³⁶, dedicada a ofrecer un servicio para la transformación rápida y fácil de todo tipo de datos en cualquier tipo de servicio, sin ningún límite en cuanto a volumen de los mismos.

Esta iniciativa trata de recopilar bajo una misma plataforma todos los portales de datos abiertos a nivel mundial, donde profesionales, estudiantes o ciudadanos pueden recurrir para explorar, visualizar y utilizar los datos compartidos.

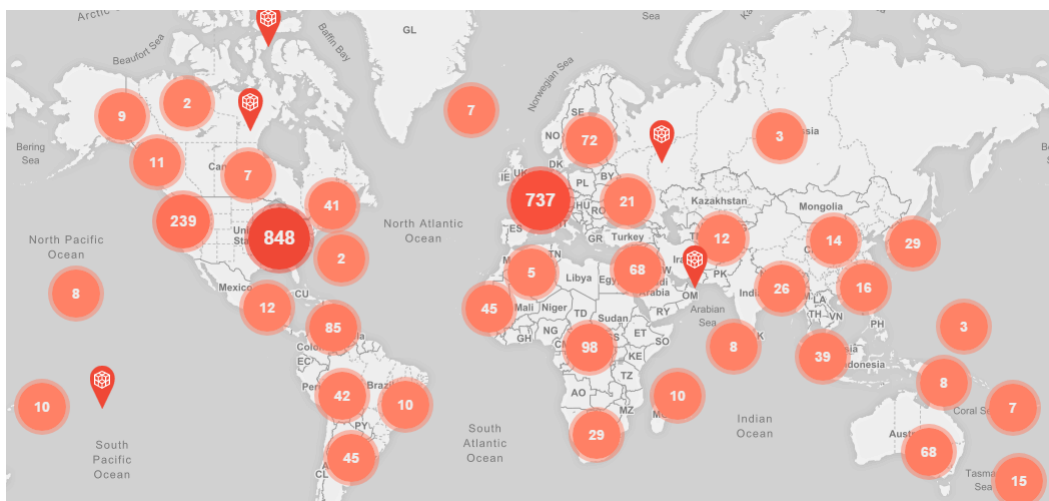


Figura 3.8: Mapa mundial de portales Open Data

³⁴<https://www.gov.uk/>

³⁵<https://goo.gl/gswkxd>

³⁶<https://www.opendatasoft.es/>

Metodología

EN este capítulo va a describir el método de trabajo que se ha llevado a cabo para la gestión del proyecto, también se explicará cómo se ha llevado a cabo el proceso de desarrollo, así como los distintos roles que desempeñan los implicados en el proyecto y se hará una breve descripción del marco tecnológico empleado.

4.1 Programación Extrema

La metodología empleada para el desarrollo de este proyecto ha sido la denominada **Programación Extrema** o **Extreme Programming (XP)** [KB04]. Se trata de una metodología ágil concebida y desarrollada por Kent Beck y que trata de hacer más énfasis en la adaptabilidad que en la previsibilidad, diferenciando esta metodología de las tradicionales. Esta visión considera que los cambios de requisitos sobre la marcha son un aspecto natural en el desarrollo de proyectos software, en los que el equipo de desarrollo es pequeño y en el que el entorno de trabajo sufre de cambios en las especificaciones de el producto. Esta metodología se basa en la simplicidad; la base de este enfoque, un diseño simple agiliza el desarrollo y facilita el mantenimiento, la comunicación; el código resulta más fácil de entender para el resto de programadores de el equipo, cuanto más sencillo sea, y el reciclado continuo de código.

A continuación se enumeran las características esenciales de XP organizadas en los cuatro pilares fundamentales: historias de usuario, roles, prácticas y procesos.

4.1.1 Historias de Usuario

Se trata de las especificaciones de las funcionalidades deseadas de el sistema, comparables con un caso de uso, y que realizan los propios usuarios (clientes) y que se encuentran en el anexo D. Se trata de especificaciones iniciales y son los elementos que emplean los programadores para ir desarrollando las funcionalidades que deberá tener la aplicación o sistema. Cada historia de usuario es lo suficientemente comprensible y delimitada para que los programadores puedan llevarla a cabo en unas semanas, en caso de que la información contenida en la historia sea insuficiente, la comunicación se realizará de manera directa con el cliente. El principal requisito de este enfoque es que al comienzo de cada iteración estén registradas tanto las historias de usuario como los cambios pertinentes.

4.1.2 Roles

Para el desarrollo de este proyecto se ha contado únicamente con dos personas; el estudiante y el director de proyecto. A continuación se describen los roles que ha desempeñado cada uno de ellos de acuerdo con este enfoque:

En cuanto al **estudiante**:

- **Programador:** Pieza básica en este tipo de desarrollos, es el responsable de el diseño, el código y la integridad de el sistema. Debe existir una comunicación y coordinación adecuada entre los programadores y otros miembros del equipo.
- **Encargado de pruebas (Tester):** Apoya al cliente en la preparación de las pruebas funcionales, así como de su ejecución y la posterior publicación de los datos.

En cuanto al **director de proyecto**:

- **Cliente:** El cliente define las especificaciones mediante las historias de usuario, así como las pruebas funcionales para validar su implementación. Por otro lado, asigna prioridades a las historias, decidiendo cuáles aportan mayor valor al negocio.
- **Encargado de seguimiento:** Se encarga de recopilar, analizar y publicar el estado de el cumplimiento de los requisitos de el proyecto, además supervisa el cumplimiento de las estimaciones previstas en cada iteración. Por otro lado, también se ocupa de controlar la marcha de las pruebas funcionales.
- **Entrenador:** Es responsable de el proceso en su conjunto, identifica posibles desviaciones en el desarrollo de las iteraciones y se encarga de guiar al grupo si es necesario. Puede intervenir directamente para atajar un posible problema si es necesario.

Otros roles especificados en esta metodología como el Gestor, el Consultor, etc. no han sido necesarios para la realización del proyecto.

4.1.3 Prácticas

A continuación se describen algunas de las prácticas que se describen en XP y que se consideran relevantes para el desarrollo de este proyecto:

- **Pequeñas entregas:** La idea es producir rápidamente versiones del sistema tan pequeñas como sea posible, siempre y cuando contengan los requisitos de negocio más importantes y sean operativas, aunque no dispongan de toda la funcionalidad esperada, de modo que cada versión sea lanzada en no mas de 3 meses.
- **Metáfora:** Se trata de una historia que todo individuo involucrado puede contar acerca de cómo funciona el sistema, de modo que cualquier persona puede entender el objeto de el sistema o la aplicación que se encuentre en desarrollo.
- **Diseño sencillo:** Cada vez que se implementa una nueva funcionalidad del sistema, es realizada de la forma más sencilla posible, y después de su implementación, se trata

de que el sistema sea lo más simple posible sin perder funcionalidad, de modo que sea más fácil añadir nuevas funcionalidades al sistema para que pueda disponer de nuevas características en el futuro.

- **Pruebas:** Siempre se trata de que no haya ninguna funcionalidad en el sistema que no haya sido probada, mientras los programadores se encargan de escribir las pruebas, los clientes se encargan de realizar las mismas.

4.1.4 Proceso

Durante el desarrollo de un proyecto utilizando este enfoque, el cliente se encuentra mucho más cercano a éste, se podría decir casi, que forma parte de él. Se elimina la fase inicial de captura de requisitos y se permite que se vayan definiendo de forma ordenada en el tiempo que dura el proyecto. Para este enfoque, esta fase gira entorno a la lista de características que el cliente quiere que estén disponibles en el sistema final [Jef00].

1. **Planificación Global:** Se realiza en base a una Visión Global del mismo y a una hoja de ruta.
2. **Exploración:** Se plantean a grandes rasgos las historias de usuario de interés para la entrega. Se evalúan tecnologías a utilizar, cambios de arquitectura, etc. Estas historias de usuario se pueden ver como las diferentes tarjetas incluidas en la herramienta Trello.
3. **Planificación:** El cliente decide la prioridad de cada una de las historias de usuario, de manera que los programadores puedan realizar una estimación sobre el esfuerzo necesario para completarlas. Se realiza un acuerdo sobre el contenido de la primera entrega, para después determinar un cronograma de manera conjunta con el cliente. Cada una de las entregas no puede demorarse por más de tres meses.
4. **Implementación:** Durante esta fase el programador se ocupa de desarrollar las historias de usuario que aportan valor al producto, utilizando mecanismos eficientes para establecer contacto continuo con el cliente y poder resolver dudas en caso de que se produzcan.
5. **Producción:** Se realizan pruebas adicionales, así como una evaluación de de la versión realizada y una puesta en marcha en un servidor. Por otro lado se realiza la documentación de la versión.
6. **Mantenimiento:** Mientras la primera versión se encuentra en fase de producción, el proyecto debe continuar con el desarrollo de nuevas iteraciones que añadan nuevas características descritas en las historias de usuario.

Por último, mientras se completan las iteraciones de este ciclo, tanto cliente como el programador, aprenden. Durante este tiempo éste no debe verse presionado, para no perder calidad en el software implementado, del mismo modo, el cliente debe estar en la obligación de asegurar que el sistema tenga el mayor valor de negocio posible con cada iteración.

4.2 Herramientas de ayuda

A continuación se presentan las diferentes herramientas que se han empleado para realizar con éxito, la metodología anteriormente expuesta.

4.2.1 Road Map y Vision Board

Con el objetivo de tener claro en todo momento de cuál es el resultado final esperado de la herramienta, así como de la planificación que hay realizada sobre el proceso de trabajo, se han empleado dos herramientas concretas para estos fines; *Product Road Map* y *Product Vision Board*.

El primero¹, tiene como propósito informar de la dirección y el progreso conseguidos a los equipos de desarrollo y otras partes externas interesadas, a grandes rasgos, muestra los pasos realizados que han llevado a conseguir ese progreso. No tiene por qué incluir todas las características del producto final, como tampoco errores específicos detectados. La idea es tener una descripción de cómo el producto en cuestión va creciendo.

El segundo², tiene como objetivo describir, visualizar y validar la visión y estrategia de producto. Trata de reunir en un sólo documento, cuáles son las motivaciones para llevar a cabo el producto (**Vision**), cuáles son los clientes potenciales o a qué mercado está dirigido (**Target Group**), el propósito del producto (**Needs**), cuáles son las características que lo diferencian de la competencia (**Product**) y por último, por qué vale la pena (**Business Goals**) para la empresa.

4.2.2 Kanban

Cuando hablamos de *Kanban*, nos referimos a un sistema de trabajo que proporciona un mejor flujo de trabajo, al dividir el trabajo en varias fases perfectamente delimitadas [And10].

Kanban es una palabra que viene de el japonés y que significa “cartel o panel”, elemento clave de este método de producción, el cual se rige por dos objetivos; lograr un producto de calidad y acabar con los “cuellos de botella”, que pueden darse en ciertas fases de trabajo en las que se convierte en primordial, la rapidez por encima de la calidad de el producto.

En cuanto a las reglas que se siguen en este método de trabajo, son cuatro las que se describen en esta metodología de trabajo:

¹<https://www.aha.io/roadmapping/guide/product-roadmap>

²<http://www.romanpichler.com/blog/the-product-vision-board/>

- **Empieza con lo que haces ahora:** Este método de producción no te dice cómo tienes que hacer tu trabajo, se limita a ayudar al equipo en las decisiones sobre si el trabajo que se está realizando es correcto o por el contrario hay que realizar algún ajuste.
- **Aceptación:** No por el hecho de que si algo funciona bien es mejor dejarlo así, no existe la posibilidad de que pueda funcionar mejor. Utilizando este método, todos los miembros de el equipo tienen que estar dispuestos a realizar cambios en cualquier momento, siempre y cuando se realicen poco a poco y con sentido común.
- **Respeto:** Tanto el proceso en curso, como los roles y responsabilidades de cada uno, es imprescindible para este método que cada uno de los miembros de el equipo sepa cuales son sus funciones, de manera que siempre sepan qué hacer en todo momento.
- **Liderazgo:** Otro elemento a tener en cuenta, debe haber iniciativa así como una correcta gestión tanto del equipo como de las tareas, volviendo al punto anterior, cada miembro de el equipo debe saber cuál es su función y cómo ejecutarla.

En cuanto a su aplicación en equipos de desarrollo software, se han obtenidos resultados positivos con la práctica ágil.

Tablero Kanban

Esta herramienta se utiliza para visualizar las tareas a realizar y optimizar el flujo de trabajo en un equipo de desarrollo. Su función es garantizar que este flujo se unifique y que todos los factores de los que dependa o que lo bloqueen sean identificados y resueltos.

En este tablero se localizan las tarjetas de las que anteriormente se ha hablado, las cuales se pueden encontrar en tres estados esenciales dentro de éste: por hacer (**To do**), en el que la tarea contenida en la tarjeta, se encuentra propuesta para su desarrollo en la iteración que se esté realizando en ese momento. En curso (**In progress**), en el que la tarea ha sido escogida para que el equipo pertinente la lleve a cabo. Desde éste, pasa a hecho (**Done**), lo que significa que la tarea ha sido llevada a cabo con éxito y puede pasar a la fase de producción. En otras ocasiones, se habla dos estados adicionales; una vez que una tarea ha sido elegida para llevarse a cabo no siempre pasa al estado *Done*, puede pasar que esta tarea dependa de otras para que pueda ser completada con éxito o que esté a la espera de pasar por unas pruebas determinadas. Por último, también se habla de un estado en el que una tarjeta se puede encontrar mientras se encuentra en el tablero; **Backlog**, en el que una tarea se encuentra programada para que sea llevada a cabo, pero que a diferencia de *To do*, no ha sido propuesta para su desarrollo.

4.2.3 Trello

Se trata de una herramienta gratuita enfocada a la gestión de proyectos, consiguiendo que la colaboración entre equipos de desarrollo o miembros dentro de un equipo, sea más sencilla.

Básicamente consiste en una página web donde existe una serie de listas dispuestas de manera horizontal, de modo que con un simple vistazo es posible comprobar todo lo que hay en un proyecto.

Por otro lado, cada uno de los *items* dentro de cada una de las listas, denominados *cards* (tarjetas), contienen las tareas a llevar a cabo, pudiendo contener imágenes u otros archivos adjuntos. Dependiendo de la prioridad de cada una de las tareas, éstas se encontrarán en la correspondiente lista, pudiendo mover las tarjetas entre diferentes listas de manera sencilla.

