



Universidad Politécnica  
de Madrid



**Escuela Técnica Superior de  
Ingenieros Informáticos**

Grado en Ingeniería Informática

Trabajo Fin de Grado

**Extracción de Datos de Contratación  
Pública del Estado y Transformación en  
Ficheros en Formato CSV**

Autor: George Alexandru Stan

Tutor(a): Óscar Corcho García

Madrid, mayo 2021

Este Trabajo Fin de Grado se ha depositado en la ETSI Informáticos de la Universidad Politécnica de Madrid para su defensa.

*Trabajo Fin de Grado*

*Grado en Ingeniería Informática*

*Título:* Extracción de Datos de Contratación Pública del Estado y  
Transformación en Ficheros en Formato CSV  
mayo 2021

*Autor:* George Alexandru Stan

*Tutor:*

Óscar Corcho García  
Departamento de Inteligencia Artificial  
ETSI Informáticos  
Universidad Politécnica de Madrid

# Agradecimientos

A mi madre por su apoyo incondicional y por haberme dado la idea de realizar esta carrera.

A mi padre por no dejarme bajar el ritmo y enseñarme que los estudios son lo primero, porque para todas las cosas de la vida hay tiempo.

A mi hermana por haber creído en mí siempre y asumir varias de mis responsabilidades para poder centrarme en mis estudios.

A mis amigos y a mi novia por animarme siempre y ayudarme a despejar la mente en momentos difíciles.

Por último, a Óscar por darme su confianza y la oportunidad de realizar este trabajo bajo su supervisión.

“Debo estar dispuesto a renunciar a lo que soy con el fin de convertirme en lo que seré.”

Albert Einstein

## Resumen

En los últimos años, en todos los ámbitos de negocio se ha buscado sin cesar una mayor transparencia de las actividades realizadas, los bienes y servicios empleados y el capital invertido en las mismas, y sobre todo en el sector relacionado con los contratos y licitaciones Públicas. Es por esto, que desde la primera publicación de un estándar relacionado con los datos de las Administraciones Públicas este ha sido mejorado, continuando con la búsqueda de una mayor transparencia de los datos. Este estándar desarrollado por el Ministerio de Economía y Hacienda se denomina CODICE, y es un estándar nacional que marca las directrices a seguir por las entidades implicadas en el desarrollo de un Proceso de Licitación Pública en la generación de documentos y ficheros con los datos del mismo.

El segundo objetivo, pero no menos importante, de este afán por conseguir una mayor transparencia de estos procesos es permitir que los datos publicados sean más accesibles y permitir su uso y redistribución por un mayor número de personas, ya sean usuarios noveles o profesionales del área administrativa. Para ello, las Administraciones Públicas actualizan los datos de los diferentes contratos en la Plataforma de Contratación Pública del Estado, que a su vez, publica los documentos, ya en forma de datos abiertos, en el Portal Institucional del Ministerio de Hacienda mediante ficheros comprimidos que en su interior contiene archivos en formato XML.

A raíz de esto, y con la motivación de dotar a los datos de una mayor legibilidad, el objetivo de este proyecto es la extracción de los datos de Contratación Pública y generar ficheros con formato .csv, ya que los datos organizados en estructuras de tablas permiten un mayor entendimiento a los usuarios noveles o que estén poco familiarizados con las tecnologías.

Para alcanzar este objetivo, se ha realizado un análisis previo de la estructura de los ficheros de Contratación Pública y de los diferentes campos integrados en los mismos tal y como se explica más adelante en este documento para definir las opciones de generación de ficheros y saber qué ficheros generar y cuántos.

# Abstract

In recent years, in all areas of business there has been a constant search for greater transparency of the activities carried out, the goods and services used and the capital invested in them, and especially in the sector related to public contracts and tenders. This is why, since the first publication of a standard related to Public Administration data, it has been improved continuing the search for greater data transparency. This standard developed by the Ministry of Economy and Finance is called CODICE, and it is a national standard that sets the guidelines to be followed by the entities involved in the development of a Public Bidding Process for the generation of documents and files.

The second objective, but no less important, of this effort to achieve greater transparency in these processes, is to make the published data more accessible and allow its use and redistribution by a greater number of people, whether they are novice users or professionals in the administrative area. To this end, the Public Administrations update the data of the different contracts on the State Public Procurement Platform, which in turn publishes the documents, already in the form of open data, on the Institutional Portal of the Ministry of Finance through compressed files that contain files in XML format.

As a result of this, and with the aim of making the data more readable, the objective of this project is to extract the Public Procurement data and generate files in .csv format, since the data organized in table structures allow a better understanding for novice users or those who are unfamiliar with IT.

To achieve this objective, a previous analysis of the structure of the Public Procurement files and of the different fields integrated in them has been carried out as explained later in this document, in order to define the file generation options and to know which files to generate and how many.

# Tabla de contenidos

<b>1</b>	<b>Introducción</b>	<b>1</b>
<b>2</b>	<b>Estado del Arte</b>	<b>2</b>
2.1	Datos Abiertos en la Plataforma de Contratación Pública	5
2.1.1	¿Qué son los Datos Abiertos?	5
2.1.2	CODICE	8
2.1.2.1	<entry>	11
2.1.2.2	<cac-place-ext:ContractFolderStatus>	12
2.1.2.3	<cac-place-ext:LocatedContractingParty>	13
2.1.2.4	<cac:Party>	13
2.1.2.5	<cac:PostalAddress>	13
2.1.2.6	<cac:Contact>	14
2.1.2.7	<cac:ProcurementProject>	15
2.1.2.8	<cac:TenderResult>	16
2.1.2.9	<cac:WinningParty>	17
2.1.2.10	<cac:AwardedTenderedProject>	17
2.1.2.11	<cac:TenderingTerms>	18
2.1.2.12	<cac:TenderingProcess>	18
2.1.2.13	<cac:LegalDocumentReference>	19
2.1.2.14	<cac:TechnicalDocumentReference>	20
2.1.2.15	<cac-place-ext:ValidNoticeInfo>	21
2.2	Open Contracting Data Standard (OCDS)	23
2.2.1	Estructura de OCDS	24
2.2.1.1	Paquetes de entrega de proceso de contratación	25
2.2.1.2	Entrega de datos de proceso de contratación	26
<b>3</b>	<b>Desarrollo</b>	<b>38</b>
3.1	Análisis	39
3.1.1	Estructura de los ficheros de Contratación Pública del Estado	39
3.1.2	Restricciones y limitaciones de software	39
3.1.3	Requisitos de software	40
3.2	Diseño	41
3.2.1	Diagramas de flujo de software	42
3.2.2	Herramientas y lenguaje de desarrollo	44
3.2.2.1	Eclipse	44
3.2.2.2	Sublime Text y NotePad++	45

3.2.2.3	Java.....	46
3.3	Implementación de la Extracción de los Datos.....	46
3.3.1	API DOM.....	49
3.3.2	Estructura de ficheros de Código Fuente .....	50
3.3.3	Control de versiones.....	51
3.3.4	Pruebas realizadas .....	52
<b>4</b>	<b>Resultados y conclusiones .....</b>	<b>53</b>
4.1	Resultados.....	53
4.2	Conclusiones.....	54
<b>5</b>	<b>Análisis de Impacto .....</b>	<b>56</b>
5.1	Impacto personal.....	56
5.2	Impacto empresarial.....	56
5.3	Impacto social .....	57
5.4	Impacto económico.....	57
5.5	Impacto medioambiental .....	57
5.6	Impacto cultural.....	58
<b>6</b>	<b>Bibliografía .....</b>	<b>59</b>

# 1 Introducción

Las Administraciones Públicas en España actualizan todos los datos relativos a los Procesos de Contratación en la Plataforma de Contratación Pública del Estado. El motivo de esto es garantizar el mayor nivel de transparencia posible y la accesibilidad a dichos datos al mayor número de usuarios, ya sean estos usuarios noveles o profesionales en el Área Administrativa. La idea o el objetivo de garantizar un mayor nivel de acceso y transparencia a los datos se fundamenta en el estándar CODICE creado por el Ministerio de Economía y Hacienda del Gobierno de España. Este estándar establece la estructura y formato de los ficheros generados en los Procesos de Licitación Pública, así como su formato de publicación. Estos ficheros generados se recogen y se publican en el Portal Institucional del Ministerio de Hacienda como datos abiertos.

A parte del estándar desarrollado por el Ministerio de Economía y Hacienda, también se ha desarrollado un estándar internacional denominado OCDS, que, aunque defina una estructura y formato diferente que el estándar CODICE, guarda cierta similitud en cuanto a los campos definidos para un Proceso de Contratación, lo que permitiría en un futuro un mapeo de los datos de ambos estándares.

En función de lo explicado, y para permitir un mayor entendimiento de los datos abiertos para los usuarios noveles que sientan curiosidad por los Procesos de Licitación y para aquellos profesionales que quieran realizar tareas de tratamientos de datos y limpieza de los mismos, se propone el objetivo de crear un programa capaz de extraer los datos de los ficheros de Contratación Pública y generar ficheros en formato .csv, mostrando así los datos en un formato mayormente conocido como es el de tablas. Para cumplir este objetivo, se establecen los objetivos previos de análisis de la estructura de los ficheros del estándar CODICE, y el análisis del estándar OCDS, permitiendo así la posibilidad de realizar un mapeo entre ambos formatos en un futuro. Además, se establecen los objetivos de comprobar la integridad de los datos y de los ficheros generados, para garantizar una correcta estructura de los mismos y fiel a los datos publicados por el Ministerio.

Este documento se encuentra dividido en 4 secciones, siendo estas el *Estado del Arte*, que contiene la explicaciones de la Plataforma de Contratación Pública del Estado, los datos abiertos y los análisis previos de los estándares mencionados; la sección de *Desarrollo*, en la que se explica el proceso seguido para el desarrollo del programa y las herramientas empleadas para ello; *Resultados y Conclusiones y Análisis de Impacto*, en el que se detalla el impacto que puede tener este proyecto en la sociedad y en diferentes ámbitos, y la explicación de los resultados obtenidos y la conclusión final.

## 2 Estado del Arte

Al hablar sobre la Contratación Pública en España, implícitamente se hace referencia al término licitación, que, según la Academia Real Española en su primera entrada, es un sistema por el que se rige la Administración Pública cuando se quiere otorgar un determinado contrato a aquella empresa que ofrece las mejores condiciones. En su segunda definición, se entiende como el proceso previo por el que se otorgan permisos por parte de la administración para el empleo de bienes públicos o su uso de forma anormal. En resumen, podemos entender la licitación como los pasos a seguir por parte de las Administraciones Públicas para llevar a cabo determinadas tareas mediante contratos, cumpliendo siempre con las normativas vigentes sobre la Contratación Pública. Hoy en día sigue vigente la Ley 40/2015, de 1 de octubre, del Régimen Jurídico del Sector Público, tal y como se recoge en el Boletín Oficial del Estado. Los principios que se mencionan en dicha ley son:

- Eficacia
- Jerarquía
- Descentralización
- Desconcentración
- Coordinación
- Transparencia
- Planificación
- Dirección

A estos hay que sumar los ya mencionados en leyes anteriores, para poder proceder con la definición de los mismos a continuación.

- **Principio de necesidad y eficacia**, en el que se estipula que la licitación debe ser justificada por una causa de interés general, basada en el análisis y la identificación clara de lo que se pretende conseguir con la misma, y finalmente elegir aquella que sea idónea, para ser garantía de logro de aquello que se busca.
- **Principio de proporcionalidad**, donde se exige que la licitación debe ser regulada para cubrir las necesidades generadas, previamente habiendo constatado que no existen otras medidas con menos restricciones de derechos o que conlleven una imposición menor de obligaciones a quienes sean destinadas.
- **Principio de Seguridad Jurídica**, con el que se quiere aplicar la normativa vigente con el fin de generar un cuadro de normas que facilite la realización de dicha licitación por las empresas involucradas.
- **Principio de Transparencia**, que conlleva la publicación periódica de aquella información relevante y necesaria para poder asegurar la transparencia de la licitación relacionada con las Administraciones Públicas. En adición a esto, dichas Administraciones Públicas deberán permitir el acceso por parte del público a los documentos relacionados con el marco de normas y condiciones de la licitación.

- **Principio de eficiencia**, que busca en la medida de lo posible asegurar un trato razonable de los recursos públicos y evitar cargas excesivas no necesarias para el cumplimiento de las licitaciones.
- **Principio de Sostenibilidad financiera**, con el que se pretende garantizar la capacidad, en términos de capital, de realizar un compromiso acorde con la deuda pública.
- **Principio de Estabilidad presupuestaria**, gracias al que se establece que las Administraciones Públicas deberán mantener el estado de equilibrio estructural.
- **Principio de No Discriminación**, mediante el que se quiere garantizar que ningún operador económico participante en las licitaciones no será discriminado a razón de su lugar de residencia. En este principio se introduce el término de Operador Económico que, según la Wikipedia, se puede entender como la figura encargada de aplicar aquellas medidas de seguridad que resulten más eficaces y rigurosas, sin inferir en el tráfico del comercio internacional de mercancías [1].
- **Principio de Confianza Mutua**, en el que queda estipulado que las autoridades competentes involucradas en el desarrollo de la licitación confiarán y reconocerán las actividades realizadas por cualesquiera de las partes.
- **Principio de Libertad Establecimiento/Eficacia Nacional**, con el que se garantiza que cualquier empresa que se encuentre ejerciendo una actividad de servicios legalmente y que esté establecida en España, podrá ejercer dicha actividad desde cualquier ubicación dentro del territorio español.
- **Principio de Necesidad y proporcionalidad**, en el que se establece que cualquier medida restrictiva aplicada a la licitación deberá de ser debidamente justificada como resultante de la necesidad de salvaguardar datos de interés general.

Una vez explicados los principios que toda Administración Pública establecida dentro del territorio nacional debe de cumplir a la hora establecer una nueva licitación de Contratación Pública, podemos empezar a explicar qué es lo que un usuario puede encontrar al acceder a la Plataforma de Contratación del Sector Público [2]. Antes de seguir, es necesario remarcar que, aunque este sitio web sea del estado, no todas las Comunidades Autónomas se encuentran alojadas dentro de la Plataforma, sino por el contrario, hay una serie de comunidades y ciudades cuyas propias Plataformas de Contratación se encuentran conectadas con la Plataforma de Contratación del Sector Público. A continuación, se muestra una tabla con las Comunidades Autónomas y sus entidades dependientes alojadas y aquellas comunidades y ciudades con plataformas conectadas.

<b>Alojadas</b>	<b>Plataformas Conectadas</b>
Junta de Comunidades de Castilla-La Mancha	Plataforma de Contratación de la Generalitat de Cataluña
Generalitat Valenciana	Plataforma de Contratación del Gobierno Vasco
Govern de Illes Balears	Plataforma de Contratación de La Rioja
Junta de Extremadura	Plataforma de Contratación de la Comunidad de Madrid
Gobierno de Cantabria	Plataforma de Contratos Públicos de Galicia
Gobierno de Aragón	Plataforma de Contratación de la Junta de Andalucía
Ciudad Autónoma de Melilla	Plataforma de Contratación del Gobierno de Navarra
Junta de Castilla y León	
Gobierno de Canarias	
Ciudad Autónoma de Ceuta	
Región de Murcia	
Principado de Asturias	

*Tabla 1 Comunidades y Ciudades de la Plataforma de Contratación del Sector Público*

Al abrir el sitio web, los usuarios tienen de entrada una serie de filtros, y accesos, ya sean Empresas u Organismos Públicos.



*Ilustración 1 Filtros y Accesos*

Los filtros, como se puede ver en la Ilustración 1, son “Licitaciones”, que permite acceder a las últimas licitaciones publicadas; “Publicaciones”, que permite tener acceso a una serie de filtros para realizar búsquedas de licitaciones por formulario, contratos menores y búsquedas guiadas de licitaciones según qué criterios; y por último “Perfil Contratante”, que permite acceder a la información

relativa a los distintos Organismos presentes en la Plataforma de Contratación con sus respectivas licitaciones en caso de tenerlas.

En la Ilustración 1 también se pueden observar, tal y como se ha mencionado al principio de la sección, las opciones de acceso para Empresas, en el campo “Empresas” y Organismos Públicos, en el campo “Organismos Públicos”. En ambos casos los accesos se realizan mediante un ID de usuario y su respectiva contraseña. Para el desarrollo de este trabajo, se pueden obviar estas dos opciones de acceso, ya a que son meramente informativas y carecen de valor. Lo realmente interesante del portal se puede encontrar en el apartado de “Publicaciones”, donde los usuarios pueden acceder a las licitaciones publicadas y ver una serie de atributos pertenecientes a las mismas, como pueden ser el título, la categoría de la licitación o el importe necesario para llevarla a cabo. Pero, aun así, estos datos no sirven de mucho al usuario, debido a que no se pueden realizar labores de análisis o tratamiento de los datos desde el punto de vista de los programadores, y es aquí donde tienen especial lugar los términos Datos Abiertos y CODICE.

## **2.1 Datos Abiertos en la Plataforma de Contratación Pública**

En la Plataforma de Contratación Pública del Estado, al final de la barra de menú principal, se encuentra publicado un campo titulado “Datos Abiertos”, pero el enlace actualmente no se encuentra disponible. El objetivo de esta sección es introducir a los lectores de este proyecto y a aquellas personas que se sientan atraídas por los datos y el tratamiento de estos, a lo que los datos abiertos son, las utilidades que tienen, y la motivación de porqué estos existen, además de explicar la estructura de los Datos Abiertos de Contratación Pública que se encuentran publicados en forma de archivos comprimidos dentro del Portal Institucional del Ministerio de Hacienda [3], que recoge todas aquellas licitaciones empezando desde enero de 2014.

### **2.1.1 ¿Qué son los Datos Abiertos?**

En lo referente a este trabajo y a las Administraciones Públicas mencionadas en el apartado anterior, los datos abiertos buscan fomentar la transparencia en estas con el fin de aumentar la confianza en dichas instituciones gracias a que se muestran los trabajos realizados y cómo se desarrollan y gestionan los recursos públicos invertidos en los mismos, por medio de los portales de datos abiertos.

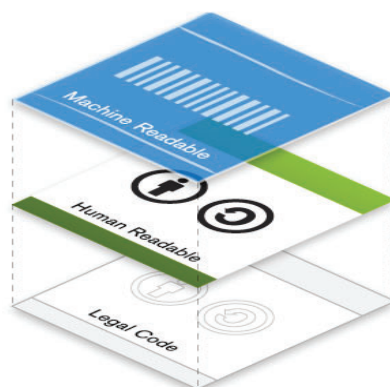
Según el portal de Administración Pública y Estado [4], los datos abiertos son aquellos datos disponibles para su uso, reutilización y redistribución por cualquier entidad, ya sean particulares o empresas, con la condición de hacer referencia al autor o fuente. A esta primera definición, se puede añadir la particularidad de que los datos pueden ser usados, reutilizados y redistribuidos sin ninguna restricción legal, social o tecnológica, según la Open Knowledge Foundation (OKF) [5].

Los Datos Abiertos se rigen por los 3 principios explicados a continuación:

- **Principio de disponibilidad y acceso**, cuyo fin es mantener los datos disponibles, los costes de su reproducción mínimos, debiendo optar siempre por la descarga de Internet, y ajustar los datos en la medida de lo posible para que sean cuanto más modificables y legibles.
- **Principio de reutilización y redistribución**, con el que se busca obtener el mayor rendimiento de los datos cuando son reutilizados o redistribuidos gracias al trabajo previo de estructuración de dichos datos, permitiendo siempre a las máquinas la lectura de los mismos.
- **Principio de participación universal**, en el que se estipula que los datos deben de ser disponibles, reutilizables y redistribuibles por cualquier persona o entidad sin ningún tipo de discriminación.