A continuación se muestra el estado de el tablero kanban empleado para el desarrollo de este proyecto:

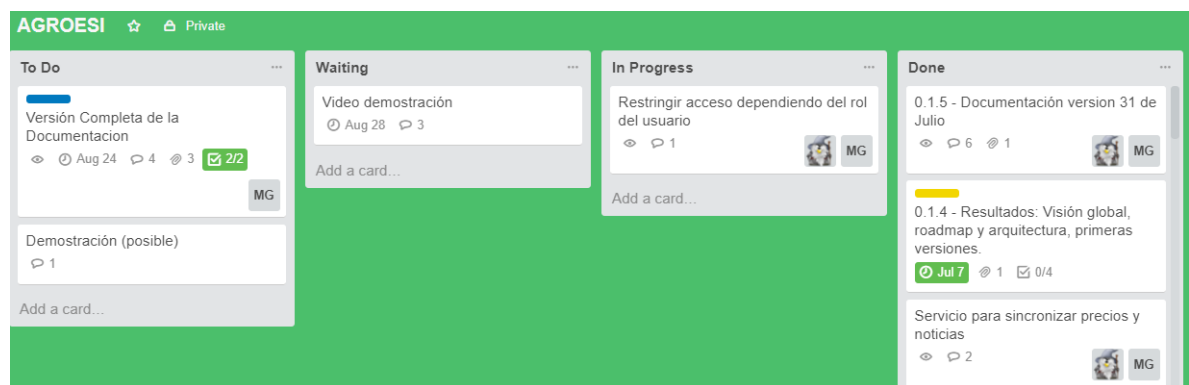


Figura 4.1: Tablero kanban creado con Trello

Como se puede observar en la cabecera de cada una de las listas es posible observar algunos de los estados descritos en la sección 4.2.1, de modo que resulta más cómodo determinar el estado de una tarea dependiendo de la lista en la que se encuentre.

Sin embargo, también presenta algunos inconvenientes, el principal, la necesidad de estar conectado a Internet para poder utilizar esta herramienta.

4.3 Marco tecnológico

EN esta sección se describirán los medios utilizados tanto hardware como software, así como las características que se consideren más relevantes. También se describirán los requisitos técnicos, especialmente software, para el despliegue de la herramienta, así como el resto de elementos que son importantes para el correcto funcionamiento de ésta.

4.3.1 Medios Hardware

El sistema hardware utilizado es un ordenador portátil con las siguientes características: Procesador Intel Core i5-3317U 1.7GHz, 8GB RAM DDR3.

4.3.2 Medios Software

En este apartado se describirán las herramientas software que se utilizarán.

- **Exploración de fuentes y extracción de datos:** Para la exploración de las fuentes de datos se ha empleado “AbotX”³, el cual extiende la funcionalidad y configuraciones que proporciona el conocido WebCrawler escrito en C# “Abot”⁴. Esta extensión se basa en el uso de 2 componentes principales; “CrawlerX” y “ParallelCrawlerEngine”. Mientras que CrawlerX es una sola instancia del crawler, ParallelCrawlerEngine se encarga de crear y gestionar las múltiples instancias existentes.

Para la extracción de los datos se ha empleado “AngleSharp”⁵, una librería para .NET que proporciona la capacidad de extraer el contenido de archivos que contienen datos en formato hiper texto como HTML o XML. El núcleo de esta librería, encargado de la conversión, basa su funcionamiento en la especificación oficial proporcionada por W3C. El resultado es una perfecta representación en formato DOM del archivo proporcionado.

Para el manejo de archivos .pdf se ha utilizado “iTextSharp”⁶, una librería de código abierto para .NET, que permite a los desarrolladores crear, adaptar y mantener documentos en este formato, de modo que es posible añadir estas funcionalidades a proyectos software de manera sencilla. También dispone de una versión para usar en lenguaje Java.

Para el manejo de archivos Excel, se utiliza la librería “ExcelDataReader”⁷. Una librería ligera y rápida, escrita en C# y que es compatible tanto con sistemas Linux como Windows. Además es compatible con diferentes formatos y versiones de este tipo de ficheros.

³<https://abotx.org/>

⁴<https://github.com/sjdirect/abot>

⁵<https://anglesharp.github.io/>

⁶<https://sourceforge.net/projects/itextsharp/>

⁷<https://github.com/ExcelDataReader/ExcelDataReader>

- **Almacenamiento de datos:** Para el almacenamiento de los datos se ha utilizado ElasticSearch⁸ (ES), un motor de búsquedas distribuido, NoSQL, que puede realizar análisis de los datos que almacena en tiempo real. Se suele emplear para búsquedas de textos, búsquedas estructuradas, análisis y combinaciones de las 3 opciones mencionadas anteriormente. Es posible su instalación y uso tanto en un sólo ordenador o portátil como en servidores, además, es compatible con lenguajes de programación tales Java, C#, Python o Ruby.

ES, es de código abierto y basa su funcionamiento sobre *Apache Lucene*⁹, una librería, también de código abierto (con licencia de Apache) para motores de búsquedas de textos, y que permite a los desarrolladores de distintos lenguajes de programación, incluir funcionalidades de indexación y búsquedas de texto. El núcleo de su funcionamiento reside en el concepto de Documento (*Document*) que a su vez contiene Campos (*Fields*).

Sin embargo, ES es mucho más que esta librería y otras tantas funcionalidades añadidas:

- **Almacén de documentos:** Se trata de un almacén de documentos distribuido en tiempo real en el que cada campo es almacenado y puede ser buscado.
- **Motor de búsquedas:** Es capaz de realizar búsquedas a la vez que realiza análisis en tiempo real.
- **Escalabilidad:** Capaz de escalar a cientos de servidores y petabytes de datos estructurados y no estructurados.

ES está orientada a documentos, esto es, que almacena objetos enteros o documentos, pero no sólo los guarda, los indexa de forma que puedan ser accedidos. Los documentos son serializados utilizando JSON, y son accedidos utilizando una API RESTful sobre el protocolo HTTP. Esto significa que se pueden utilizar comandos tales como “GET”, “PUT”, “POST” o “DELETE”, para poder interactuar con ES.

En cuanto al esquema que sigue; los documentos pertenecen a un “Type” (tipo), dependiendo de la naturaleza del objeto o documento almacenado y este tipo se sitúa dentro de un “Index” (Índice). Si se compara ES con las bases de datos tradicionales basadas en SQL, un cluster ES, puede contener múltiples índices, que representan BBDD SQL, y que a su vez pueden contener múltiples tipos, tablas siguiendo el esquema SQL.

Los tipos a su vez mantienen múltiples documentos (filas SQL), y cada documento contiene múltiples campos (columnas SQL).

⁸<https://www.elastic.co/>

⁹<https://lucene.apache.org/core/>

- **Framework aplicación web:** Para el desarrollo de la aplicación web que se encargará de la visualización y el análisis de los datos se ha optado por utilizar *ASP.NET MVC*, en el cual debemos distinguir dos elementos clave; **ASP.NET** y **MVC**.

En cuanto a *ASP.NET*, es un modelo unificado de desarrollo web, desarrollado y comercializado por Microsoft que incluye las características necesarias para poder implementar aplicaciones y sitios web así como servicios web, con el mínimo código posible. Utilizando este modelo, las paginas web se construyen mediante formularios web, lo que conocemos como paginas *.aspx*. Este tipo de archivos contienen generalmente etiquetas HTML, pero también incluyen *Controles Web*, que se procesan en el lado del servidor, y *Controles de usuario* que albergan todo el contenido estático y dinámico requerido por la página web.

Por otro lado, se encuentra *MVC*, acrónimo de (Modelo, Vista, Controlador), se trata de un modelo de diseño que permite separar la lógica y el acceso a los datos de modo que toda la lógica de dominio y el acceso a datos es realizada por el *Modelo*, las *Vistas* contienen el código que se manda al navegador, generalmente código HTML del lado del servidor, y que se ocupa de las tareas de presentación. Por su parte los *controladores* reciben las peticiones del navegador y en base a éstas, deciden qué vista mostrar y con qué datos.

Este modelo hace también que la realización de pruebas para este tipo de aplicaciones sea más fácil, debido a esta separación de la lógica de la aplicación, pudiendo centrarse en componentes más específicos.

Este framework como conjunto de los 2 elementos anteriormente descritos, surge como una alternativa a ASP.NET, aunque esté construido sobre éste, pero en el que en vez de construir las páginas web mediante el uso de *Webforms*, se realiza mediante el uso del modelo MVC

Otras de las características que proporciona este marco de trabajo son las siguientes:

- **Conectividad y personalización:** Los componentes que se utilizan en este marco, están diseñados para que se puedan reemplazar y personalizar con facilidad.
- **Enrutamiento:** Amplia compatibilidad con el enrutamiento de ASP.NET, este componente permite crear aplicaciones web que contienen direcciones URL comprensibles y que admiten búsquedas y que no tienen que incluir las extensiones de los nombres de archivo.
- **Compatibilidad:** Las características ya existentes de ASP.NET se pueden utilizar en este framework, tales como diferentes tipos de autenticación, uso de roles, autorizaciones para direcciones URL, etc.

- **Entorno de trabajo:** Como principal entorno de desarrollo se ha usado Visual Studio 2015. El Sistema Operativo empleado ha sido Windows 10 (64 bits).
- **Elaboración documental:** Para la elaboración de la documentación se ha utilizado LaTeX mediante el uso de la herramienta TeXstudio¹⁰.
- **Testeado de aplicación:** Para realizar las pruebas se ha planteado el uso de de la herramienta para realizar pruebas en aplicaciones web, Selenium¹¹, sin embargo, por diversas circunstancias, finalmente sólo se ha podido ejecutar la aplicación para comprobar su funcionamiento en los exploradores web; Google Chrome¹² y Microsoft Edge¹³.
- **Control de versiones:** Para realizar el control de versiones, se ha empleado GitHub

Los enlaces que contienen el código a cada uno de los componentes de la plataforma se adjuntan a continuación:

- **Aplicación Web:** <https://github.com/EvilMick/AgroEsi>
- **Crawler y extractor:** <https://github.com/EvilMick/CrawlerX>
- **Servicio noticias:** <https://github.com/EvilMick/ActualizaES>
- **Servicio ovino,vacuno y cereales:** <https://github.com/EvilMick/Lonjas>

¹⁰<https://www.texstudio.org/>

¹¹<http://www.seleniumhq.org/>

¹²<https://www.google.es/chrome/browser/desktop/index.html>

¹³<https://www.microsoft.com/es-es/windows/microsoft-edge>

Capítulo 5

Resultados

EL capítulo contiene los resultados que se han obtenido del proceso de desarrollo de este proyecto. Primero se dará al lector una visión general de este proyecto, la planificación realizada para cada una de las versiones así como los objetivos que se pretendían cubrir con cada una de ellas. También se especifica el marco en el que se ha desarrollado y el análisis para evaluar el proceso de implementación.

5.1 Visión General

Como ya se comentó en el capítulo 4, sección 2.2, para ofrecer un resumen en general y poder describir en qué consiste este proyecto, y la planificación empleada, se ha utilizado un *Vision Board*, y un *Road Map*, los cuales se detallan a continuación:

5.1.1 Vision Board

Se pretende crear una herramienta para el análisis, visualización y seguimiento de los precios de diferentes productos agrícolas.

Grupo objetivo

Esta herramienta está dirigida a que cualquier profesional relacionado con los sectores agrario y/o ganadero, pueda tener acceso a ella con el fin de que les sea de utilidad en sus actividades diarias.

Necesidades

Las enfermedades, las plagas, o las condiciones meteorológicas, convierten a estos sectores en negocios de altísimo riesgo. Por este motivo es fundamental que los profesionales dispongan de herramientas para que puedan anticiparse al futuro e intentar minimizar el impacto de cualquier agente externo.

Producto

El objetivo es crear una aplicación web que permita a los usuarios la visualización y el seguimiento de los precios de diferentes productos agrícolas que se encuentran almacenados en una base de datos a la que está conectada, y que además integra servicios de noticias relacionadas con ambos sectores así como un servicio meteorológico.

Beneficios

El valor de esta herramienta reside en el intento de proporcionar a los agricultores y ganaderos el máximo posible de mecanismos que les permitan obtener el mayor beneficio posible a sus negocios. Del mismo modo, cooperativas y otros intermediarios que participan en la cadena de comercio de estos productos, puedan conocer los precios en lonja que alcanzan los productos agrícolas y ganaderos, para poder establecer unos precios lo más competitivos posibles.

5.1.2 Road Map

El *Road Map* del proyecto se describe a continuación:

Fecha: 7 de abril			
Nombre	Objetivo	Características	Métricas
Versión 1. Visualización básica	Implementación de un ciclo completo del proyecto, es decir, adquisición y transformación de algunos datos para su después visualización desde la aplicación web.	- Elaborar una lista de fuentes de datos sobre los precios de los productos agrícolas. - Elaborar una lista de sitios web que ofrezcan un canal RSS. - Exploración de fuentes de datos mediante proceso de crawling. - Transformación de los datos mediante técnicas de webscraping. - Almacenamiento de los datos. - Visualización básica de los datos desde la aplicación web.	-¿Se obtienen los datos de la lista de fuentes acordada? -¿Los datos se transforman correctamente siguiendo el formato más recomendable? -¿Se almacenan los datos?

Fecha: 7 de mayo			
Nombre	Objetivo	Características	Métricas
Versión 1.0.1 : El Torno	Una vez obtenido un resultado positivo de las pruebas realizadas a la versión anterior. Se empieza a trabajar con datos reales procedentes de la lista de fuentes elaborada previamente. Para esta versión se utilizan tanto datos sobre precios como noticias del sector ganadero solamente.	<ul style="list-style-type: none"> - Comprobación de que todos los datos previstos están almacenados sin errores. - Adquisición de noticias. - Adquisición de precios. - Diseño inicial de la aplicación web. 	<ul style="list-style-type: none"> - ¿Están todos los precios y noticias almacenados? - ¿Los datos se pueden visualizar correctamente? - ¿Es válido el diseño de la aplicación web?

Fecha: 7 de junio			
Nombre	Objetivo	Características	Métricas
Versión 1.0.2 : Zona de los montes	El desarrollo se centra en la adquisición de datos referentes a precios y noticias del sector agrícola.	<ul style="list-style-type: none"> - Comprobación que los datos previstos están almacenados sin errores. - Adquisición de precios. - Adquisición de estadísticas. - Adquisición de noticias y etiquetado en base a su contenido. - Adquisición de precios. - Los usuarios se pueden registrar e iniciar sesión. 	<ul style="list-style-type: none"> - ¿Están todos los datos correctamente almacenados? - ¿Un usuario puede registrarse e iniciar sesión en la aplicación web? - ¿Están todas las noticias correctamente etiquetadas?

Fecha: 7 de julio			
Nombre	Objetivo	Características	Métricas
Versión 1.0.3 : Ciudad Real	Finaliza el proceso de adquisición de datos y es posible su visualización en la herramienta, además se añade una funcionalidad para poder consultar la previsión meteorológica en Ciudad Real así como un histórico de datos de diferentes años. Por otro lado, se crean dos servicios encargados de actualizar de manera periódica las noticias y los precios contenidos en la base de datos.	-Ha terminado el proceso de implementación de toda la estructura de la aplicación web. - Es posible consultar la previsión meteorológica así como otras estadísticas. - La aplicación es completamente funcional.	- ¿Ha terminado el proceso de implementación de la aplicación web? -¿Es posible visualizar datos y estadísticas meteorológicas? -¿Se añaden nuevos datos de manera periódica?

5.1.3 Distribución del proyecto

El proyecto lo componen 3 elementos principales cada uno de los cuales ha sido implementado y mejorado según avanzaba el desarrollo del mismo e iban apareciendo nuevos requisitos y objetivos a completar:

- **Crawler:** Implementado con la librería Abot y su extensión AbotX, este componente se ocupa de explorar las diferentes fuentes de datos, ya sean noticias, estadísticas o precios, y obtener los enlaces URL de los recursos que son de interés, para procesarlos y almacenarlos en la base de datos.
- **Extractor:** Compuesto por diferentes librerías para tratar los recursos proporcionados por el crawler, este componente se encarga de extraer los datos dependiendo de su naturaleza, después, son normalizados siguiendo un formato previamente establecido, antes de ser almacenados en la base de datos.
- **Aplicación web:** Aplicación web implementada utilizando el modelos de desarrollo web ASP.net utilizando el patrón de diseño MVC, encargada de mostrar y analizar los datos almacenados en el data lake.

5.2 Arquitectura

Para la construcción de la aplicación web se ha optado por utilizar el *framework* para este tipo de aplicaciones, **ASP.NET MVC**, el cual como su nombre indica, implementa el patrón *modelo-vista-controlador* (MVC). Ésta última característica no es propia de Microsoft, existen otros frameworks populares como *Ruby on Rails*, *Spring* o *Apache Struts*.

Este modelo arquitectónico, separa las aplicaciones en estos tres componentes principales:

- **Modelo:** Este componente de la aplicación se encarga de implementar la lógica del dominio de datos de la aplicación. A menudo los objetos de un modelo son los que se encargan de almacenar y/o recuperar datos almacenados, del estado de un modelo.
- **Vista:** Las vistas son los componentes que se encargan de mostrar la interfaz de la aplicación al usuario y son las que se encargan de la visualización de los datos de los diferentes modelos.
- **Controlador:** Se encargan de gestionar la interacción del usuario con la aplicación trabajando con el modelo, para después seleccionar una vista que represente la interfaz de usuario. En este tipo de arquitectura las vistas sólo muestran información, es el controlador el que administra y responde a las interacciones por parte del usuario.

El concepto clave de este framework, es el de **enrutamiento**, el cual permite a las aplicaciones web aceptar peticiones URL, de forma que éstas son gestionadas por el controlador y la vista adecuada dependiendo de la URL.

Otro de los elementos clave de este framework, es el motor de vistas *Razor*. En ASP.NET, las vistas son código HTML en las que se puede incrustar fragmentos de código que se ejecuta en el lado del servidor y que sirven para mostrar más información. El código de *Razor* se ejecuta en el servidor antes de que la página se envíe al explorador.

Esta sintaxis proporciona una alternativa a los formularios web de ASP.NET, los cuales se centran en controles de servidor web que generan código HTML de manera automática. En cambio utilizando *Razor*, las páginas funcionan más directamente como paginas HTML comunes, donde el desarrollador puede de manera virtual crear todo el formato.

Otra de las utilidades es *Nuget*, un sistema de administración de paquetes gratuito y de código abierto, que permite a los desarrolladores instalar de una manera más fácil y sencilla bibliotecas de terceros en los proyectos de desarrollo de aplicaciones en .NET.

A continuación se puede comprobar la secuencia de ejecución¹ que sigue cualquier aplicación utilizando las tecnologías ya descritas:

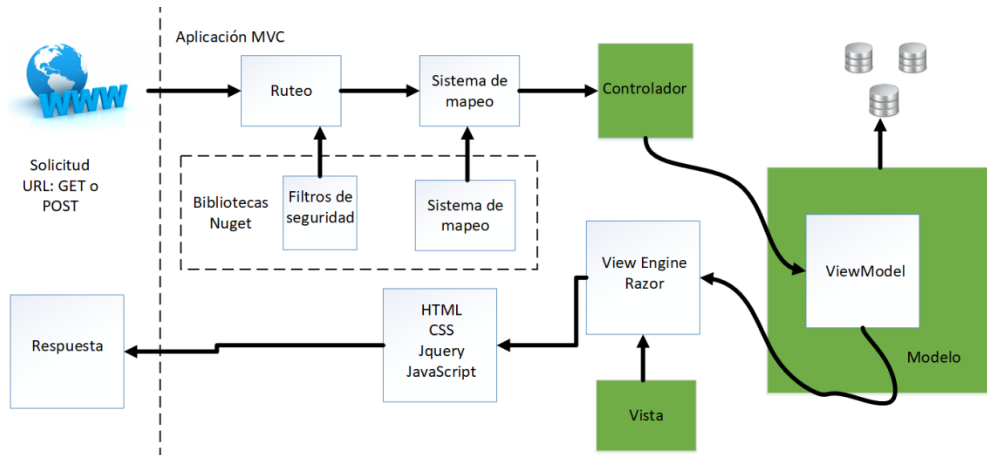


Figura 5.1: Esquema aplicaciones ASP.NET MVC

Lo que se produce es que cuando una aplicación MVC comienza su ejecución, se llama al método *Application_start()*, éste a su vez llama a *RegisterRoutes()*, el cual se encarga de crear la tabla de rutas. Por defecto, esta tabla contiene una única ruta *Default*, de modo que cuando un usuario realiza una acción y se genera la correspondiente dirección URL, esta ruta se encarga de buscar el nombre del controlador utilizando el primer segmento de la misma, el segundo segmento de la ruta se corresponde con la acción del controlador que se quiere invocar, y el tercer segmento se corresponde con los parámetros que recibe la acción invocada.

Con el objetivo de aclarar aún mas al lector, de la secuencia de eventos que se producen en una aplicación construida sobre este modelo arquitectónico, y qué entidades son las que intervienen en cada momento, se adjunta un diagrama de secuencia:

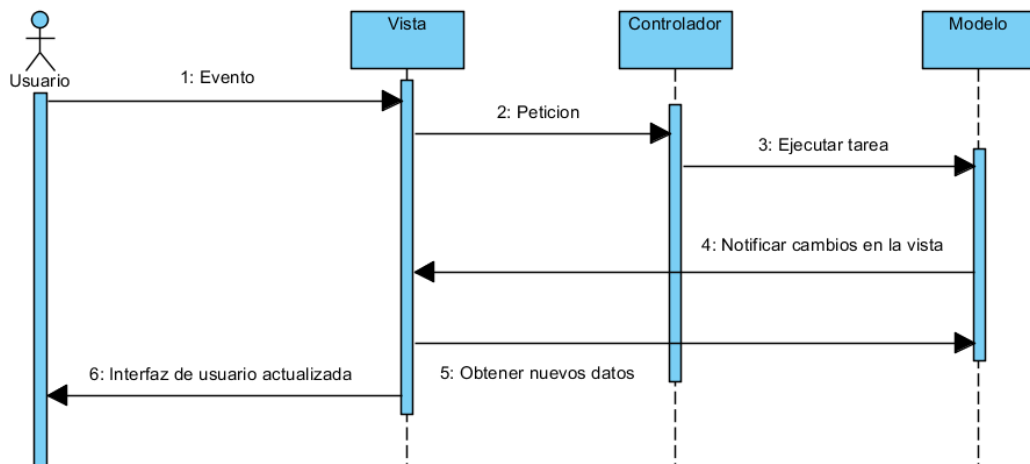


Figura 5.2: Diagrama de secuencia MVC

¹<http://msaspnetmvc.blogspot.com.es/2015/03/la-arquitectura-de-mis-proyectos-mvc.html>

En el caso de la herramienta que se está implementando, la sucesión de eventos que se produce, es la misma que la descrita en la figura anterior, al ejecutar cualquiera de los siguientes módulos funcionales:

- **Noticias:** El usuario selecciona del menú que le aparece si quiere ver las noticias del sector agrícola o del sector ganadero, una vez seleccionada la opción, el controlador correspondiente se encarga de buscar en la base de datos una lista de noticias lo más actuales posibles, y con ayuda del modelo *Noticia*, se manda esta lista a la vista que se encarga de mostrárselas al usuario. Ocurre lo mismo si el usuario decide realizar alguna búsqueda, solo que en este caso, la vista proporciona los parámetros de búsqueda al controlador, el resto de eventos que se producen, no cambian.
- **Precios:** El escenario, es el mismo que el caso anterior, el usuario, selecciona del menú si desea ver los precios del sector agrícola o del sector ganadero, el controlador correspondiente se ocupa de cargar los últimos precios de la base de datos y con la ayuda del modelo *Precio*, crea la lista de precios que recibe la vista, la cual se encarga de mostrar los gráficos, los indicadores y las tablas con las relaciones de datos sobre los precios seleccionados, también, en el caso concreto de la ganadería, el controlador carga de la base de datos, una lista de titulares de noticias relacionadas con este sector y que con ayuda del modelo *Noticia* crea la lista que también recibe y que gracias al motor de vistas **Razor** se ha podido incrustar código C# para manejar estas listas y poder generar el código necesario para su visualización.

En el caso de las búsquedas sobre datos de precios, dependiendo del producto, el usuario introduce; producto, variedad y campaña, en la vista que esté utilizando, y la vista se encarga de recoger los datos introducidos por el usuario, para que el controlador correspondiente se encargue de realizar la búsqueda y utilizar los modelos necesarios para devolver los resultados a la vista y el usuario pueda ver los resultados producidos por la base de datos.

- **Clima:** El escenario vuelve a ser el mismo, el usuario selecciona del menú la opción clima, el controlador correspondiente se encarga de cargar los últimos datos climáticos almacenados correspondientes al último año almacenado en la base de datos con la ayuda del modelo *Clima*, por otro lado se ejecuta el código correspondiente al *widget* que se encuentra incrustado en la vista, para cargar la previsión meteorológica de los próximos días, junto con la lista de años disponibles para que el usuario pueda realizar consultas a la base de datos y poder ver en los gráficos disponibles, las variaciones de precipitaciones y temperatura a lo largo de el año, así como otras estadísticas almacenadas.

A continuación se muestran los modelos, vistas y controladores que se han creado para cubrir las diferentes funcionalidades de la aplicación:

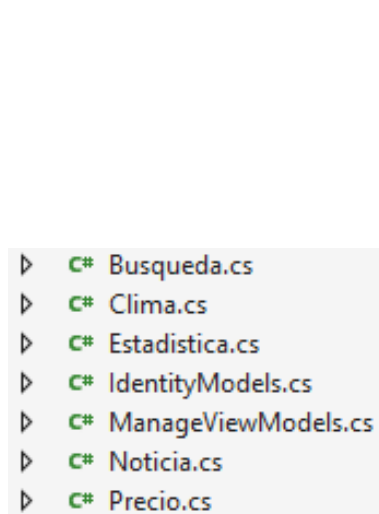


Figura 5.3: Lista de modelos

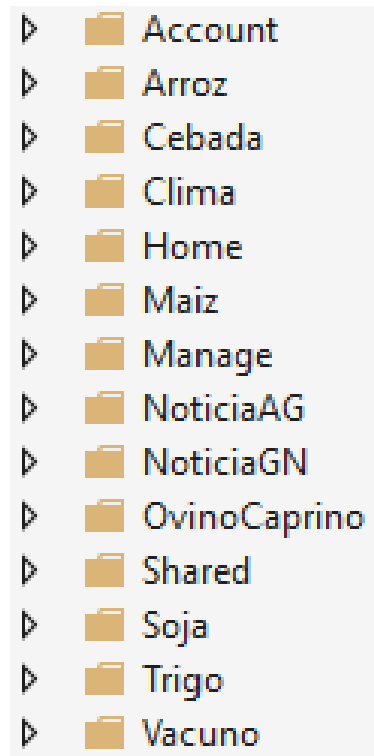


Figura 5.4: Lista de vistas

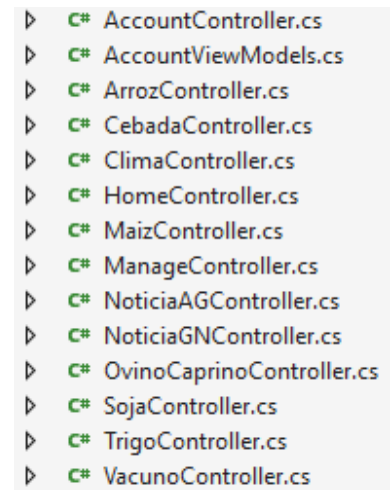


Figura 5.5: Lista de controladores

Ahora se detalla el trabajo realizado para cada una de las versiones, los requisitos de usuario (historias de usuario) que se mencionan en cada una de las mismas, se encuentran disponibles en el anexo D, donde se describen más en detalle.

5.3 Versión 1: Visualización básica

Durante el desarrollo de esta versión se realizan todas las tareas necesarias para obtener algunos datos referentes a precios, de la lista de fuentes de datos, mediante el uso del crawler. Después de los recursos obtenidos, realizar la posterior extracción y normalización de los datos antes de guardarlos en la base de datos, para después mostrarlos en la aplicación web.