Esta filosofía busca seguir el mismo principio que el software libre o el código abierto, y con ello alcanzar el fin común de todos, siendo este el conocimiento libre.

Los Datos Abiertos pueden ser utilizados por cualquier persona o entidad, siempre que se haga referencia al autor o a las fuentes de las que se han obtenido, pero ¿cómo saber si unos datos son o no abiertos? Normalmente, cuando se habla de datos abiertos, y en relación con el desarrollo de este proyecto y las distintas Comunidades Autónomas, estos pueden encontrarse sin ninguna condición específica, o por el contrario se pueden encontrar bajo una licencia abierta, siendo esta mayoritariamente Creative Commons 3.0 Attribution 3.3 Unported, o CC BY 3.0 [6]. La característica fundamental de estas licencias es la capacidad de concesión de permisos que conceden a los autores o entidades encargadas de su publicación, además de poder ser propietarios legales de los derechos intelectuales. Según el artículo de Gestión de datos de investigación de la Biblioteca CEPAL (Comisión Económica para América Latina y el Caribe) [7], las licencias Creative Commons se rigen por tres niveles o capas, siendo estos el Código Legal, el nivel Legible por las personas y el nivel Legible por las máquinas o computadoras. Cada una de estas especifica la información legal desde el punto de vista de las Leyes Orgánicas, el lenguaje entendido por los seres humanos y el lenguaje entendido por los sistemas de la información.



*Ilustración 2 Tres capas Creative Commons*

Dependiendo el tipo de uso que se pueda dar a los datos abiertos, Creative Commons establece distintos tipos de licencias, explicadas a continuación:

- **Licencia CC BY (Atribución)**, que permite a los usuarios o entidades la distribución, tratamiento y creación a partir de los datos abiertos unos nuevos sin excluir fines comerciales con la singularidad de que se reconozca la creación pionera.



*Ilustración 3*  
CC BY

- **Licencia CC BY-SA (Atribución-CompartirIgual)**, que permite el mismo uso que la anterior con la condición de que las obras creadas a partir de los datos abiertos se publiquen con la misma licencia.



*Ilustración 4*  
CC BY-SA

- **Licencia CC BY-ND (Atribución-SinDerivadas)**, que permite reutilizar los datos incluidos fines comerciales, pero sin permitir la distribución de estos habiendo sido manipulados y referenciando siempre al autor.



*Ilustración 5*  
CC BY-ND

- **Licencia CC BY-NC (Atribución-NoComercial)**, que permite un uso equivalente a la licencia CC BY, aunque con la excepción de que no se puede emplear para fines comerciales.



*Ilustración 6*  
CC BY-NC

- **Licencia CC BY-NC-SA (Atribución-NoComercial-CompartirIgual)**, completamente igual que la licencia anterior y además se debe de otorgar a los datos tratados la misma licencia que los datos originales.



*Ilustración 7*  
CC BY-NC-SA

- **Licencia CC BY-NC-ND (Atribución-NoComercial-SinDerivadas)**, actualmente la más restrictiva, debido a que se permite exclusivamente la descarga y la redistribución manteniendo siempre la referencia al autor. No se permite su uso comercial ni su modificación.



*Ilustración 8*  
CC BY-NC-ND

- **Licencia CC0**, con la que se busca renunciar a los derechos de propiedad de las publicaciones y entregando tales al dominio público.



*Ilustración 9*  
CC0

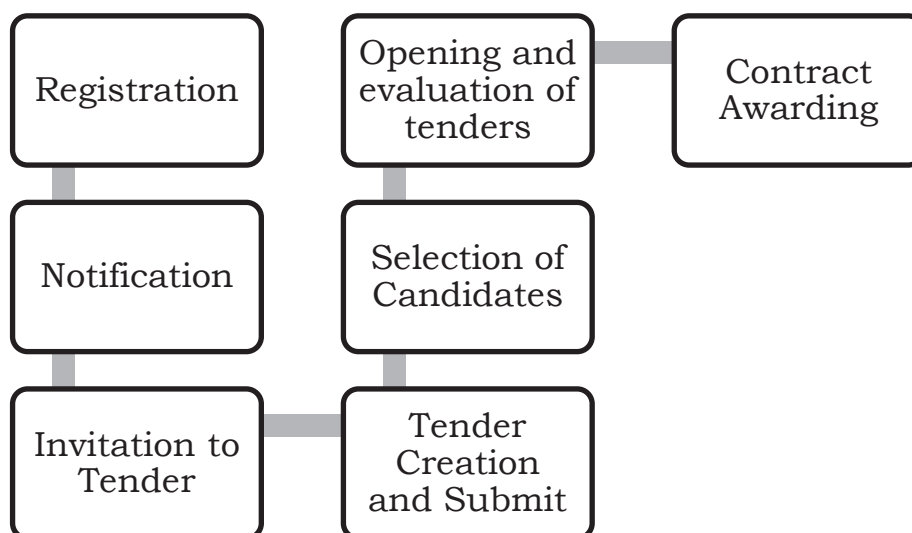
Para dar por finalizada esta sección, se quiere recalcar que las licencias son la manera clara de establecer y hacer llegar a los usuarios todas las condiciones a seguir para poder llevar a cabo el uso de los Datos Abiertos elegidos.

## 2.1.2 CODICE

CODICE, o Componentes y Documentos Interoperables para la Contratación Electrónica, es un proyecto desarrollado en el marco del Ministerio de Economía y Hacienda con el fin de garantizar a las Administraciones Generales del Estado la existencia de servicios de la información implicados en el desarrollo de documentos y mensajes que se rigen por un estándar común, cuyo objetivo es la creación de una estructura estandarizada para aquellos sistemas informáticos necesarios en el proceso de desarrollo de soluciones para la contratación electrónica, y para los sistemas informáticos de agentes económicos que puedan estar implicados en el proceso de licitación emprendido por los Órganos de Contratación Pública, en base a leyes y legislaciones internacionales que rigen dichas operaciones. El objetivo de este proyecto desarrollado por el Ministerio de Economía y Hacienda es dotar al Sector Público de una serie de documentos, mensajes y componentes XML estandarizados internacionalmente, por medio de la creación de la librería o modelo CODICE.

En una primera instancia se crearon los documentos CODICE, gracias a los que se ha conseguido la creación de la Plataforma de Contratación Pública del Estado y el posterior envío de mensajes entre los participantes de los procesos de licitación. Este desarrollo obtuvo su auge en el año 2007 con la publicación de la versión 1.06, aunque su comienzo queda marcado por el año 2006, en el que se creó un primer modelo para los documentos de Licitación Pública, y el desarrollo de un vocabulario siguiendo las líneas europeas en materia de procesos de contratación para el Sector Público. Este modelo XML se inspiró en los modelos de documentos ya existentes creados por UN/CEFACT o en las librerías UBL 2.0.

Además de establecer un modelo de documentos, este proyecto también establece el proceso por el que se rige la Contratación Pública en España, que viene ilustrado en el siguiente diagrama:



*Ilustración 10 Diagrama de desarrollo de licitaciones*

Cada uno de los pasos previos a la Adjudicación del Contrato e incluido este último, se detallan a continuación:

- **Registration (Registro)**, proceso en el que se establece el envío de toda la información necesaria acerca de los operadores económicos con el fin de su inscripción en el Registro Oficial de Licitadores y Empresas Clasificadas del Estado en la Dirección General del Patrimonio de Estado.
- **Notification (Notificación)**, proceso con el que se produce la publicación de la documentación relacionada a la licitación y el envío del Anuncio de Licitación al Boletín Oficial del Estado u otros diarios.
- **Invitation to Tender (Invitación a Licitación)**, proceso gracias al que se produce la intercomunicación entre los operadores económicos y el Órgano de Contratación.
- **Tender Creation and Submit (Creación y envío de la Oferta)**, proceso que posibilita la creación de una Oferta y su posterior publicación en el Órgano de Contratación.
- **Selection of Candidates (Selección de los candidatos)**, proceso que da acceso a la selección de aquellos candidatos en concursos con procedimientos abiertos y sin negociación pública.
- **Opening and Evaluation Tendering (Apertura y Evaluación de las Ofertas)**, proceso con el que se produce el intercambio de información entre los Órganos Adjudicadores y Evaluadores con el licitador para el comienzo de la apertura y análisis de las condiciones propuestas.
- **Contract Awarding (Adjudicación del Contrato)**, proceso por el que se producen envíos de información relativa a la adjudicación provisional del contrato y el posterior envío de información y documentación una vez se da por adjudicado definitivamente el contrato.

Posteriormente, se han creado las versiones del proyecto denominada CODICE 2.0, CODICE 2.01 y CODICE 2.02 debido a que, desde el momento de su primera publicación, CODICE ha sido adoptado rápidamente por el Sector Público y el Privado, además de la publicación de nuevas leyes de regulación de Contratos en el Sector Público, por lo que se ha encontrado la necesidad de actualizar y mejorar dicho modelo con uno que permita las opciones de aceptación o exclusión de una determinada licitación de cualesquiera de los candidatos participantes. Estas nuevas opciones de mejora han surgido gracias a las aportaciones de los Ministerios y agencias de la Administración del Estado, el Boletín Oficial del Estado o el Comité Técnico de UBL de OASIS.

## CODICE2

*Ilustración 11 CODICE 2*

Gracias a CODICE, se permite la generación de una serie de documentos en materia de Contratación Pública, que se describen a continuación.