5.3.1 Contexto de la versión

Gracias al uso de los ya mencionados “datos abiertos”, cada vez son más los sitios web que ponen a disposición de los usuarios sus propios datos y estadísticas, o también sitios web que recopilan conjuntos de datos ya sea a nivel regional², o nivel nacional³.

En nuestro caso, las lonjas actualizan sus páginas web, de manera periódica, con los precios que alcanzan las diferentes materias primas de las que se ocupan. Debido a la cantidad de fuentes de datos así como de los propios conjuntos de datos a tratar, era primordial, elaborar una lista de fuentes de datos a la que ceñirse a la hora de trabajar. Esta lista ha incluido, fuentes como las páginas web de lonjas como la de Albacete o Talavera de la Reina, a nivel local, o la página web del ministerio de Agricultura, Pesca, Alimentación y Medioambiente, a nivel nacional. También se han podido obtener precios a través de la lonja de Chicago para algunos productos más concretos y sobre los que apenas había datos.

En este punto del desarrollo, donde comienzan las tareas de extracción de datos, procedentes de las distintas fuentes, hay que plantearse hasta qué punto es legal este proceso, el cual, es realizado sin el consentimiento ni notificación a los administradores o responsables de las páginas web. De hecho, existen casos de juicios entre compañías, con motivo del uso de este procedimiento⁴. Una de las principales razones alegadas durante los juicios, por parte de las compañías demandantes, era que esta técnica, podía permitir a la otra parte, ofrecer el mismo servicio proporcionado por la primera. Puede resultar sencillo, pensar que este escenario puede suceder en el caso de esta herramienta, hay que aclarar que en ningún momento ése, haya sido el motivo principal para su desarrollo, al contrario, el que la aplicación se pueda mantener en funcionamiento en el futuro, depende precisamente de la labor que las páginas web de estas organizaciones llevan a cabo semanalmente. Añadir también que el límite de la legalidad de esta técnica, en opinión de el autor del presente documento, lo ponen los datos, es decir, dejando a un lado sean cual sean las intenciones con la extracción de datos mediante esta técnica, no se puede tener la misma consideración con datos referentes a personas físicas que con otros datos.

²<http://data-castillalamancha.opendata.arcgis.com/>

³<http://datos.gob.es/>

⁴<https://www.enrique-dans.com/2017/08/web-scraping-legal-ilegal-o-depende.html/amp>

En última instancia, el desarrollo de esta versión ha tratado de ser una respuesta a una “duda tecnológica”, ya que pese a saber los componentes clave del proyecto y el papel que iba a jugar cada uno de ellos, era necesario definir un marco tecnológico sobre el que trabajar.

5.3.2 Desarrollo de la versión

Durante el desarrollo de esta versión se crea la primera versión de los 3 elementos principales que componen el proyecto y que se han mencionado anteriormente.

- La versión(1.0.0) del crawler incluye la instalación y configuración de las librerías necesarias para comprobar su correcto funcionamiento. Éstas son *Abot* y su extensión *AbotX*. El crawler se prueba con enlaces a fuentes de datos provisionales.
- La versión(1.0.0) del extractor, por el momento sólo es capaz de extraer el contenido de ficheros html (“scraping”), por medio de la librería *AngleSharp*.
- Se crea el proyecto de *Visual Studio* utilizando *ASP.net* con versión(1.0.0) que contendrá la aplicación web, y se conecta a ES, utilizando la librería *NEST* para poder empezar a cargar datos de prueba previamente almacenados.

Se instala y configura el servicio de *ElasticSearch* para poder almacenar los datos. También se crea una clase escrita en C# encargada de interactuar con la misma (*Broker*).

Adquisición de datos

La primera toma de contacto con los datos ha sido con los datos procedentes de la página web del ministerio de agricultura del gobierno español⁵, en los que se pueden visualizar el precio medio de diferentes materias primas. Además cada materia prima contenía un enlace a un breve histórico de precios en diferentes campañas, por lo que ha sido la situación ideal para empezar a realizar las primeras pruebas a nuestro *crawler*.

El proceso de adquisición ha sido muy simple, una vez proporcionada la dirección URL que contiene todos los datos, el crawler se encarga de explorar todos los enlaces que ésta, a su vez, contiene, por lo que con un simple filtro en base al contenido de las direcciones obtenidas, ha resultado sencillo obtener las direcciones deseadas. Después, con un simple cliente *HTTP*, disponible en la API de C# y una clase *FileItem* que contiene la dirección origen y la dirección destino, se ha descargado cada recurso de manera local en el equipo utilizado. En este caso en formato *html*, para su posterior procesamiento.

⁵<https://goo.gl/z1HaVQ>

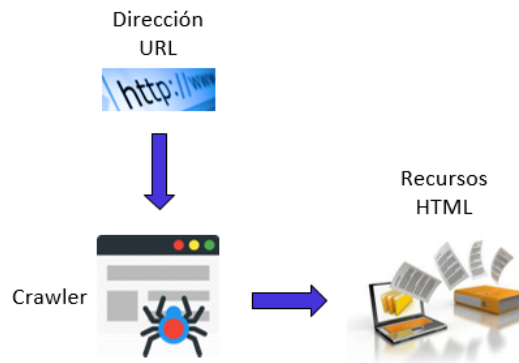


Figura 5.6: Adquisición de datos

Procesamiento de datos

Una vez se tienen los recursos disponibles en el equipo, es el turno de extraer los datos, el *extractor* de datos, lo van a conformar todas las librerías y clases implementadas por el alumno, encargadas de la extracción de los datos sea cual sea su formato. En el caso de esta versión, los recursos sólo se encuentran en formato .html, por lo que para extraer datos de este tipo de formato, se ha usado, *AngleSharp*⁶. Por otro lado, y utilizando una notación orientada a objetos, se ha implementado una clase *Precio* que contiene información como el precio, fecha, campaña, o el nombre del producto entre otros.

Almacenamiento de datos

Para almacenar los datos, y debido a que los datos proceden de distintas fuentes y en distintos formatos, se ha optado por utilizar un esquema NoSQL, ya que era innecesario utilizar un modelo relacional, porque no existen relaciones entre las entidades que se van a almacenar. Por otro lado ES, permite guardar objetos enteros como documentos en formato JSON, en los que cada uno de los campos puede ser accedido. El mapeo realizado durante esta versión se encuentra en el apartado C.1 del capítulo “Mapeos de Elasticsearch”, localizado en anexos.

Durante las pruebas de búsquedas desde la aplicación web, el alumno se ha dado cuenta de que ES, al servir principalmente como motor de búsquedas, toda cadena que almacena, al realizar una búsqueda, el analizador interno, realiza las búsquedas a nivel de palabras, por lo que toda cadena que incluya caracteres tales como espacios, guiones, comas, puntos, etc. las separa utilizando este tipo de caracteres, es por este motivo por el que se ha optado por codificar con un identificador único, todo producto y su respectiva variedad; quedando de la siguiente manera:

⁶<https://anglesharp.github.io/>

Descripción	Código	Descripción	Código
Leche de cabra	LC	Leche de oveja con D.O.	LOCDO
Leche de oveja sin D.O.	LOSDO	Cabrito Basto de 7 a 10kg	C7B10
Cabrito Fino de 7 a 9kg	C7F9	Cordero con IGP de 10 a 15kg	CM10CI15
Cordero con IGP de 15 a 19kg	CM15CI19	Cordero con IGP de 19 a 23kg	CM19CI23
Cordero con IGP de 23 a 25kg	CM23CI25	Cordero con IGP de 25 a 28kg	CM25CI28
Cordero con IGP de 28 a 34kg	CM28CI34	Cordero con IGP media 10kg	CMCIM10
Cordero sin IGP de 10 a 15kg	CSII10YO15	Cordero sin IGP de 15 a 19kg	CSI15YO19
Cordero sin IGP de 19 a 23kg	CSI19YO23	Cordero sin IGP de 23 a 25kg	CSI23YO25
Cordero sin IGP de 25 a 28kg	CSI25YO28	Cordero sin IGP de 28 a 34kg	CSI28YO34
Cordero sin IGP media 10kg	CSIYOM10	Ternero cruzado 1 ^a	TO200C1
Ternera cruzada 1 ^a	TA200C1	Ternero cruzado 2 ^a	TO200C2
Ternera cruzada 2 ^a	TA200C2	Ternero del país	TO200PA
Ternera del país	TA200PA	Maíz	MZ
Trigo Blando	TB	Trigo Duro	TD
Cebada	CBD	Soja	SJ
Arroz Índica	AZI	Arroz Japónica	AZJ

Cuadro 5.1: Relación de productos y códigos

Visualización de datos

Para la visualización de los datos, y como ya se ha mencionado anteriormente se ha creado una aplicación web utilizando ASP.net que se va a encargar de conectarse al servicio de ES y que dispone de distintos tipos de vistas, dependiendo del tipo de dato que solicite el usuario en cada momento. Para esta versión la aplicación no dispone de vistas muy elaboradas ni tampoco gráficos. Por otro lado, se instala la librería necesaria para poder establecer la comunicación con la base de datos, *NEST*⁷ y se configura para su uso. Un vez establecida ésta y comprobado que se pueden obtener y visualizar datos, se da por concluido el trabajo.

⁷<https://github.com/elastic/elasticsearch-net>

5.4 Versión 1.0.1: El Torno.

Con el conocimiento adquirido durante el desarrollo de la versión anterior y teniendo ya claro el “cómo” a la hora de adquirir, procesar, almacenar y visualizar los datos. Se decide por empezar con el tratamiento de datos referentes a precios así como de noticias del sector ganadero.

5.4.1 Contexto de la versión

El desarrollo de esta versión y una vez realizada la lista de fuentes de datos, se va a centrar en la adquisición, procesamiento, almacenamiento y visualización tanto de precios como de noticias relacionadas solamente con el sector ganadero, por el momento.

5.4.2 Desarrollo de la versión

Durante el desarrollo de esta versión se actualizan o mejoran los 3 elementos clave del proyecto de acuerdo con el objetivo principal de desarrollo.

- El crawler durante esta versión no es modificado por lo que se mantiene en la versión(1.0.0). Esto es debido a que gracias a los controles disponibles en el cuadro de mandos de la página web de la lonja de Albacete, fue posible generar un único fichero HTML que contenía el histórico de precios desde el año 1999, hasta la última fecha de la que había datos en ese momento. En cuanto a los precios referentes al sector vacuno, también fue posible generar un histórico con suficientes datos con un número de archivos reducido.
- El extractor, ha sido modificado y se han añadido nuevas librerías para la extracción de datos debido a que han aparecido nuevos formatos a tratar, en concreto ficheros *.PDF*, por lo que se ha añadido; *itextsharp*. La versión pasa a ser (1.0.1).
- En cuanto a la aplicación web, también es modificada su estructura y diseño, así como las funcionalidades que ésta es capaz de realizar. La versión pasa a ser (1.0.1).

Adquisición de datos

Los históricos de datos con los que se ha decidido trabajar han sido los que se encuentran disponibles en las páginas web de las lonjas de Albacete y de Talavera de la Reina, de las siguientes materias primas; cordero, cabrito, leche de oveja y de cabra, con y sin denominación de origen y ternera. En el caso de la lonja de Albacete, figura 5.7⁸, se han podido obtener datos desde 1999 hasta la actualidad, en el caso de la lonja de Talavera de la Reina, figura 5.8⁹, para precios pertenecientes al sector ovino/caprino, se han almacenado datos del sector vacuno del periodo 2002-2016.

⁸<https://goo.gl/VXqxPg>

⁹<https://goo.gl/86Bm6r>

Histórico de precios

Mesa: Categoría: Producto:

Fecha inicial: Fecha final:

Datos del día 17/08/2017. Semana 33

Mesa	Categoría	Producto	Mínimo	Máximo	Medida
CEREALES	AVENA	Avena Blanca		141,00	precio/t
CEREALES	AVENA	Avena Rubia		141,00	precio/t
CEREALES	CEBADA	Para pienso de más de 62 kg/hl		156,00	precio/t
CEREALES	CEBADA	Para pienso de menos de 62 kg/hl		149,00	precio/t

Figura 5.7: Lonja de Albacete

Figura 5.8: Lonja de Talavera

En el apartado de las noticias, y por el momento, se ha almacenado el histórico desde 2009 de noticias referentes al sector ovino/caprino, procedente del portal de noticias “OVIespaña”, figura 5.9¹⁰. Para esta labor se ha llevado a cabo una descarga automática de las páginas con las noticias en ficheros formato html.

¹⁰<https://goo.gl/LzjKyB>



Figura 5.9: Portal de noticias “OVIESPAÑA”

Procesamiento de datos

Una vez descargados todos los ficheros, se han implementado nuevas clases que han pasado a formar parte del conjunto de las mismas que conforman el extractor, utilizando la nueva librería para el tratamiento de ficheros PDF, y se ha usado la librería ya instalada durante el desarrollo de la versión para el tratamiento de datos en formato HTML.

Por otro lado se ha implementado otra clase, denominada “Noticia”, para crear instancias de ésta con todos los datos relevantes sobre las noticias extraídas.

Almacenamiento de datos El proceso para crear el nuevo tipo de datos a almacenar, en este caso noticias, se describe en el apartado C.2 del capítulo “Mapeos de ElasticSearch”, localizado en anexos.

Visualización de datos

Una vez almacenados los datos, es necesario implementar las vistas necesarias para la correcta visualización tanto de precios como de noticias. En el caso de los precios, durante el desarrollo de esta versión, también se instala “Chart.MVC¹¹”, para que la aplicación web pueda generar gráficos a partir de los datos y el usuario pueda ver una manera más clara las variaciones de los precios.

Por otro lado, también se añade en el lateral derecho del cuadro de mandos creado una pequeña columna con los titulares más recientes que hay de este sector almacenados en la base de datos, con un enlace a la fuente original de la misma, figura 5.10.