<b>Documento CODICE</b>	<b>Nombre</b>	<b>Descripción</b>
AwardingNotification	Notificación de adjudicación	Mensaje enviado a los participantes de una licitación para comunicarles el resultado de la misma
Call For Tenders	Pliegos	Documentación relativa a la licitación, contratación, participación y adjudicación
ContractAwardNotice	Anuncio de Adjudicación	Anuncio publicado acerca del resultado de una licitación
ContractModificationNotice	Anuncio de Modificación de Contrato	Anuncio publicado acerca de las modificaciones de condiciones de una licitación
ContractNotice	Anuncio de Licitación	Anuncio con el que se da publicidad a la licitación
Guarantee	Garantía	Documento que permite construir una garantía provisional o definitiva de seguro de caución
GuaranteeCertificate	Certificado de Garantía	Certificado emitido sobre una garantía
GuaranteeCertificateRequest	Solicitud de Certificado de constitución de garantía	Solicitud sobre el estado de constitución de una garantía
PriorInformationNotice	Anuncio de información previa	Anuncio de intención de licitación dentro de un plazo de 12 meses

*Tabla 2 Extracto de tabla de Guía de implementación de documentos CODICE 2.0 [8]*

<b>Documento CODICE</b>	<b>Nombre</b>	<b>Descripción</b>
Tender	Oferta	Documentación con las condiciones relativas a una oferta determinada
TenderEnvelope	Sobre de oferta	Sobre con documentación que un licitador presenta para la adjudicación de una oferta
TenderQualification	Documentación Administrativa	Documentación presentada por parte del licitador para acreditar el cumplimiento de los requisitos de participación
TenderQualificationResponse	Resultado de admisión/exclusión	Mensaje enviado a un licitador acerca de su admisión/exclusión del procedimiento
TenderReceipt	Justificante de recepción de la oferta	Documento devuelto al licitador para demostrar la presentación a la misma

*Tabla 3 Continuación Extracto de tabla de Guía de implementación de documentos CODICE 2.0 [8]*

Una vez explicados el proceso de desarrollo de una licitación con sus respectivas fases y los documentos que se crean antes, durante y después de la licitación, se procede a explicar los componentes, nodos y estructuras XML que establece el estándar. Para ello, se toma de referencia la documentación empleada en el desarrollo de este proyecto, referente a los contratos de Licitación Pública.

#### **2.1.2.1 <entry>**

Nodo que hace referencia a un nuevo contrato. En el desarrollo del programa implicado en la extracción de los datos relativos a cada contrato, este nodo se emplea para la identificación de la totalidad de contratos. Contiene 6 nodos hijos, los 5 primeros a modo de resumen del contrato, y el último con la información relativa a todos los datos del proceso de adjudicación del mismo, que son:

- **<id>**, contiene la url de la Plataforma de Contratación Pública del Estado con el id del contrato como parámetro añadido.
- **<link>**, contiene una referencia a la Plataforma de Contratación Pública del Estado con la información detallada del proyecto.
- **<summary type="text">**, contiene un resumen con la información más relevante del contrato.
- **<title>**, define el título del contrato.
- **<updated>**, contiene la última fecha de actualización del contrato.
- **<cac-place-ext:ContractFolderStatus>**, contiene toda la información detallada y categorizada por nodos del contrato.

```
<entry>
  <id>https://contrataciondelestado.es/sindicacion/licitacionesPerfilCon
  tratante/789966</id>
  <link
  href="https://contrataciondelestado.es/wps/poc?uri=deeplink:detalle_li
  citacion&idEvl=HwExWI%2FHuQQQK2TEfXGy%2BA%3D%3D"/>
  <summary type="text">Id licitación: 4270012027200; Órgano de
  Contratación: Jefatura de la Sección Económico-Administrativa 27 -
  Base Aérea de Getafe; Importe: 50377.14 EUR; Estado: RES</summary>
  <title>Adquisición de repuestos para el ILS AMS 2100</title>
  <updated>2013-01-11T09:41:37.139+01:00</updated>
  <cac-place-ext:ContractFolderStatus>
```

### 2.1.2.2 <cac-place-ext:ContractFolderStatus>

Como se ha mencionado con anterioridad, este nodo contiene toda la información relativa y detallada de la adjudicación del contrato de Licitación Pública. Contiene además el posible estado en el que se encuentre o el listado de anuncios publicados. Se procede a explicar cada uno de los nodos hijos con sus respectivos nodos internos, y al finalizar se muestra una ilustración para dar a los lectores una visión global de todos los elementos del nodo.

- **<cbc:ContractFolderID>**, contiene el número de expediente que se corresponde a la licitación. Hay que destacar que este número es únicamente aplicable a los documentos de licitación electrónica.
- **<cbc-place-ext:ContractFolderStatusCode>**, contiene el código referente al estado del expediente, es decir es un código empleado para relacionar el expediente a cada una de las fases mostradas en el diagrama de Desarrollo de Licitaciones.

```
<cbc:ContractFolderID>4270012027200</cbc:ContractFolderID>
  <cbc-place-ext:ContractFolderStatusCode languageID="es"
  listURI="https://contrataciondelestado.es/codice/cl/2.04/SyndicationCo
  ntractFolderStatusCode-2.04.gc">RES</cbc-place-
  ext:ContractFolderStatusCode>
  <cac-place-ext:LocatedContractingParty>
```

### 2.1.2.3 <cac-place-ext:LocatedContractingParty>

Contiene un elemento encargado de mostrar la listURI que describe el órgano involucrado en la contratación junto con la información relativa a la ubicación del mismo dentro de la estructura organizativa o entidad.

- <cbc:ContractingPartyTypeCode listURI="">, nodo contenedor de la listURI mencionada.

```
<cac-place-ext:LocatedContractingParty>
  <cbc:ContractingPartyTypeCode
    listURI="http://contrataciondelestado.es/codice/cl/1.04/ContractingAuthorityCode-1.04.gc">1</cbc:ContractingPartyTypeCode>
</cac:Party>
```

### 2.1.2.4 <cac:Party>

Contiene toda la información relativa a los datos del operador económico.

- <cbc:WebsiteURI>, contiene la URL del órgano a cargo de la contratación.
- <cac:PartyIdentification>, contiene un identificador que hace referencia a un valor con significado semántico empleado por las aplicaciones que trabajen con CODICE.
- <cbc:ID schemeName="">, nodo contenedor del identificador mencionado.
- <cac:PartyName>, contiene el nombre del operador económico.
- <cbc:Name>, nodo contenedor del nombre mencionado.

```
<cac:Party>
  <cbc:WebsiteURI>http://www.ejercitodelaire.mde.es/Contratacion</cbc:WebsiteURI>
  <cac:PartyIdentification>
    <cbc:ID schemeName="DIR3">EA0003016</cbc:ID>
  </cac:PartyIdentification>
  <cac:PartyName>
    <cbc:Name>Jefatura de la Sección Económico-Administrativa 27 -
      Base Aérea de Getafe</cbc:Name>
  </cac:PartyName>
  <cac:PostalAddress>
```

### 2.1.2.5 <cac:PostalAddress>

Contiene la información relativa a la dirección física del operador económico.

- <cbc:CityName>, contiene el nombre de la ciudad en la que se encuentra establecido el operador.
- <cbc:PostalZone>, contiene el código postal del operador.
- <cac:AddressLine>, contiene la dirección (calle, avenida, plaza, etc.) del operador.
- <cbc:Line>, nodo contenedor de la dirección mencionada.

- **<cac:Country>**, contiene el id de identificación y el nombre del país en el que está situado el operador económico.
- **<cbc:IdentificationCode listURI="">**, nodo contenedor del identificador mencionado.
- **<cbc:Name>**, nodo contenedor del nombre del país.

```

<cac:PostalAddress>
  <cbc:CityName>Getafe</cbc:CityName>
  <cbc:PostalZone>28901</cbc:PostalZone>
  <cac:AddressLine>
    <cbc:Line>Negociado de Contratación de la SEA de la Base Aérea
      de Getafe, Plaza Coronel Polanco, s/n</cbc:Line>
  </cac:AddressLine>
  <cac:Country>
    <cbc:IdentificationCode listURI="http://docs.oasis-
      open.org/ubl/os-ubl-2.0/cl/gc/default/CountryIdentificationCode-
      2.0.gc">ES</cbc:IdentificationCode>
    <cbc:Name>España</cbc:Name>
  </cac:Country>
</cac:PostalAddress>
<cac:Contact>

```

#### 2.1.2.6 <cac:Contact>

Contiene la información relativa a las distintas vías de contacto con el operador económico existentes.

- **<cbc:Name>**, contiene el nombre del operador económico.
- **<cbc:Telephone>**, contiene el teléfono de contacto del operador económico.
- **<cbc:Telefax>**, contiene el fax de contacto del operador económico.
- **<cbc:ElectronicMail>**, contiene el correo electrónico del operador económico.

```

<cac:Contact>
  <cbc:Name>Jefatura de la Sección Económico-Administrativa 27 - Base
    Aérea de Getafe</cbc:Name>
  <cbc:Telephone>917798214</cbc:Telephone>
  <cbc:Telefax>917798624</cbc:Telefax>
  <cbc:ElectronicMail>sea27contratacion@ea.mde.es</cbc:ElectronicMail>
</cac:Contact>

```

### 2.1.2.7 <cac:ProcurementProject>

Define un proyecto de compra y contiene la información definida en los pliegos del proyecto original.