¹¹<https://github.com/martinobordin/Chart.Mvc>

Evolución de precios:

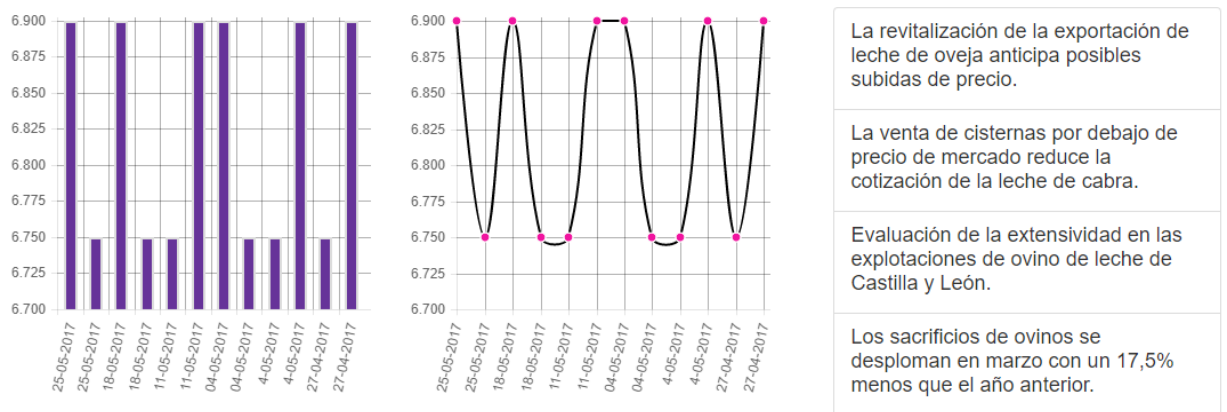


Figura 5.10: Gráficos y columna de noticias.

También se añade una tabla con la relación de datos sobre los precios, figura 5.12, así como una serie de indicadores, figura 5.11, para indicar al usuario valores como el número de campañas y cuáles, de las que existen datos, el precio más reciente, y el máximo y mínimo históricos.

Cuadro de mandos:



Figura 5.11: Indicadores de precios.

Producto	Variiedad	Precio	Tipo precio	Medida	Fecha
Leche	de cabra	6,9	max	precio/Hgdo	25-05-2017
Leche	de cabra	6,75	min	precio/Hgdo	25-05-2017
Leche	de cabra	6,9	max	precio/Hgdo	18-05-2017
Leche	de cabra	6,75	min	precio/Hgdo	18-05-2017
Leche	de cabra	6,75	min	precio/Hgdo	11-05-2017
Leche	de cabra	6,9	max	precio/Hgdo	11-05-2017

Figura 5.12: Relación de datos de cada precio.

5.5 Versión 1.0.2: Zona de los montes

Se trata de una versión que aunque da comienzo después de la versión (1.0.1), gracias al desarrollo de la versión inicial y la consiguiente respuesta a esa duda tecnológica que se planteaba, podría haberse llevado a cabo en paralelo con la versión anterior, ya que mientras la primera, se ocupa del tratamiento de datos referentes al sector ganadero, ésta se ocupa de los datos, estadísticas y noticias relacionados con el sector agrícola.

5.5.1 Contexto de la versión

El desarrollo de esta versión se centra en la adquisición, procesamiento, almacenamiento y visualización de precios, noticias y estadísticas relacionadas, en este caso con el sector agrícola.

5.5.2 Desarrollo de la versión

Durante el desarrollo de esta versión se actualizan o mejoran, si es necesario, los 3 elementos clave del proyecto de acuerdo con el objetivo principal de desarrollo.

- El crawler durante esta versión no es modificado, y por el momento sigue en la versión(1.0.0). Como en el caso anterior, no es necesario su uso a la hora de adquirir y manejar los recursos necesarios para esta versión ya que con apenas varios ficheros se ha podido crear el histórico de datos para los cultivos que se manejan en la aplicación, en caso de la soja, al igual que pasó con los datos de precios de ovino, con los controles presentes en la página web de la fuente elegida, fue posible generar un único fichero con los datos, por lo que se ha considerado innecesario su uso.
- El extractor, durante el desarrollo de la versión, ha sido modificado, debido en gran medida a que aparecen recursos en formato *Excel*, por lo que es necesario el uso de una librería nueva, *ExcelDataReader* para el manejo de datos con este formato, además también es necesario la implementación de nuevas funciones que se encarguen de extraer los datos de nuevos recursos en formato PDF y en formato HTML, con las librerías instaladas en anteriores versiones. La versión pasa a ser (1.0.2).
- En cuanto a la aplicación web, también se modifica su estructura, se añaden nuevas vistas para mostrar los nuevos datos referentes al sector agrícola y diseño, así como las funcionalidades que ésta es capaz de realizar. La versión pasa a ser (1.0.2).

Adquisición de datos

Debido a la gran cantidad de productos agrícolas y hortofrutícolas, se ha optado por elegir algunos de los productos más representativos y que más importancia tienen en el comercio exterior, como son; maíz, trigo, arroz, soja y cebada. Tanto los informes con los históricos de precios, como estadísticas sobre superficies cultivadas, producción, importaciones y exportaciones, se han podido obtener de la página web del ministerio de agricultura, el cual además de mantener estos históricos, es el que publica los precios semanales que alcanzan estos productos en diferentes lonjas de la geografía española.

Los datos obtenidos son los siguientes:

Producto	Campañas (<i>Precio</i>)	Campañas (<i>Estadística</i>)
Maíz	2011/12-2015/16	2014/15-2016/17
Cebada	2011/12-2015/16	2014/15-2016/17
Trigo	2011/12-2015/16	2014/15-2016/17
Arroz	2010/11-2015/16	2010-2016
Soja	2007-2017	2010-2015

Cuadro 5.2: Datos cereales

En el apartado de las noticias, se ha podido obtener un histórico desde el año 2013 hasta la fecha, de 2 portales distintos; *Agroinfo*¹² y *efeagro*¹³. Esta adquisición de datos también ha sido realizada de manera automática, al igual que se hizo en el versión (1.0.1).

Procesamiento de datos

Una vez descargados todos los ficheros, se han implementado nuevas clases que han pasado a formar parte del conjunto de las mismas que conforman el extractor, utilizando la nueva librería para el tratamiento de ficheros Excel, además se ha implementado nuevas funciones para el tratamiento de los datos en formato PDF, debido a que éstos estaban estructurados de manera distinta a otros previamente tratados. En el caso concreto de la soja, los datos han sido extraídos de un fichero html utilizando la librería instalada durante la versión (1.0.0).

Durante el desarrollo de ésta versión surge la necesidad de hacer un etiquetado automático de noticias en base a su contenido. Pese a que muchas de las noticias procedentes de sus respectivas fuentes, venían etiquetadas, ha sido necesaria la implementación de un procedimiento encargado de realizar esta labor, ya que las noticias de años anteriores no lo estaban.

Este procedimiento ha consistido en determinar si un conjunto de palabras clave *keywords*, se encontraban presentes en el cuerpo de la noticia, de modo que si alguna de éstas, estaba presente en el contenido de la noticia, servía como etiqueta para la misma.

¹²<http://www.agroinformacion.com/category/agricultura/>

¹³<http://www.efeagro.com/category/agricultura/>

Para los datos referentes a las estadísticas, se ha creado una nueva clase “Estadística”, sobre la que crear diferentes instancias para cada producto, con atributos como, superficie, producción, importaciones y exportaciones, así como el código de cada producto.

El caso del procesamiento de noticias, además de lo ya comentado del etiquetado automático en base a su contenido, se ha seguido el mismo procedimiento que en el caso de las noticias relacionadas con el sector ganadero.

Almacenamiento de datos

Dado que aparece un nuevo tipo de datos que se quiere almacenar también en la base de datos, para que la herramienta ofrezca más valor a su uso, se ha decidido almacenar datos de producciones, o importaciones/exportaciones relacionados con los cultivos elegidos, bajo un nuevo tipo de datos *Estadística*. El correspondiente mapeo se encuentra en el apartado C.3 del capítulo “Mapeos de ElasticSearch”, localizado en anexos.

Visualización de datos

Una vez almacenados los nuevos datos, se han implementado las correspondientes vistas tanto para la visualización de precios como de noticias, cuyo menú ahora, las separa en 2 vistas diferentes *Agricultura* y *Ganadería*, además ahora cada noticia aparece etiquetada y se puede ver la fecha de publicación.



Figura 5.13: Ejemplo de *Noticia*.



Figura 5.14: Menú noticias.

5.6 Versión 1.0.3: Ciudad Real

Se trata de la última versión realizada durante el desarrollo de este proyecto. Para completar las funcionalidades que la herramienta es capaz, se añade una última para que los agricultores y ganaderos puedan consultar la previsión meteorológica. Por otro lado y no menos importante, se intenta proveer de mecanismos a la plataforma, que sean capaces de actualizar los datos contenidos en la base de datos, tanto precios, como noticias.

5.6.1 Contexto de la versión

El desarrollo de esta versión, se centra en la adquisición, procesamiento, almacenamiento y visualización, de datos climáticos en Ciudad Real, para su posterior visualización en la aplicación web.

Como ya se ha comentado en otras secciones de este documento y según la opinión del autor de éste, el factor climático es quizá el que más influye a la hora de obtener buenas cosechas por parte de los agricultores por lo que una buena previsión, accesible a cualquier agricultor, así como el análisis del comportamiento de el clima en una zona, es fundamental para tratar de que éste sea la más previsible posible y así los agricultores puedan planificar las siembras y cosechas.

En segundo lugar, el hecho de que en versiones anteriores los datos hayan sido descargados de forma manual, sobre todo, para crear un histórico de datos de precios y noticias, no implica que en el futuro tenga que haber una persona encargada de la actualización de los datos. Por este motivo, se ha optado por la implementación de dos servicios para Windows, los cuales corren en segundo plano y se encargan de actualizar de actualizar la mayor parte de los precios y noticias almacenados, por sí solos.

5.6.2 Desarrollo de la versión

Durante el desarrollo de esta versión se han actualizado o mejorado, en caso necesario, los 3 elementos clave del proyecto:

- El crawler durante esta versión ha sido modificado y utilizado, para la obtención de todos los datos climáticos como ya se explicará en siguientes apartados, por lo que la versión pasa a ser la (1.0.1).
- El extractor, durante el desarrollo de la versión, ha sido modificado, para añadir nuevas funciones encargadas de extraer los datos procedentes de archivos .html, mediante scrapping. La versión pasa a ser (1.0.3).
- La aplicación web, también se modifica su estructura, se añaden nuevas vistas para mostrar los nuevos datos referentes al clima, y continúan las tareas de mejora en la visualización de los datos, así como a la hora de seleccionar datos concretos contenidos en la base de datos. La versión pasa a ser (1.0.3).

Como ya se ha mencionado con anterioridad, también se implementan dos servicios encargados de actualizar los precios y las noticias; uno de ellos que se ejecuta todos los días para la actualización de noticias, tanto las referentes al sector ganadero como al agrícola, el segundo, es suficiente con que se ejecute una vez por semana, ya que éste se encarga de actualizar los precios en lonja, los cuales se suelen publicar cada miércoles/jueves. Al tratarse de unas versiones iniciales la versión es la (1.0.0). Se explicarán más en detalle en el anexo E, junto con el nuevo esquema de funcionamiento de la plataforma.

Adquisición de datos

Los datos climáticos¹⁴ que finalmente han sido almacenados en la base de datos, debido a la cantidad, y a su organización, ha sido necesario el uso del crawler para extraer todos los enlaces de datos, que comprendían desde el año 1973, hasta la actualidad, además por cada año, se han extraído diferentes valores medios que se han considerado relevantes, de cada uno de los meses del año.

En el caso de las fuentes de datos a las que se conectará el servicio; para las noticias, se emplearán los canales RSS que ofrecen las fuentes ya utilizadas para crear el histórico de noticias tanto del sector ganadero¹⁵, como del sector agrícola^{16 17}.

En el caso de los precios, por el momento solo se actualizan los de la soja¹⁸, que se actualizan todos los días, para el resto de materias primas, por el momento no se han encontrado fuentes con datos lo suficientemente actualizados.

Procesamiento de datos

Una vez obtenidos todos los archivos .html que contienen los datos climáticos, es el turno de extraer los datos, gracias al conocimiento adquirido en el desarrollo de las otras versiones, solo ha sido necesaria la adaptación del código, para poder utilizar *AngleSharp*, y así poder extraer los datos.

En el caso de las noticias, una vez construido el objeto .XML, a partir de el enlace de los diferentes canales RSS de noticias, se han podido construir los diferentes objetos de tipo *Noticia* y *Precio*, que son los que se encargan de encapsular los datos y son los que se guardan en la base de datos.

Para almacenar los datos climáticos se ha creado una nueva clase *Clima*, encargada de encapsular los valores medios de datos tales como; temperatura media, temperaturas máximas y mínimas, humedad, presión atmosférica, cantidad de precipitaciones o velocidad del viento.

¹⁴<https://www.tutiempo.net/clima/ws-83480.html>

¹⁵<https://goo.gl/TeJitm>

¹⁶<http://www.agroinformacion.com/feed/>

¹⁷<https://goo.gl/CDjzXJ>

¹⁸<http://www.ambito.com/economia/mercados/granos/info/?id=Soja>

Almacenamiento de datos

Para el almacenamiento de los datos climáticos, al ser un nuevo tipo de datos, es necesario definirlo en ES, así como el nuevo índice que va a contener estos datos. El mapeo correspondiente se encuentra en el apartado C.4 del capítulo “Mapeos de ElasticSearch”, localizado en anexos.

Visualización de datos

Una vez almacenados los datos, se ha implementado una nueva vista para la visualización de los datos, bajo su correspondiente pestaña en el menú principal de la aplicación web. Después de seleccionar el año, se muestran 2 gráficos con las variaciones de temperaturas y precipitaciones a lo largo del año, junto con una tabla con la relación de los valores medios de los diferentes factores climáticos comentados anteriormente. También se ha incluido un *widget* que muestra el pronóstico para un intervalo de cuatro días.

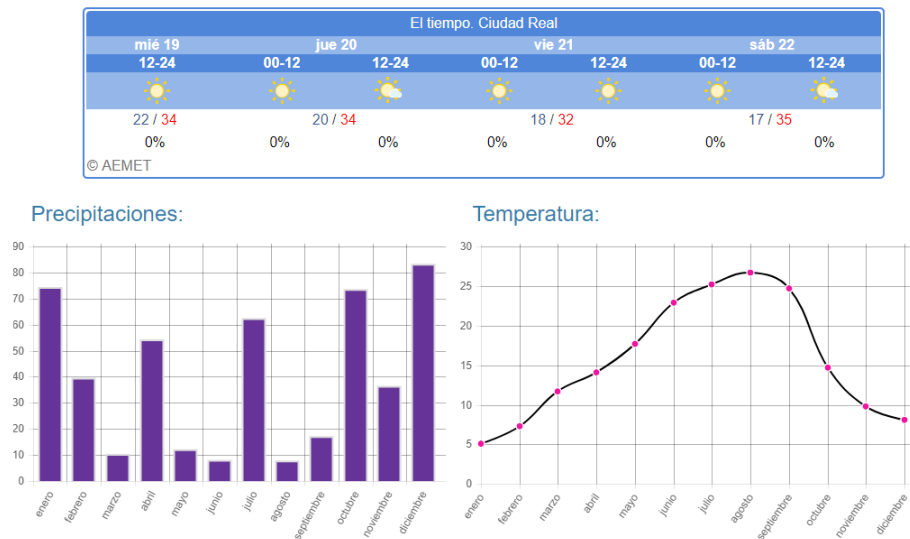


Figura 5.15: Gráficos

Capítulo 6

Conclusiones

EN este capítulo se pretende detallar el grado de cumplimiento conseguido con el trabajo realizado hasta la fecha y los objetivos que se marcaron al comienzo.

Como ya se ha detallado a lo largo de este documento, el objetivo principal ha sido el de implementar una plataforma web para el análisis, visualización y seguimiento de los precios de diferentes productos agrícolas. En la actualidad, los agricultores y los ganaderos no disponen de demasiadas herramientas que unifiquen en un sólo lugar, los precios de las materias primas en las que puedan estar interesados, por lo que si así lo desean, tienen que dirigirse a las lonjas o bien llamando o bien de manera online, para ver cómo se están comportando los mercados.

AgroEsi, no sólo permite ver los precios de diferentes productos según se actualizan de manera periódica en las lonjas, esta herramienta permite:

- Analizar el comportamiento de un mercado determinado a lo largo de un año, o de un mes dependiendo de la materia en concreto, no sólo mediante la visualización de los precios en diferentes gráficos si no que también se pueden visualizar otras estadísticas almacenadas en la base de datos.
- Servicio con las noticias diarias almacenadas en la base de datos, relacionadas con la agricultura y la ganadería, para que los agricultores puedan estar al día de todo lo que ocurre fuera de sus explotaciones.
- Acceso al historial climático de Ciudad Real, así como un servicio diario con la previsión meteorológica.

Dado que es fundamental para un uso útil de la aplicación, que la información contenida en la base de datos esté actualizada, se han implementado dos servicios; uno que se ejecuta todos los días encargado de recopilar todas las noticias publicadas a lo largo del día, así como el precio de la soja, que también se actualiza cada día. Y otro, encargado de recopilar los precios en las lonjas y que se actualizan cada semana.

A continuación se muestra una tabla con la lista de de objetivos específicos propuestos para este proyecto y el trabajo que ha sido necesario para conseguirlo:

Listado de objetivos	
Objetivo	Trabajo necesario
Proceso de Crawling	Con ayuda de librerías para C#, se ha podido implementar un proceso por el cual es posible explorar una página web y filtrar los que ésta contiene.
Mecanismos de adquisición de datos	Implementación de dos servicios encargados de obtener los datos de las distintas fuentes, dependiendo del producto, así como de la periodicidad con la que se publican datos de los mismos.
Interacción con ES	Implementación de una clase <i>broker</i> utilizando el lenguaje de programación C# que contiene la configuración necesaria para obtener conexión con Elastic-Search, de modo que en cualquier momento que se necesite interacción con la misma, se pueda hacer uso de ésta.
Aplicación web	Implementación de una plataforma web utilizando la tecnología ASP.NET junto con el modelo arquitectónico MVC, para que se puedan visualizar los datos contenidos en la base de datos y resulten de utilidad para los usuarios.

Cuadro 6.1: Objetivos del proyecto

6.1 Opinión personal

Como hijo de ganadero y agricultor, creo que la aplicación puede resultar de utilidad en ciertas ocasiones; por ejemplo cuando se acercan las campañas de verano o navidad, y un ganadero quiere confirmar que efectivamente son los momentos en los que el cordero lechal o el cabrito alcanzan los precios más altos del año, del mismo modo, un agricultor puede consultar el mejor momento para realizar sus cosechas, tanto por precio como por condiciones climáticas. También puede servir, en mi caso, por mi padre, cuando sube el precio del pienso que utiliza y la respuesta que recibe es que una o varias materias primas presentes en el pienso han subido de precio debido a sequías u otros factores, de manera que mediante el uso de esta herramienta pueda comprobar si esa subida de precio está realmente justificada. Sin embargo, creo que tanto para agricultores como para ganaderos, puede acabar siendo simplemente un portal de noticias que ofrece un servicio para consultar la previsión meteorológica, ya que no son ellos los que ponen los precios a las materias primas que venden y por tanto mucha utilidad a esta funcionalidad no van a poder sacarle. Por otro lado, creo que cooperativas o empresas que se dedican a la compra-venta de materias primas sí que podrían sacarle partido a esta funcionalidad, para poder ofrecer unos precios lo más competitivos posibles.

Para terminar, si al principio del documento comentábamos la necesidad de actualizar las infraestructuras para poder aplicar TIC, a estos sectores, creo también en la obligación de un cambio de mentalidad en agricultores y ganaderos, para que empiecen a confiar en las posibilidades que ofrecen.

6.2 Trabajo futuro

Durante los siguientes apartados de esta sección se pretenden describir las mejoras que se le pueden hacer a la propia plataforma web, así como otras funcionalidades o servicios que se pueden añadir.

Aunque la mayor parte de los objetivos, han sido cubiertos, a continuación se detallan una serie de tareas que se pueden realizar para completar aún mas la plataforma y la consecuente mejora en la experiencia de uso de los usuarios.

6.2.1 Aplicación Android

La movilidad es fundamental en negocios de este tipo, por lo que los profesionales no pueden tener que utilizar un portatil con la consiguiente conexión física a internet que ello conlleva, por lo que es mejor tener una versión de la plataforma que pueda correr en dispositivos móviles que tienen acceso a internet casi en cualquier lugar siempre que las redes móviles lo permitan.

Por este motivo se plantea el diseño e implementación de una aplicación para Android, que implemente una versión con funcionalidad reducida de la plataforma junto a un sistema de preferencias de modo que el usuario sólo reciba notificaciones de los precios alcanzados en lonjas así como noticias relacionadas, de acuerdo con sus preferencias.

6.2.2 Análisis Predictivo

Debido al avance de las tecnologías TIC, hoy en día es posible almacenar grandes cantidades de datos de cualquier tipo y origen, pero tener toda esta información almacenada no resulta útil para una una organización si no es capaz de extraer conocimiento de éstos. Es necesario la aplicación de técnicas y el uso de herramientas que sean capaces de detectar patrones o tendencias en los datos [ET17].

En el caso de *AgroEsi* ocurre lo mismo, se tiene una buena cantidad de registros de precios, noticias y datos climáticos, y pese a que es posible realizar consultas y pueden verse la comparativa de precios en el tiempo, esto para algunos profesionales puede no ser suficiente. Para intentar resolver esta situación se plantea para el futuro, la aplicación de técnicas de análisis predictivo, un área de la minería de datos encargada de extraer conocimiento de los datos y uso para poder predecir patrones y tendencias de comportamiento. De modo que se puedan resolver preguntas del estilo; ¿Cuál es el factor que más influye en el precio del cereal? o ¿Cuál es la época del año en la que más caro está el precio del cordero?

6.2.3 Datos de nuevos mercados

Para el desarrollo de este proyecto se ha optado por centrarse en los mercados tanto agrónomos como ganaderos que predominan en la zona de el autor del mismo, sin embargo, en Castilla - La Mancha y en la provincia de Ciudad Real, no sólo se cultivan cereales, también existen zonas donde los cultivos que predominan son las olivas y la vid.

Por este motivo, también se plantea añadir a la base de datos precios y otras estadísticas que resulten de interés para estos cultivos, de manera que más profesionales se puedan beneficiar de el uso de esta herramienta, del mismo modo no hay que olvidarse de seguir actualizando los precios de los mercados ya presentes en la plataforma.

6.2.4 Gestión de usuarios

Aunque el desarrollo se haya centrado en conseguir una plataforma funcional para que los usuarios puedan realizar sus consultas de datos. Dado que se ha incluido un proceso para que los usuarios se puedan registrar e iniciar sesión en la aplicación para poder utilizar todas las funcionalidad que ésta ofrece, es conveniente implementar una vista donde el administrador pueda ver la lista de usuarios que utilizan la aplicación y poder realizar las acciones convenientes según corresponda.

6.2.5 Migración a Elastic Cloud

Durante el desarrollo de la aplicación, la base de datos se ha encontrado instalada de forma local en el equipo utilizado durante todo el proceso de desarrollo, sin embargo, la compañía que desarrolla Elasticsearch, ofrece un servicio de pago¹, que se apoya en Amazon Web Services y que permite acceder a los datos sea cual sea la ubicación de la página web a través de Internet.

¹<https://www.elastic.co/cloud>

ANEXOS

Anexo A

Fuentes de datos

A continuación se muestran las distintas relaciones de fuentes de datos que se han empleado para la adquisición de datos referentes a precios, noticias, estadísticas y clima:

Nombre	Dirección URL	Descripción
Lonja de Talavera	https://goo.gl/86Bm6r	Página web de la lonja de Talavera
Lonja de Albacete	https://goo.gl/VXqxPg	Página web de la lonja de Albacete
Mapama	https://goo.gl/A1uQYf	Página web del gobierno español
Indexmundi	https://goo.gl/Rrresh	Página web del precio diario de la soja
Ambito	https://goo.gl/jg6FY5	Página web del precio diario de la soja

Cuadro A.1: Fuentes de precios y estadísticas.

Nombre	Dirección URL	Descripción
Oviespaña	https://goo.gl/LzjKyB	Noticias de ovino-caprino.
Efeagro	https://goo.gl/3UGbNM	Noticias de agricultura y ganadería.
Agroinformación	https://goo.gl/LUPi3d	Noticias de agricultura y ganadería.

Cuadro A.2: Fuentes de noticias.

Nombre	Dirección URL	Descripción
Tutiempo	https://goo.gl/6aHoiH	Histórico de datos climáticos.
Aemet	https://goo.gl/MZuUHv	Previsión diaria del tiempo.

Cuadro A.3: Fuentes de clima.

Anexo B

Manual de usuario

Durante este capítulo se tratará de explicar a grandes rasgos los requisitos necesarios para poder ejecutar la aplicación, después se realizará una descripción sobre cómo funciona la aplicación, cómo está organizado el contenido que se puede visualizar en la misma y cómo interactuar con la plataforma para poder obtener los resultados que persiga el usuario.

B.1 Requisitos de uso

El principal requisito, es que la ejecución de los diferentes elementos que conforman la plataforma es que sea un entorno Windows (mínimo Windows 7, o versiones superiores), ya que si se quiere tener actualizada la información contenida en la base de datos, los servicios que se han implementado, se instalan junto con los que ya ofrece este sistema operativo.

Otro requisito, mientras se utilice Elasticsearch de manera local, es que también se instala como un servicio para Windows, y para ello es necesario tener como mínimo instalado Java 8 (versión 1.8.0).

Para ejecutar la aplicación, de manera local, es necesario tener instalado el framework para .NET (versión 4.5) que suele venir incluido en Visual Studio.

B.2 Página principal. Sin inicio de sesión

Cuando el usuario se conecta a la plataforma, utilizando cualquier explorador web, se carga la página principal de la plataforma con el menú de la aplicación, el cual sólo muestra algunas opciones, en caso de que el usuario no se haya registrado e iniciado sesión en la aplicación:



Figura B.1: Opciones disponibles sin iniciar sesión

También se muestran los últimos titulares almacenados en la base de datos tanto de agricultura como de ganadería:

AgroEsi, tu herramienta para poder estar al tanto de todo lo que sucede fuera de tu negocio.

Últimos titulares:

Agricultura	Ganadería
Los ecologistas piden la ayuda del Ejército para erradicar la xylella, pero rechazan la tala masiva	Etiopía: los ganaderos afectados por la sequía necesitan ayuda urgente
Vuelve La Verema del Cava, un mes de enoturismo en el Penedès	España encara un otoño crítico del que dependerá la entrada en sequía severa
La finca La Cocosa da nuevo paso para situarse a la vanguardia en tecnología agrícola	La feria italiana "Cheese 2017" estará dedicada a la leche cruda
	El Banco de Inglaterra seguirá emitiendo billetes con grasa animal

Figura B.2: Últimos titulares

B.3 Clima

A partir de aquí un usuario puede dirigirse a la categoría de la sección de noticias que más le interese o bien puede acceder a la pestaña *Clima* para consultar la previsión diaria o las estadísticas climáticas almacenadas en la base de datos; para cada año se muestran las precipitaciones y temperaturas medias mensuales, en dos gráficos distintos así como otras estadísticas relevantes en una tabla:

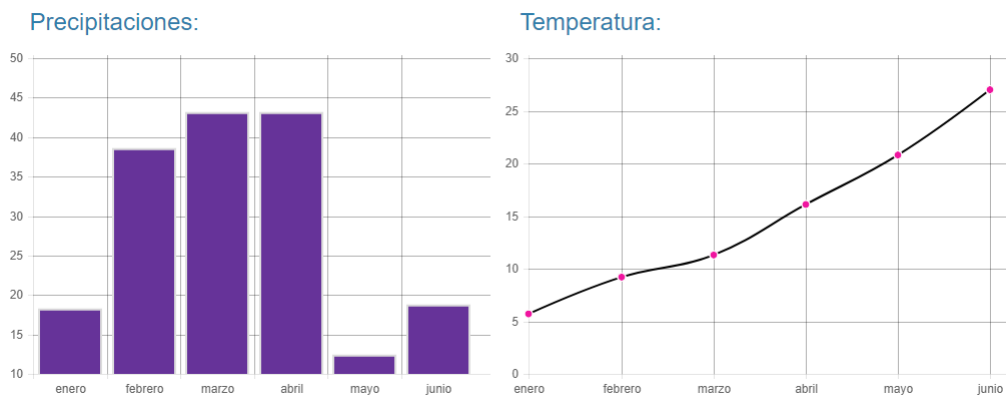


Figura B.3: Precipitaciones y temperaturas medias mensuales

Mes	T(°C)	TM(°C)	Tm(°C)	P.Atm(hPa)	H(%)	PP(mm)	V(Km)	Vm(Km/h)	VM(Km/h)
enero	5,7	10,9	1	1023,7	75,8	18,29	16	6	11,7
febrero	9,2	14	5,2	1020,5	76,5	38,61	16,2	9,4	16
marzo	11,3	17,8	5,5	1020,5	65,8	43,16	18,8	8,3	15,3
abril	16,1	23,2	8,7	1017,4	46,1	43,18	19,1	8,6	15,5
mayo	20,8	27,8	13,1	1014,2	44,5	12,45	19,6	7,7	14,3
junio	27	34,1	19,3	1013,5	32,4	18,79	20	8	15,7

Figura B.4: Otras estadísticas climáticas

B.4 Noticias

Las noticias, se encuentran divididas en las dos categorías que se han tratado a lo largo de este documento; *Agricultura* y *Ganadería*; ambas vistas siguen el mismo esquema; un pequeño cuadro de mandos, para poder seleccionar noticias de diferentes años y meses, así como una opción de búsqueda haciendo clic en cualquiera de las etiquetas de cada noticia, además si se hace clic en la fecha de la noticia también se realiza una búsqueda de noticias de ese mes y año.