- <cbc:Name>, contiene el nombre del proyecto de compra.
- <cbc:TypeCode listURI="">, contiene el código que describe el órgano involucrado en el proyecto de compra junto con la información relativa a la ubicación del mismo dentro de la estructura organizativa.
- <cbc:SubTypeCode listURI="">, contiene el subcódigo que describe el órgano involucrado en el proyecto de compra junto con la información relativa a la ubicación del mismo dentro de la estructura organizativa.
- <cac:BudgetAmount>, contiene la información relativa al valor total estimado para el contrato y el presupuesto disponible para la licitación. En concreto se informa acerca del presupuesto total para el correcto desarrollo del contrato. Es imprescindible destacar que todos los nodos contenedores de información de importes necesitan declarar el atributo @currencyID para definir la moneda en la que se especifican dichos importes.
- <cbc:EstimatedOverallContractAmount currencyID="">, contiene la estimación del valor total del contrato siguiendo siempre las leyes vigentes.
- <cbc:TotalAmount currencyID="">, contiene la cantidad total presupuestada con el importe neto junto con la adición de impuestos y costes derivados del mismo.
- <cbc:TaxExclusiveAmount currencyID="">, contiene únicamente el presupuesto neto sin ningún tipo de añadidos.
- <cac:RequiredCommodityClassification>, contiene el código empleado para la clasificación correcta del servicio o producto, siendo este en la mayoría de los casos el código CPV<sup>1</sup>.
- <cbc:ItemClassificationCode listURI="">, nodo contenedor del código explicado.
- <cac:RealizedLocation>, contiene el lugar de realización del proyecto de compra.
- <cbc:CountrySubentity>, especifica el nombre del lugar de realización del proyecto de compra.
- <cac:Address>, contiene la dirección del lugar en que se desarrolla el proyecto de compra.
- <cac:Country>, contiene el país donde se desarrolla el proyecto de compra.
- <cbc:IdentificationCode listURI="">, contiene el código del país mencionado.
- <cbc:Name>, contiene el nombre del país mencionado.

---

<sup>1</sup> Los códigos CPV o Common Procurement Vocabulary son aquellos implicados en la clasificación e identificación de todos aquellos contratos con posibilidad de ser incluidos en procesos de Licitación Pública en la Unión Europea. [13]

```

<cac:ProcurementProject>
  <cbc:Name>Adquisición de repuestos para el ILS AMS 2100</cbc:Name>
  <cbc:TypeCode
    listURI="http://contrataciondelestado.es/codice/cl/2.02/ContractCode-
    2.02.gc">1</cbc:TypeCode>
  <cbc:SubTypeCode
    listURI="http://contrataciondelestado.es/codice/cl/1.04/WorksContractC
    ode-1.04.gc">2</cbc:SubTypeCode>
  <cac:BudgetAmount>
    <cbc:EstimatedOverallContractAmount
      currencyID="EUR">50377.14</cbc:EstimatedOverallContractAmount>
    <cbc:TotalAmount currencyID="EUR">60956.34</cbc:TotalAmount>
    <cbc:TaxExclusiveAmount
      currencyID="EUR">50377.14</cbc:TaxExclusiveAmount>
  </cac:BudgetAmount>
  <cac:RequiredCommodityClassification>
    <cbc:ItemClassificationCode
      listURI="http://contrataciondelestado.es/codice/cl/1.04/CPV2007-
      1.04.gc">32500000</cbc:ItemClassificationCode>
  </cac:RequiredCommodityClassification>
  <cac:RealizedLocation>
    <cbc:CountrySubentity>ESPAÑA</cbc:CountrySubentity>
    <cbc:CountrySubentityCode
      listURI="http://contrataciondelestado.es/codice/cl/2.0/NUTS-
      2009.gc">ES</cbc:CountrySubentityCode>
    <cac:Address>
      <cac:Country>
        <cbc:IdentificationCode listURI="http://docs.oasis-
        open.org/ubl/os-ubl-
        2.0/cl/gc/default/CountryIdentificationCode-
        2.0.gc">ES</cbc:IdentificationCode>
        <cbc:Name>España</cbc:Name>
      </cac:Country>
    </cac:Address>
  </cac:RealizedLocation>
</cac:ProcurementProject>

```

### 2.1.2.8 <cac:TenderResult>

Contiene los resultados relativos a la licitación, es decir los anuncios de adjudicación y formalización.

- **<cbc:ResultCode listURI="">**, contiene el código relativo al resultado, pudiendo ser este equivalente a una adjudicación, formalización, renuncia, desistimiento, o directamente que no se rellene quedando el campo vacío.
- **<cbc:Description>**, contiene una descripción de los motivos de la adjudicación.
- **<cac:Contract>**, contiene la información relativa al contrato.
- **<cbc:IssueDate>**, contiene la descripción de la fecha de inicio del contrato.

```

<cac:TenderResult>
  <cbc:ResultCode
    listURI="http://contrataciondelestado.es/codice/cl/2.02/TenderResultCo
    de-2.02.gc">8</cbc:ResultCode>
  <cbc:Description>Por ser la única oferta presentada y ajustarse a lo
  exigido en los pliegos que rigen el contrato</cbc:Description>
  <cac:Contract>
    <cbc:IssueDate>2012-12-07</cbc:IssueDate>
  </cac:Contract>
  <cac:WinningParty>

```

### 2.1.2.9 <cac:WinningParty>

Contiene la información acerca de la empresa o entidad ganadora de la licitación y por tanto a la que se le adjudica dicho contrato.

- <cac:PartyIdentification>, contiene el NIF de la entidad o empresa a la que se adjudica el contrato.
- <cbc:ID schemeName="">, nodo contenedor del NIF mencionado,
- <cac:PartyName>, contiene el nombre de la entidad o empresa ganadora de la licitación.
- <cbc:Name>, nodo contenedor del nombre mencionado.

```

<cac:WinningParty>
  <cac:PartyIdentification>
    <cbc:ID schemeName="NIF">B83028084</cbc:ID>
  </cac:PartyIdentification>
  <cac:PartyName>
    <cbc:Name>DEIMOS SPACE S.L.</cbc:Name>
  </cac:PartyName>
</cac:WinningParty>
<cac:AwardedTenderedProject>

```

### 2.1.2.10 <cac:AwardedTenderedProject>

Contiene información aportada por el licitador que haya sido adjudicada a la entidad ganadora de la licitación, en concreto, el presupuesto para el desarrollo del contrato y el presupuesto junto con los importes derivados.

- <cac:LegalMonetaryTotal>, contiene los importes relativos al proyecto.
- <cbc:TaxExclusiveAmount currencyID="">, contiene el importe total sin impuestos.
- <cbc:PayableAmount currencyID="">, contiene el importe total a abonar por el desarrollo del proyecto.

```

<cac:AwardedTenderedProject>
  <cac:LegalMonetaryTotal>
    <cbc:TaxExclusiveAmount
      currencyID="EUR">50377.08</cbc:TaxExclusiveAmount>
    <cbc:PayableAmount
      currencyID="EUR">60956.27</cbc:PayableAmount>
  </cac:LegalMonetaryTotal>
</cac:AwardedTenderedProject>
</cac:TenderResult>
<cac:TenderingTerms>

```

### 2.1.2.11 <cac:TenderingTerms>

Contiene las condiciones asociadas a la licitación que vienen estipuladas en los pliegos.

- **<cbc:VariantConstraintIndicator>**, especifica el valor verdadero o falso dependiendo de si la empresa o entidad que se presenta a una determinada licitación acepta cambios en las ofertas ya presentes.
- **<cac:Language>**, indica el idioma en que vienen descritas las condiciones de la licitación.
- **<cbc:ID>**, nodo contenedor del identificador asociado al idioma de descripción de las condiciones de licitación.

```

<cac:TenderingTerms>
  <cbc:VariantConstraintIndicator>>false</cbc:VariantConstraintIndicator>
  <cac:Language>
    <cbc:ID>es</cbc:ID>
  </cac:Language>
</cac:TenderingTerms>
<cac:TenderingProcess>

```

### 2.1.2.12 <cac:TenderingProcess>

Contiene información relativa al proceso de licitación de la actividad deseada, es decir toda aquella información que contenga alguna particularidad en especial.

- **<cbc:ProcedureCode listURI="">**, contiene información acerca del tipo del procedimiento, pudiendo ser este abierto, negociado con o sin publicidad, contrato menor, diálogo competitivo o restringido.
- **<cbc:UrgencyCode listURI="">**, contiene información relativa al tipo de tramitación que puede tener la licitación, entre los que se distinguen 3 niveles de prioridad, siendo estos ordinaria, urgente o de emergencia.
- **<cac:TenderSubmissionDeadlinePeriod>**, contiene la información mediante la que se establece un periodo temporal de aceptación de ofertas para una determinada licitación.

- **<cbc:EndDate>**, establece la fecha de terminación del periodo de aceptación de ofertas.
- **<cbc:EndTime>**, establece la hora de terminación del periodo de aceptación de ofertas.

```

<cac:TenderingProcess>
  <cbc:ProcedureCode
    listURI="https://contrataciondelestado.es/codice/cl/2.04/SyndicationTenderingProcessCode-2.04.gc">3</cbc:ProcedureCode>
  <cbc:UrgencyCode
    listURI="http://contrataciondelestado.es/codice/cl/1.04/DiligenceTypeCode-1.04.gc">1</cbc:UrgencyCode>
  <cac:TenderSubmissionDeadlinePeriod>
    <cbc:EndDate>2012-11-13</cbc:EndDate>
    <cbc:EndTime>12:00:00</cbc:EndTime>
  </cac:TenderSubmissionDeadlinePeriod>
</cac:TenderingProcess>
<cac:LegalDocumentReference>

```

Antes de seguir con la descripción de los últimos nodos presentes en el contrato, es necesario explicar la existencia de 3 tipos diferentes de documentos que pueden ser añadidos al contrato a modo de Anexo. Esto es así debido que CODICE genera documentos XML que se rigen por un estándar, pero en ciertas ocasiones surge la necesidad de adjuntar información que no puede ser insertada directamente en XML. A continuación, se mencionan los 3 tipos de Anexos que CODICE permite:

- **<cac:LegalDocumentReference>**, que permite el anexo de los Pliegos relativos a las Cláusulas Administrativas Particulares pertenecientes a la licitación (PCAP).
- **<cac:TechnicalDocumentReference>**, que permite el anexo de los Pliegos relativos a las Prescripciones Técnicas Particulares asociadas al Proyecto (PPT).
- **<cac:AdditionalDocumentReference>**, que permite establecer como anexo el documento necesario indiferentemente de su formato o composición.

### 2.1.2.13 <cac:LegalDocumentReference>

Contiene la documentación de los Pliegos de Cláusulas Administrativas Particulares a modo de Anexo. Esta información anexada puede venir en forma de documentos con cualquier formato, ya sea PDF, Microsoft Office Word y relacionados.