Figura B.5: Organización de noticias

Por cada noticia, se carga una miniatura, si está disponible en la fuente original, el título de la misma que contiene el enlace a la fuente, una serie de etiquetas con palabras clave de la misma y la fecha en que fue publicada.

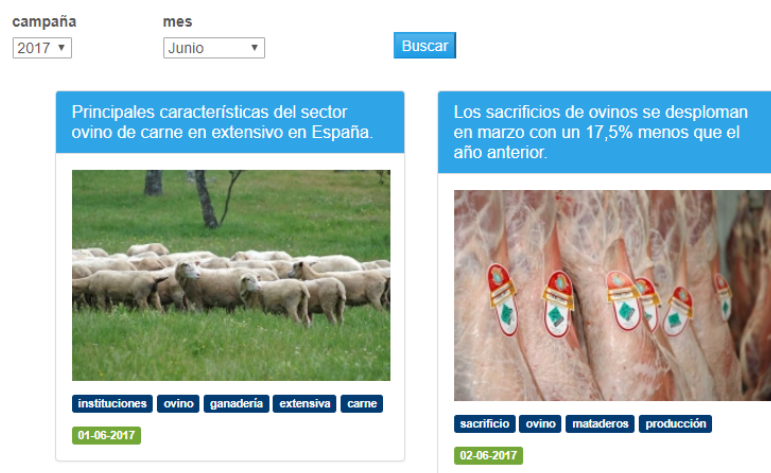


Figura B.6: Ejemplo de búsqueda

B.5 Página principal. Con inicio de sesión

Una vez, un usuario decide registrarse en la plataforma para poder ver los precios de las diferentes materias primas tanto productos ganaderos como agrícolas. Al iniciar sesión, en el menú de la página principal aparecen las correspondientes opciones pasando a ser así:



Figura B.7: Opciones disponibles después iniciar sesión

B.6 Agricultura

En esta sección del menú principal se encuentran las opciones para ver las precios de; *Cebada, Trigo, Maíz, Soja y Arroz*. Todas las vistas comparten el mismo patrón, un pequeño cuadro de mandos, para seleccionar el precio:

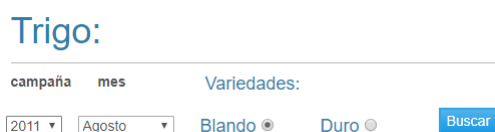


Figura B.8: Cuadro de mandos

Unos indicadores que muestran los precios más alto y más bajo de todos los precios almacenados para ese producto, el número de campañas de las que se disponen precios en la base de datos, así como el precio más reciente almacenado:

Los precios se muestran utilizando gráficos de barras y de líneas, en el caso de el arroz y del trigo, dado que se almacenan precios de las variedades *Japónica e Índica*, para el caso de el arroz y *Blando y Duro*, para el caso de el trigo, cuando se busca el precio para una determinada fecha, también se muestran los precios que tuvo la otra variedad, para ese mismo intervalo de tiempo:

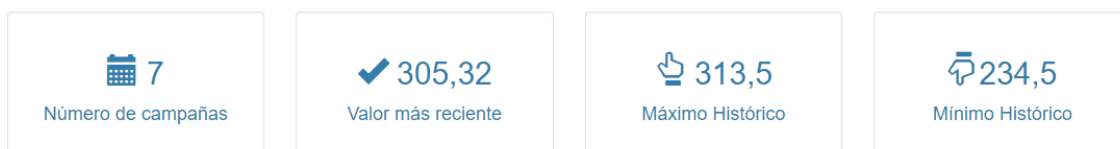


Figura B.9: Cuadro de mandos

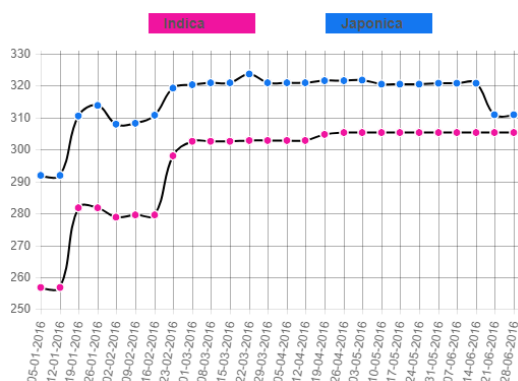


Figura B.10: Gráfico precios Arroz.

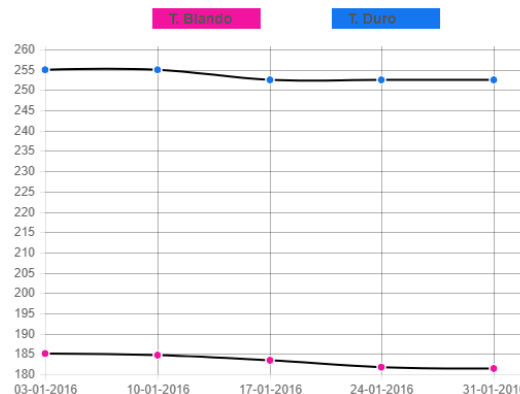


Figura B.11: Gráfico precios Trigo.

Por otra parte, también se adjunta en cada una de las respectivas vistas de cada producto agrícola, una tabla con estadísticas acerca de producciones, importaciones/exportaciones y superficie cultivada en España, en campañas anteriores:

B.7 Ganadería

Esta sección del menú principal se divide a su vez en dos categorías; *Ovino-Caprino* y *Vacuno*. Ambas vistas comparten el mismo esquema que las de la sección anterior, un pequeño cuadro de mandos para seleccionar el producto, el intervalo de tiempo así como los mismos indicadores descritos en la sección de agricultura:

Cuadro de mandos:



Figura B.12: Indicadores ganadería

La vista de ovino-caprino se encarga de mostrar los datos de precios referentes a leche de cabra y oveja con y sin denominación así como los precios de cabrito y cordero. Por otra parte la vista de vacuno, se ocupa de mostrar los precios de las diferentes categorías de ternera.

Anexo C

Mapeos de Elasticsearch

EN el siguiente capítulo se muestran todos los mapeos de datos utilizados durante el desarrollo de la herramienta. Este proceso¹ es fundamental cuando se utiliza ES, ya que define cómo se va almacenar cada documento, ya que dentro de un índice, pueden existir uno o más tipos de documentos, con este mecanismo podemos agrupar los documentos en función de diferentes criterios. Con este proceso, también se puede determinar el formato que va a tener cada uno de los campos, dentro de los documentos almacenados. A continuación se describen cada uno de los mapeos que se fueron creando a medida que avanzaba el proceso de desarrollo de la herramienta.

C.1 Versión 1: Visualización básica

Durante el desarrollo de esta versión se crea el primer índice, así como el primer tipo de documentos. Con ayuda del plugin para Chrome, *Sense*², se crea el índice *agroesi*, el cual alberga el tipo de documentos *precio*:

```
PUT agroesi
{
  "mappings": {
    "precio": {
      "properties": {
        "codigo": {"type": "text"},
        "precio": {"type": "double"},
        "dia": {"type": "integer"},
        "semana": {"type": "integer"},
        "nomMes": {"type": "text"},
        "numMes": {"type": "integer"},
        "año": {"type": "integer"},
        "fecha": {"type": "date", "format": "dd-MM-yyyy"}
      },
      "tipoPrecio": {"type": "text"},
      "fuente": {"type": "text"},
      "medida": {"type": "text"}
    }
  }
}
```

Figura C.1: Mapeo ES tipo “precio”.

¹<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping.html>

²<https://www.elastic.co/blog/found-sense-a-cool-json-aware-interface-to-elasticsearch>

C.2 Versión 1.0.1: El Torno

A medida que avanza el desarrollo de esta versión aparece un nuevo tipo de dato, *noticia*, por lo que se crea el correspondiente índice con el nombre de *rss*:

```
PUT rss/_mapping/noticia
{
  "properties":{
    "codigo": {"type":"text"},
    "titulo": {"type":"text"},
    "link": {"type":"text"},
    "descripcion": {"type":"text"},
    "etiquetas": {"type":"text"},
    "dia": {"type":"integer"},
    "numMes": {"type":"integer"},
    "nomMes": {"type":"text"},
    "año": {"type":"integer"},
    "fecha": {"type":"date","format":"dd-MM-yyyy"},
    "imagen": {"type":"text"}
  }
}
```

Figura C.2: Mapeo ES tipo “noticia”.

C.3 Versión 1.0.2: Zona de los montes

Durante el desarrollo de esta versión, también se decide realizar la adquisición de un nuevo tipo de datos, estadísticas referentes a los diferentes cultivos de los que hay datos de precios en la base de datos, por este motivo, es necesario la creación de un nuevo tipo de datos, bajo su correspondiente índice; en este caso, se crea el tipo de datos *estadística* bajo el índice *cereales*:

```
PUT cereales
{
  "mappings": {
    "estadística":{
      "properties":{
        "codigo": {"type":"text"},
        "campaña": {"type":"text"},
        "superficie": {"type":"double"},
        "produccion": {"type":"double"},
        "importaciones": {"type":"double"},
        "exportaciones": {"type":"double"}
      }
    }
  }
}
```

Figura C.3: Mapeo ES estadísticas

C.4 Versión 1.0.3: Ciudad Real

Con el desarrollo de la última versión se añade el módulo funcional encargado de ofrecer la previsión meteorológica, así como la posibilidad de realizar consultas sobre el histórico de datos para Ciudad Real, que se encuentra almacenado en la base de datos, para ello ha sido necesario la creación de un nuevo mapeo, *clima*, bajo el índice *meteo*:

```
PUT meteo
{
  "mappings": {
    "clima": {
      "properties": {
        "año": {"type": "integer"},
        "numMes": {"type": "integer"},
        "nomMes": {"type": "text"},
        "Tmedia": {"type": "double"},
        "Tmax": {"type": "double"},
        "Tmin": {"type": "double"},
        "PresionATM": {"type": "double"},
        "Humedad": {"type": "double"},
        "Lluvia": {"type": "double"},
        "Visibilidad": {"type": "double"},
        "VelocidadMed": {"type": "double"},
        "VelocidadMax": {"type": "double"}
      }
    }
  }
}
```

Figura C.4: Mapeo ES Clima

C.5 Estadísticas ES:

Al concluir el proceso de carga de datos, por el momento, de precios, noticias y estadísticas; a continuación se muestra el número de registros de cada tipo, proporcionados por ES:

```
"hits": {
  "total": 26038,
  "max_score": 1,
  "hits": [
```

Figura C.5: Número total de registros de tipo *Precio*.

```
"hits": {
  "total": 4857,
  "max_score": 1,
  "hits": [
```

Figura C.6: Número total de registros de tipo *Noticia*

```
"hits": {
  "total": 24,
  "max_score": 1,
  "hits": [
```

Figura C.7: Número total de registros de tipo *Estadística*

Anexo D

Historias de usuario

Portales de noticias que ofrezcan un canal RSS			
Sprint	Prioridad	Descripción	Pruebas
0.0.1	Alta	Para acotar los recursos con los que trabajar con el fin de evitar hacer un malgasto de tiempo, necesito tener una lista con los portales de noticias que disponen de servicio RSS.	Una vez, realizada la lista, comprobar si el canal de difusión que ofrecen, se actualiza de manera periódica todos los días.

Portales que publiquen precios agrícolas y ganaderos			
Sprint	Prioridad	Descripción	Pruebas
0.0.1	Alta	Dado que son muchos los portales necesito acotar esta lista de portales y que además tengan datos sobre los precios ya sea semanales o diarios, de las distintas materias primas de las que se va a ocupar la herramienta desarrollada.	Una vez, realizada la lista, comprobar que efectivamente con la lista de fuentes escogida, es posible obtener precios actuales así como un histórico con datos publicados en años anteriores.

Transformación de datos mediante técnicas de webscrapping			
Sprint	Prioridad	Descripción	Pruebas
0.0.1	Media	En caso de que no exista la opción de descargar los datos que me interesan, es necesario realizar la adquisición de datos directamente de las páginas web. Por este motivo necesito tener un mecanismo encargado de realizar esta labor cuando ese escenario se produzca. Por otro lado necesito buscar las tecnologías que mejor se adapten a las necesidades de este proyecto.	En caso de no poder obtener los datos de la manera tradicional, ya sea ficheros <i>Excel</i> o <i>PDF</i> , existe un mecanismo que es capaz de extraer los datos directamente de una página web.

Almacenamiento de datos			
Sprint	Prioridad	Descripción	Pruebas
0.0.1	Baja	Todo los datos adquiridos es necesario almacenarlos en algún tipo de base de datos. Como estudiante debo aprender que no sólo es posible almacenar datos usando un sistema tradicional basado en SQL, también existen bases de datos que también utilizan otros paradigmas, en el que datos semi-estructurados pueden ser también almacenados en una base de datos.	La tecnología escogida , permite almacenar los diferentes modelos de datos; precios, noticias, etc. y después es posible recuperar estos datos.

Visualización de los primeros resultados			
Sprint	Baja	Descripción	Pruebas
0.0.1	Media	De nada sirve tener datos almacenados si después no existe ningún tipo de herramienta que permita su visualización. Como desarrollador tengo que utilizar la mejor tecnología posible para las necesidades del proyecto en el que trabajo, y así poder ofrecer un servicio útil a los usuarios.	Con la tecnología empleada, es posible implementar una aplicación web capaz de conectarse a la base de datos y mostrar los datos que ésta almacena.

Documentación			
Sprint	Baja	Descripción	Pruebas
0.1.1	Media	Como desarrollador también tengo que dejar constancia por escrito de todo el trabajo realizado, cuando una iteración termina.	

Caso real			
Sprint	Prioridad	Descripción	Pruebas
0.1.1	Media	Como desarrollador tengo que enfocar la implementación de la herramienta con el objetivo de que en el futuro ésta será utilizada por personas reales con distintas necesidades.	

Control de versiones			
Sprint	Prioridad	Descripción	Pruebas
0.1.1	Media	Como desarrollador, tengo que utilizar tecnologías que me permitan gestionar los cambios que se producen en el código de la plataforma, a medida que avanza su desarrollo.	Es posible comprobar el historial de cambios que ha habido, durante el desarrollo de la plataforma.

Captura y almacenado de precios sobre ganadería			
Sprint	Prioridad	Descripción	Pruebas
0.1.1	Alta	Como desarrollador, tengo que ser capaz de implementar todos los mecanismos necesarios para la adquisición y almacenamiento de datos sobre precios relacionados con la ganadería.	Los datos se encuentran almacenados en la base de datos y es posible su visualización desde la aplicación web.

Captura y almacenado de noticias sobre ganadería			
Sprint	Prioridad	Descripción	Pruebas
0.1.1	Alta	Como desarrollador, tengo que ser capaz de implementar todos los mecanismos necesarios para la adquisición y almacenamiento de noticias relacionadas con la ganadería.	Los datos se encuentran almacenados en la base de datos y es posible su visualización desde la aplicación web.

Captura y almacenado de precios sobre agricultura			
Sprint	Prioridad	Descripción	Pruebas
0.1.2	Alta	Como desarrollador, tengo que ser capaz de implementar todos los mecanismos necesarios para la adquisición y almacenamiento de datos sobre precios relacionados con la agricultura.	Los datos se encuentran almacenados en la base de datos y es posible su visualización desde la aplicación web.

Captura y almacenado de noticias sobre agricultura			
Sprint	Prioridad	Descripción	Pruebas
0.1.2	Alta	Como desarrollador, tengo que ser capaz de implementar todos los mecanismos necesarios para la adquisición y almacenamiento de noticias relacionadas con la agricultura.	Los datos se encuentran almacenados en la base de datos y es posible su visualización desde la aplicación web.

Dashboard para consultar cualquier precio			
Sprint	Prioridad	Descripción	Pruebas
0.1.2	Media	Como usuario final, tengo que ser capaz de poder visualizar datos referentes a los precios del producto, en el intervalo de tiempo que yo elija, de modo que puedo entender los datos que aparecen.	Los datos se muestran con ayuda de gráficos y otros indicadores, de modo que puedo comprender los datos que se me muestran.

Pestaña con todas las noticias			
Sprint	Prioridad	Descripción	Pruebas
0.1.2	Media	Como usuario final, puedo consultar las noticias de meses y años anteriores , de modo que me resulta fácil desplazarme entre las diferentes noticias que se encuentran organizadas por el sector al que pertenecen, y que es posible leer el contenido en la misma.	Las noticias procedentes de la base de datos, se pueden visualizar en la plataforma y se pueden filtrar por año y mes, incluso también en base a su contenido, por medio de las etiquetas de cada una de ellas.

Captura y almacenado de datos climáticos			
Sprint	Prioridad	Descripción	Pruebas
0.1.3	Alta	Como desarrollador, tengo que ser capaz de implementar todos los mecanismos necesarios para la adquisición y almacenamiento de datos climáticos de Ciudad Real.	Los datos se encuentran almacenados en la base de datos y es posible su visualización desde la aplicación web.

Mecanismos de actualización de datos			
Sprint	Prioridad	Descripción	Pruebas
0.1.3	Media	Como desarrollador tengo que implementar mecanismos que se encarguen de actualizar los datos contenidos en la base de datos, tanto precios como noticias utilizando las técnicas oportunas.	Los mecanismos proporcionados son capaces de añadir nuevos datos a la base de datos, de manera autónoma.

Anexo E

Servicios de actualización de datos

EL siguiente capítulo tiene como objetivo mostrar más en detalle cómo funcionan los servicios implementados para la actualización de los datos contenidos en ES, cómo están implementados y cuáles son los datos que son capaces de actualizar por sí mismos.

E.1 Funcionamiento

Los servicios de Windows ¹ permiten que se puedan crear programas ejecutables de larga duración que funcionan en el sistema operativo. Estos programas se pueden ejecutar cuando arranca la máquina o pueden ser iniciados por el usuario, además se pueden parar en cualquier momento o incluso ser pausados.

Resultan útiles cuando se tienen que realizar tareas repetitivas de manera periódica y que no afecten a las tareas diarias del resto de usuarios.

E.2 Implementación e instalación

El entorno de desarrollo empleado durante el desarrollo de la plataforma, *Visual Studio*, ha permitido la implementación de dos servicios, como ya se ha mencionado en otros apartados de este documento. Uno de ellos, encargado de ejecutarse todos los días, el cual se ocupa de actualizar las noticias tanto de agricultura como de ganadería, así como el precio referente a la soja, ya que todos tienen en común la periodicidad, y cambian casi todos los días.

Por otro lado, el otro servicio se ocupa de actualizar los datos referentes a los precios del sector ganadero y el resto de las materias primas del sector agrícola, descritos a lo largo de este documento, los cuales se suelen publicar cada semana.

Una vez terminados, y generado el ejecutable correctamente, es necesario su instalación junto con el resto de servicios que ofrece el sistema operativo, para ello, hay que abrir la utilidad que viene junto a Visual Studio, *Símbolo del sistema para desarrolladores de VS*, una vez localizado el directorio del servicio, hay que ejecutar el comando “InstallUtil.exe” “ejecutable.exe”. Si no hay ningún fallo aparece lo siguiente:

¹[https://msdn.microsoft.com/es-es/library/d56de412\(v=vs.110\).aspx](https://msdn.microsoft.com/es-es/library/d56de412(v=vs.110).aspx)

```

Los parámetros afectados son:
  logtoconsole =
  logfile = C:\Users\Miguel Angel\Documents\Repositorios\Lonjas\Lonja
  assemblypath = C:\Users\Miguel Angel\Documents\Repositorios\Lonjas\

La fase de confirmación finalizó correctamente.
La instalación con transacciones ha finalizado.

```

Figura E.1: Instalación del servicio.

A continuación se muestra el diagrama de clases del servicio encargado de actualizar los precios referentes al sector ovino así como los precios del trigo (blando y duro) el maíz y el resto de precios del sector ganadero:

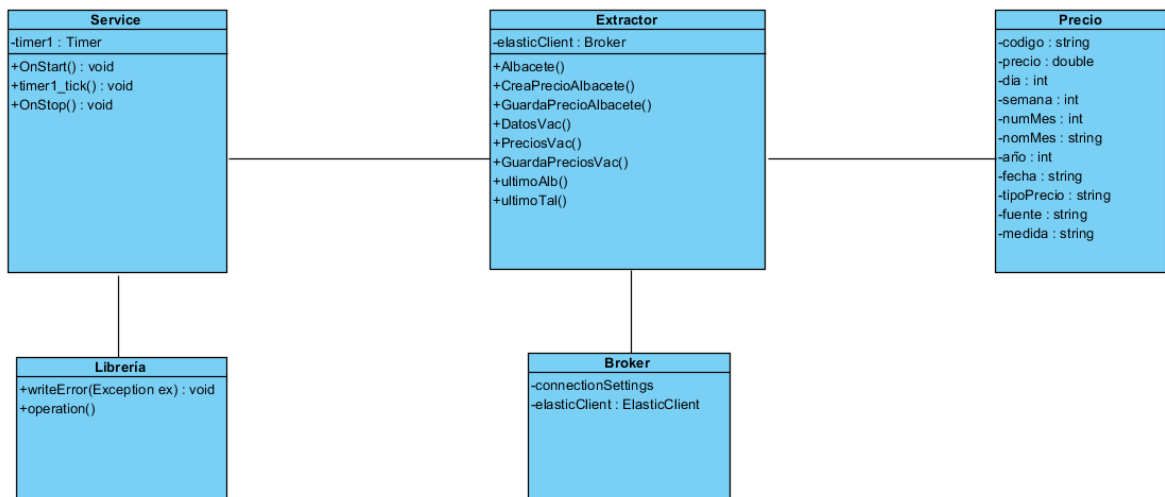


Figura E.2: Diagrama de clases (Servicio “Lonjas”).

La clase que controla el comportamiento del servicio es la denominada *Service*, donde se define la periodicidad de ejecución con el “timer” que tiene como atributo. Las llamadas a los diferentes métodos de la clase *Extractor* se realiza dentro del método “timer1_tick” que se ejecuta cada vez que pasa el intervalo de tiempo establecido. Por otro lado, la clase *Extractor*, contiene los métodos encargados de de realizar la adquisición de los datos, después se ayuda de la clase *Precio* para encapsular los datos referentes a los precios, antes de utilizar la clase *Broker* para guardar los datos en la base de datos.

El mismo escenario se produce en la ejecución del otro servicio, sólo que también dispone de una clase *Noticia* para encapsular los datos referentes a noticias tanto de agricultura como de ganadería.

E.3 Datos

Después de la puesta en marcha de ambos servicios los datos que se han actualizado sin intervención externa han sido los siguientes:

- Servicio “ActualizaES”: Encargado de la actualización de datos sobre noticias de ambos sectores así como del precio de la soja, es capaz de actualizar todos los datos sin intervención externa.
- Servicio “Lonjas”: Encargado de la actualización del resto de materias primas del sector agrícola, así como todos los referentes al sector ganadero, los resultados han sido parciales, si bien es cierto que actualiza todos los relacionados con el sector ovino-caprino, junto con el trigo blando, trigo duro y maíz. Para el sector vacuno, es necesario descargar y proporcionar el correspondiente fichero pdf, para que pueda actualizar los precios referentes a este sector. En cuanto al arroz, no se están actualizando los datos por el momento, ya que también requiere de la intervención externa y la descarga de los ficheros necesarios

E.4 Nuevo esquema de funcionamiento

Después de haber completado el desarrollo del proyecto, el esquema que se planteaba en el capítulo 1 en la figura 1.1, ha quedado obsoleto debido principalmente a que de ahora en adelante, el crawler ya no se usará más, para dejar paso a los dos servicios implementados, y que han sido descritos en el presente anexo. A continuación se muestra el nuevo esquema de la plataforma:



Figura E.3: Esquema final de la plataforma.

Referencias

- [And10] David J. Anderson. *Kanban: Successful Evolutionary Change for Your Technology Business*. Blue Hole Press; 3.8.2010 edition, 2010. ISBN 9780984521401.
- [CD10] Josep Curto Díaz. *Introducción al Business Intelligence*. Editorial UOC, S.L.; Edición: 1, 2010. ISBN 9788497888868.
- [Cor16] Francesco Corea. *Big Data Analytics: A Management Perspective*. Springer, 2016. ISBN 9783319389929.
- [CT17] Miriam Castillo Torroba. Agrimeteo: Una aplicación móvil para la gestión de explotaciones agrícolas. Technical Report 20170717, Universidad de Castilla-La Mancha. Escuela Superior de Informática, Julio 2017. 130 pp.
- [Duc11] Jon Duckett. *HTML and CSS: Design and Build Websites*. John Wiley Sons; 1st edition, 2011. ISBN 9781118008188.
- [ET17] Carlos Espino Timón. Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas open source que permiten su uso. Technical Report 20170116, Universidad de Cataluña, Enero 2017. 65 pp.
- [Fag16] Jonas Fagerberg. *ASP.NET MVC 5 - Building a Website with Visual Studio 2015 and C Sharp: The Tactical Guidebook*. CreateSpace Independent Publishing Platform, 2016. ISBN 9781535167864.
- [GT15] Clinton Gormley and Zachary Tong. *Elasticsearch: The Definitive Guide: A Distributed Real-Time Search and Analytics Engine*. O'Reilly Media; 1 edition, 2015. ISBN 9781449358549.
- [Inm16] Bill Inmon. *Data Lake Architecture: Designing the Data Lake and Avoiding the Garbage Dump*. Technics Publications Llc; Edición: First, 2016. ISBN 9781634621175.
- [JA13] Luis Joyanes Aguilar. *Big Data. Análisis De Grandes Volúmenes De Datos En Organizaciones*. Marcombo; Edición:1, 2013. ISBN 9788426720811.

- [Jef00] Ron Jeffries. *Extreme Programming Installed*. Addison-Wesley Professional; 1 edition, 2000. ISBN 9780201708424.
- [KB04] Cynthia Andres Kent Beck. *Extreme Programming Explained: Embrace Change, 2nd Edition (The XP Series)*. Addison-Wesley; 2nd edition, 2004. ISBN 9780321278654.
- [MSC13] Viktor Mayer-Schönberger and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. Hodder and Stoughton Ltd., 2013. ISBN 9781848547926.
- [PG16] Alejandro Puerta Gálvez. *Business Intelligence y las Tecnologías de la Información: 2ª Edición*. Createspace Independent Publish Platform; Edición: 2, 2016. ISBN 9781537596020.
- [Res14] CITO Research. White paper: Putting the data lake to work. Technical Report 20140410, Teradata Hortonworks, April 2014. 12 pp. URL <https://hortonworks.com/wp-content/uploads/2014/05/TeradataHortonworks.DataLake.White-Paper.20140410.pdf>.
- [Sie13] Eric Siegel. *Analítica Predictiva. Predecir el futuro utilizando Big Data*. Anaya; Títulos especiales; Edición: 1, 2013. ISBN 9788441534421.
- [Zha15] Qin Zhang. *Precision Agriculture Technology for Crop Farming*. CRC Press; 1 edition, 2015. ISBN 9781482251074.

Este documento fue editado y tipografiado con L^AT_EX empleando la clase **esi-tfg** (versión 0.20170828) que se puede encontrar en:
https://bitbucket.org/arco_group/esi-tfg

[respeta esta atribución al autor]