- **<cbc:ID>**, contiene el identificador del documento anexado, siendo este mayoritariamente el propio nombre del fichero.
- **<cac:Attachment>**, nodo que posibilita la referencia del documento.
- **<cac:ExternalReference>**, con este nodo se establece la referencia externa del documento.

- **<cbc:URI>**, contiene la URI del fichero.
- **<cbc:DocumentHash>**, contiene un Hash que se calcula a partir del documento con el objetivo de detectar cambios en el documento siendo este externo.

```

<cac:LegalDocumentReference>
  <cbc:ID>20120272-PCAP.pdf</cbc:ID>
  <cac:Attachment>
    <cac:ExternalReference>
      <cbc:URI>http://contrataciondelestado.es/wps/wcm/connect/
PLACE_es/Site/area/docAccCmpnt?srv=cmpnt&amp;cmpntname=Ge
tDocumentsById&amp;source=library&amp;DocumentIdParam=323
a8738-32a7-4aa3-93dc-7ad40190858f</cbc:URI>
      <cbc:DocumentHash>XUH4t1gDii3Nugbqq75CZUx9No8=</cbc:Docum
entHash>
    </cac:ExternalReference>
  </cac:Attachment>
</cac:LegalDocumentReference>
<cac:TechnicalDocumentReference>

```

#### 2.1.2.14 <cac:TechnicalDocumentReference>

Contiene la documentación de Pliegos de Prescripciones Técnicas Particulares del Proyecto a modo de Anexo. Esta información anexada puede venir en forma de documentos con cualquier formato, ya sea PDF, Microsoft Office Word y relacionados.

- **<cbc:ID>**, contiene el identificador del documento anexado, siendo este mayoritariamente el propio nombre del fichero.
- **<cac:Attachment>**, nodo que posibilita la referencia del documento.
- **<cac:ExternalReference>**, con este nodo se establece la referencia externa del documento.
- **<cbc:URI>**, contiene la URI del fichero.
- **<cbc:DocumentHash>**, contiene un Hash que se calcula a partir del documento con el objetivo de detectar cambios en el documento siendo este externo.

```

<cac:TechnicalDocumentReference>
  <cbc:ID>20120272-PPT.pdf</cbc:ID>
  <cac:Attachment>
    <cac:ExternalReference>
      <cbc:URI>http://contrataciondelestado.es/wps/wcm/connect/
PLACE_es/Site/area/docAccCmpnt?srv=cmpnt&amp;cmpntname=Ge
tDocumentsById&amp;source=library&amp;DocumentIdParam=123
bb746-6f0f-45e2-ac21-7e4a6469ee86</cbc:URI>
      <cbc:DocumentHash>SEGHoI9NTdBbkVA6T5xxvrjRy38=</cbc:Docum
entHash>
    </cac:ExternalReference>
  </cac:Attachment>
</cac:TechnicalDocumentReference>
<cac:place-ext:ValidNoticeInfo>

```

### 2.1.2.15 <cac-place-ext:ValidNoticeInfo>

Contiene información relativa a los anuncios vigentes sobre la licitación.

- <cbc-place-ext:NoticeTypeCode listURI="">, contiene el código con el que se identifica el tipo de anuncio, como puede ser el Anuncio de Licitación o Adjudicación entre otros.
- <cac-place-ext:AdditionalPublicationStatus>, contiene la información sobre el envío del anuncio a otro medio diferente.
- <cbc-place-ext:PublicationMediaName>, contiene el nombre del medio en el que se ha publicado el anuncio.
- <cac-place-ext:AdditionalPublicationDocumentReference>, contiene la información relativa a la publicación del anuncio en otros medios.
- <cbc:IssueDate>, contiene la fecha de publicación en el medio externo.

```
<cac-place-ext:ValidNoticeInfo>
  <cbc-place-ext:NoticeTypeCode
    listURI="http://contrataciondelestado.es/codice/cl/2.04/TenderingNo
    ticeTypeCode-2.04.gc">DOC_CAN_ADJ</cbc-place-ext:NoticeTypeCode>
  <cac-place-ext:AdditionalPublicationStatus>
    <cbc-place-ext:PublicationMediaName>Perfil del
    Contratante</cbc-place-ext:PublicationMediaName>
    <cac-place-ext:AdditionalPublicationDocumentReference>
      <cbc:IssueDate>2013-01-11</cbc:IssueDate>
    </cac-place-ext:AdditionalPublicationDocumentReference>
  </cac-place-ext:AdditionalPublicationStatus>
</cac-place-ext:ValidNoticeInfo>
</entry>
```

Para mostrar una visión general de los campos explicados, se muestra la siguiente ilustración. Cabe destacar que todos estos elementos explicados en detalle son una parte de la totalidad de campos y documentos que CODICE permite crear. Como se ha explicado en el comienzo de esta sección, se han usado a modo de guía los ficheros empleados en la creación del programa encargado de extraer los datos de los mismos y generar ficheros .csv. Toda la información relativa a CODICE se puede encontrar en las guías de desarrollo e implementación [9], [10], [11], así como en la Plataforma de Contratación Pública del Estado.

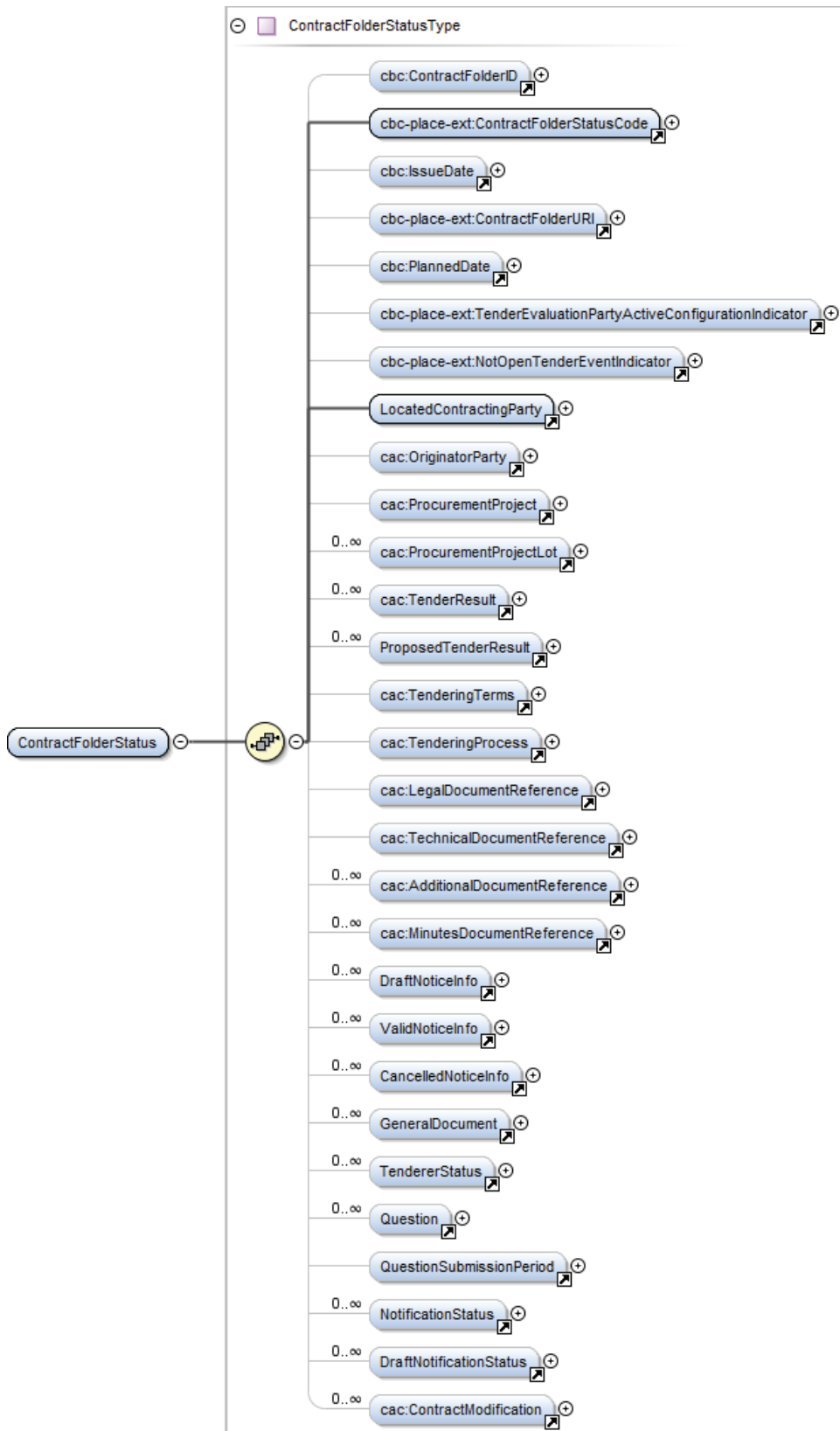


Ilustración 12 ContractFolderStatus

## 2.2 Open Contracting Data Standard (OCDS)

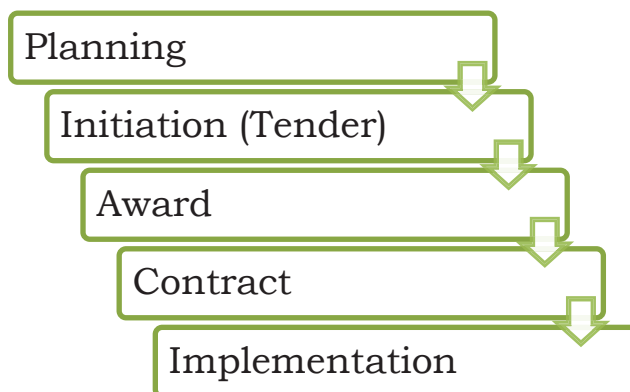
Una vez que se ha explicado la Plataforma de Contratación Pública del Estado, y el estándar desarrollado por el Gobierno de España, y en concreto el Ministerio de Hacienda y Economía, se procede a entrar en detalle en el Estándar Internacional desarrollado para la creación de documentación en las Contrataciones Públicas.

OCDS, u Open Contracting Data Standard, es un estándar creado con el fin de facilitar la tarea de creación de documentos y organización de datos dentro de las etapas por las que pasa un contrato cuando es licitado y finalmente adjudicado. Al igual que CODICE, su objetivo final es apoyar la transparencia de las actividades realizadas por las Administraciones Públicas y aumentar el acceso a los datos.



*Ilustración 13 Open Contracting Data Standard*

Este Estándar Internacional, define las siguientes etapas para los Procesos de Contratación y durante cada etapa se pueden generar unos determinados documentos:



*Ilustración 14 Fases definidas por OCDS*

A continuación, se detallan los documentos generados en cada una de las etapas.

- **Planning**, fase en la que se pueden generar los documentos “Budgets”, “Project Plans”, “Procurement Plans”, “Market Studies” y “Public hearing info”.
- **Initiation (Tender)**, fase en la que se da la opción de generar los ficheros “Tender Notes”, “Specifications”, “Line ítems”, “Values” y “Enquiries”.
- **Award**, fase en la que se permiten generar los ficheros “Details of Award”, “Bidder Information”, “Bid Evaluation”, y “Values”.

- **Contract**, fase en la que se pueden generar los documentos “Final Details”, “Signed Contract”, “Amendments” y “Values”.
- **Implementation**, fase final del proceso de contratación en el que se pueden crear los ficheros “Payments”, “Progress Updates”, “Location”, “Extensions”, “Amendments”, “Completion or Termination Info”.

Después de dar una breve introducción acerca de lo que OCDS significa y de mencionar todos los ficheros que un usuario puede encontrarse al buscar datos abiertos que siguen el estándar, se procede a explicar la estructura seguida por estos.

### 2.2.1 Estructura de OCDS

Los documentos generados en base al estándar OCDS mediante el uso del lenguaje JSON, se dividen en distintas secciones, siendo estas:

- **Sección inicial o metadatos**, muestra la información asociada a cada proceso de contratación.
- **“parties”**, contiene la información relativa a las organizaciones y entidades participantes en la licitación.
- **“buyer”**, contiene información acerca de la entidad cuyo presupuesto será empleado en el desarrollo de la actividad licitada.
- **“planning”**, contiene la información de las actividades y presupuestos para la realización de las mismas, dentro de un proceso de contratación.
- **“tender”**, contiene la información relativa al desarrollo de la licitación.
- **“awards”**, contiene los datos relativos a la adjudicación del proyecto.
- **“contracts”**, contiene información acerca de los contratos resultantes de la adjudicación de un proceso de contratación.
- **“implementation”**, contiene información relativa al estado del desarrollo del proyecto hasta la fecha de su terminación.

```
{
  "language": "es",
  "ocid": "contracting-process-id",
  "id": "release-id",
  "date": "ISO-date",
  "tag": [
    "tag-from-codelist"
  ],
  "initiationType": "tender",
  "parties": {},
  "buyer": {},
  "planning": {},
  "tender": {},
  "awards": [{}],
  "contracts": [
    {
      "implementation": {}
    }
  ]
}
```

Además de las secciones ya explicadas, el estándar OCDS tiene predefinidos una serie de campos que pueden ser empleados en cualquiera de los documentos de cada una de las fases del proceso de contratación siempre y cuando proceda. Dichos campos reutilizables se muestran a continuación:

- **“party”**, contiene la información relativa a la organización implicada en el proceso de contratación.
- **“amounts”**, contiene el importe del total presupuestado para el proceso de contratación.
- **“items”**, contiene los servicios y bienes implicados en el desarrollo del contrato, debiéndose especificar las cantidades de cada uno para no causar duplicados.
- **“period”**, contiene el intervalo de tiempo marcado por la fecha de inicio y fin.
- **“documents”**, contiene los documentos y ficheros relativos al proceso de contratación.
- **“milestones”**, contiene los hitos relativos a la finalización del proceso de contratación.

Debido a que estos campos pueden emplearse en los distintos documentos que el estándar permite generar y a que se procede a explicar cada uno de estos documentos con su estructura y elementos internos, no se muestran ejemplos de dichos campos a continuación ya que serán ejemplados en la explicación de los documentos.

#### **2.2.1.1 Paquetes de entrega de proceso de contratación**

Las distintas publicaciones de los procesos de contratación se realizan mediante paquetes de entregas, conteniendo estos todos los documentos generados para distintos procesos.

- **“uri”**, contiene la URI del paquete de entrega, dando al mismo la identificación como único en los distintos sistemas de datos globales.
- **“version”**, contiene la especificación de la versión del estándar OCDS empleado en la creación del paquete.
- **“extensions”**, contiene un resumen con las extensiones del estándar OCDS involucradas en el paquete.
- **“publishedDate”**, contiene la fecha de publicación del paquete de entrega.
- **“releases”**, contiene una lista con las entregas realizadas.
- **“publisher”**, contiene la información específica de la entidad publicadora del paquete de entrega.
- **“license”**, contiene un enlace a la especificación de la licencia que se aplica al paquete de entrega, pudiendo ser esta una de las explicadas en el apartado 2.1.1. de este documento.
- **“publicationPolicy”**, contiene un enlace a la descripción de las políticas de publicación del paquete de entrega.

```

"package":{
  "version": "1.1",
  "uri": "https://standard.open-contracting.org/latest/es/",
  "publishedDate": "2021-11-10T15:53:00Z",
  "publisher": {
    "name": "Open Data Co-operative",
    "scheme": "GB-COH",
    "uid": "9505332",
    "uri": "http://data.companieshouse.gov.uk/doc/company/09506232"
  },
  "license": "http://opendefinition.org/",
  "publicationPolicy": "https://www.open-contracting.org/resources/ocds-1-1-publication-policy-template/",
  "projects": [{}]
}

```

### 2.2.1.2 Entrega de datos de proceso de contratación

En todos los procesos de contratación, toda la información nueva referente a los mismos se publica mediante una entrega.

- **“ocid”**, u Open Contracting id es el elemento principal definido por el estándar, mediante el que se asigna un identificador único a cada nuevo proceso de contratación.
- **“id”**, contiene el identificador asignado a esta entrega de datos, debiendo este ser único en el mismo ámbito de proceso de contratación.
- **“date”**, contiene la fecha exacta en la que se publica la entrega.
- **“tag”**, contiene un código referente al tipo de documento al que se hace la entrega de datos, permitiendo así futuros filtros para buscar información.
- **“initiationType”**, contiene la especificación del código relativo al tipo de proceso de comienzo del proceso de contratación, aunque actualmente solo se admite el código ‘tender’.
- **“parties”**, como ya se ha mencionado anteriormente, este campo contiene la información acerca de las partes involucradas en el proceso de contratación.
- **“buyer”**, contiene la información de la entidad encargada de abonar el importe de los presupuestos.
- **“planning”**, contiene la descripción relativa a las fases de planeamiento del proceso.
- **“tender”**, contiene información acerca de los desarrollos de la licitación.
- **“awards”**, contiene la descripción de la entidad ganadora de la licitación y su consecuente adjudicación del proceso de contrataciones. Es importante destacar que en el proceso de contratación pueden intervenir varias adjudicaciones.
- **“contracts”**, contiene información relativa a los contratos creados y firmados a partir de la adjudicación del proceso de contratación.

- **“language”**, contiene la especificación del idioma en el que viene descrita toda la información relativa al proceso de contratación. En este caso, OCDS recomienda el uso del estándar ISO639-1.
- **“relatedProcesses”**, contiene información sobre procesos relacionados al mismo.

```
{
  "ocid": "ocds-248czf-000-00023",
  "id": "ocds-248czf-000-00023-01-tender",
  "date": "2021-04-15T09:30:00Z",
  "language": "es",
  "tag": [],
  "initiationType": "tender",
  "parties": [],
  "buyer": {},
  "tender": {}
}
```

#### 2.2.1.2.1 **“parties”**

Como se ha mencionado en la sección anterior, en este campo se referencian todas las entidades involucradas en el proceso de contratación. A continuación, se detallan todos los subcampos contenidos por este.

- **“name”**, contiene el nombre de la empresa o entidad participante.
- **“id”**, contiene la definición de un id que pueda ser referenciado desde otros apartados de la entrega.
- **“identifier”**, contiene el identificador principal de la entidad involucrada en el proceso, siendo este resultado de la extracción de la lista de identificadores relativos a la organización. Este campo contiene a su vez distintos subcampos, siendo estos:
  - **“scheme”**, contiene la referencia a la lista de identificadores de la que se extrae el id.
  - **“id”**, contiene el código identificador de la organización dentro del esquema seleccionado.
  - **“legalName”**, contiene el nombre real registrado para la entidad.
  - **“uri”**, contiene una URI relativa a la entidad.
- **“additionalIdentifiers”**, contiene un listado adjunto de distintos identificadores referentes a la entidad participante para un posible uso interno. Este campo contiene los mismos subcampos que el tipo identifier por lo que no se vuelven a explicar.
- **“address”**, contiene la dirección física de la entidad participante o una dirección en la que se pueda recibir correspondencia relativa al proceso de contratación.
  - **“streetAddress”**, contiene el nombre de la vía.
  - **“locality”**, contiene el nombre de la localidad en la que se ubica la entidad referenciada.
  - **“region”**, contiene el nombre de la región en la que se ubica la entidad referenciada.

- **“postalCode”**, contiene la especificación del código postal de la entidad.
  - **“countryName”**, contiene el nombre del país en el que se ubica la entidad referenciada.
- **“contactPoint”**, contiene la información de contacto relativa a esta entidad participante. Este tipo también contiene los siguientes subcampos:
  - **“name”**, contiene el nombre del departamento o persona de contacto de la entidad referenciada.
  - **“email”**, contiene el correo electrónico de la entidad referenciada.
  - **“telephone”**, contiene el teléfono del departamento o persona de contacto de la entidad referenciada.
  - **“faxNumber”**, contiene el número de fax del departamento o persona de contacto.
  - **“url”**, contiene el sitio web del departamento o persona de contacto de la entidad referenciada.

Cabe destacar que al ser OCDS un estándar internacional, en aquellos campos que contengan números de contacto, como puedan ser el teléfono o fax, es obligatorio insertar el prefijo de marcación internacional.

- **“roles”**, contiene la información relativa al rol que tiene la entidad en el proceso de contratación.
- **“details”**, contiene detalles adicionales acerca de la clasificación de la entidad.

```

{
  "name": "Motors ES",
  "id": "GB-GOR-XX4566",
  "identifier": {
    "scheme": "GB-GOR",
    "legalName": "Motors ES",
    "id": "XX4566",
    "uri": "https://government-
organisation.register.gov.uk/records/XX4566"
  },
  "address": {
    "postalCode": "28041",
    "countryName": "ES",
    "streetAddress": "8 Avenida de la Aviación",
    "region": "Madrid",
    "locality": "Madrid"
  },
  "contactPoint": {
    "name": "Motors Contact Departament",
    "email": "EX4566@motorses.gov.uk",
    "telephone": "+34 0123 456 7890",
    "faxNumber": "+34 0123 456 7890"
  },
  "roles": [
    "procuringEntity"
  ]
}

```

### 2.2.1.2.2 “planning”

Esta parte del documento de entrega se emplea para la descripción de antecedentes involucrados en el proceso de contratación, siendo estos detalles del presupuesto o de proyectos relacionados a este, además de la documentación relativa a la planificación e hitos de la misma.

- **“rationale”**, contiene la descripción de la toma de decisiones relativa a la adquisición.
- **“budget”**, contiene información acerca del importe presupuestado para el proceso de contratación, debido a que la información de planificación va asociada al presupuesto. Este tipo tiene los siguientes subcampos:
  - **“id”**, contiene un código identificador relativo a la entidad que abona los importes de la presupuestación.
  - **“description”**, contiene la descripción acerca del origen del presupuesto.
  - **“amount”**, contiene el importe perteneciente al proceso de contratación del presupuesto total. Este valor puede ser negativo, indicando así gastos anticipados.
  - **“project”**, contiene el nombre del proyecto mediante el que se encuentra financiado el proceso de contratación en caso de que este sea existente.
  - **“projectID”**, contiene un identificador de carácter externo relativo al proyecto del que forma parte el proceso de contratación en caso de que este sea existente.
  - **“uri”**, contiene una URI sobre los datos de presupuestación que pueda ser legible por sistemas de información.
- **“documents”**, contiene los ficheros con la información detallada de las planificaciones. Este tipo contiene los subcampos descritos a continuación:
  - **“id”**, contiene un código identificador único para el documento incluido.
  - **“documentType”**, contiene el tipo del documento incluido.
  - **“title”**, contiene el título del documento incluido.
  - **“description”**, contiene una descripción acerca del contenido del documento incluido.
  - **“url”**, contiene el enlace al documento incluido.
  - **“datePublished”**, contiene la fecha en la que el documento es publicado.
  - **“dateModified”**, contiene el registro de la fecha en la que se ha realizado una modificación en el documento.
  - **“format”**, contiene la información relativa al formato de publicación del documento.
  - **“language”**, contiene un código referente al idioma en el que se encuentra escrito el documento, siendo recomendado por el estándar OCDS el uso del estándar ISO639-1.
- **“milestones”**, contiene un listado con los hitos relativos a cada fase de la planificación. Este campo contiene a su vez los siguientes subcampos descritos:

- **“id”**, contiene un código identificador para el hito definido.
- **“title”**, contiene el título del hito descrito.
- **“type”**, contiene el tipo del hito definido.
- **“description”**, contiene la descripción relativa al hito definido.
- **“code”**, contiene un código empleado para la clasificación de los distintos eventos desarrollados en el hito definido.
- **“dueDate”**, contiene un registro con la fecha límite de terminación del hito definido.
- **“dateMet”**, contiene un registro con la fecha en la que se da por finalizado el hito.
- **“dateModified”**, contiene un registro con la fecha en la que el hito resulta modificado por diversas causas.
- **“status”**, contiene un código que describe el estado en el que se produce la modificación del hito.

```

"planning": {
  "budget": {
    "id": "1.03.1.06.02.07.26.41.1.3.2.59.76",
    "amount": {
      "amount": 30000000,
      "currency": "IDR"
    },
    "project": "Pegaso Consultance",
    "description": "APBD"
  },
  "milestones": [
    {
      "id": "1",
      "type": "preProcurement",
      "title": "Budget Plan",
      "status": "scheduled",
      "dueDate": "2021-05-19T23:29:03Z",
      "description": "New servers implementation"
    }
  ]
}

```

### 2.2.1.2.3 **“tender”**

Esta sección del documento de entrega hace referencia a la información sobre las licitaciones de aquellas organizaciones que tienen la intención de obtener contratos para realizar una determinada actividad. La sección de licitaciones puede guardar información acerca de una nueva licitación pendiente de ser recibida, o para detallar aquellas licitaciones que hayan alcanzado el estado de adjudicadas.

- **“id”**, contiene un código identificador para la licitación, pudiendo ser este equivalente al código “ocid” definido para el documento de entrega.

- **“title”**, contiene el título de la licitación, siendo recomendación por parte del estándar OCDS definir el mismo a modo de mejor resumen acerca de la licitación, para captar la atención de una mayor cantidad de entidades.
- **“description”**, contiene la descripción relativa a la licitación.
- **“status”**, contiene un código que provee la capacidad de identificar el estado en el que se encuentra la licitación. Este estado puede ser:
  - *Planning*, define un determinado proceso de contrataciones en estudio para una posible licitación futura.
  - *Planned*, define que un determinado proceso de contratación ha sido programado para una licitación, pero todavía no se ha llevado a cabo la misma.
  - *Active*, define que una licitación se encuentra en desarrollo.
  - *Cancelled*, define que una determinada licitación ha sido cancelada.
  - *Unsuccessful*, define que una licitación no ha resultado ser exitosa.
  - *Complete*, define que un proceso de licitación ha sido completado exitosamente.
  - *Retirado*, define que el proceso de licitación ha finalizado, por lo que no hay más información relativa al mismo.
- **“procuringEntity”**, contiene la información relativa a la entidad contratante que controla el proceso de licitación.
- **“items”**, contiene una lista informativa con los bienes o servicios necesarios a adquirir. Este campo contiene los siguientes subcampos:
  - **“id”**, contiene un código para hacer referencia al artículo comprado o que se tiene previsto comprar.
  - **“description”**, contiene una descripción relativa a los bienes o servicios que hayan sido comprados o que se tenga intención de comprar.
  - **“classification”**, contiene la clasificación del bien o servicio.
  - **“additionalClassification”**, contiene un listado adicional con las clasificaciones relativas al bien o servicio.
  - **“quantity”**, contiene la cantidad comprada o prevista comprar de cada bien o servicio.
  - **“unit”**, contiene la información relativa a la medida en la que se realizan las compras de los bienes o servicios.
- **“value”**, contiene la información relativa al importe máximo permitido para la contratación.
- **“minValue”**, contiene la definición del importe mínimo para la contratación.

Los campos que hacen referencia a importes o sumas de dinero contienen los subcampos **“amount”** y **“currency”**, conteniendo estos el valor y la moneda respectivamente.

- **“procurementMethod”**, contiene un código empleado para describir el procedimiento de la adquisición. Este código puede tener los valores:
  - *Open*, define que todas las entidades interesadas en la licitación pueden realizar una oferta.

- *Selective*, define que únicamente aquellas entidades que sean calificadas pueden presentar una oferta.
  - *Limited*, define que la organización a cargo de la licitación puede establecer contacto con un número definido de entidades.
  - *Direct*, define que la licitación se adjudica a una única entidad sin ningún tipo de competencia.
- **“procurementMethodRationale”**, contiene una descripción relativa al motivo de la adjudicación.
- **“mainProcurementCategory”**, contiene un código relativo a la categoría principal de la contratación, pudiendo ser esta:
  - *Goods*, define que la contratación involucra bienes y provisiones.
  - *Works*, define que la contratación involucra trabajos relacionados con obras o mantenimientos de infraestructuras.
  - *Services*, define que la contratación involucra servicios profesionales.
- **“additionalProcurementCategories”**, contiene un listado con la descripción detallada de las categorías involucradas en la contratación.
- **“awardCriteria”**, contiene información relativa a los criterios empleados en la adjudicación del contrato.
- **“awardCriteriaDetails”**, contiene información adicional relativa a los criterios de adjudicación del contrato.
- **“submissionMethod”**, contiene la definición de los métodos mediante los que se publica una determinada oferta.
- **“submissionMethodDetails”**, contiene información relativa a los detalles del método de presentación de la entrega.
- **“tenderPeriod”**, contiene la información relativa al intervalo de tiempo en el que la licitación se mantiene abierta a la entrada de ofertas.
- **“enquiryPeriod”**, contiene información relativa al intervalo de tiempo en el que las entidades interesadas en la licitación pueden realizar preguntas aclaratorias acerca de la misma.
- **“hasEnquiries”**, contiene un registro, pudiendo ser este verdadero o falso, con el fin de indicar si se han recibido consultas en el transcurso de la licitación.
- **“eligibilityCriteria”**, contiene información relativa a los criterios a seguir en la decisión de adjudicación de un contrato.
- **“awardPeriod”**, contiene la información relativa al periodo de tiempo del que dispone la organización a cargo de la licitación para publicar el ganador de la misma y adjudicar el contrato.
- **“contractPeriod”**, contiene información acerca del intervalo de tiempo estimado para la duración del contrato.
- **“numberOfTenderers”**, contiene el número total de licitaciones para un proceso de contratación.
- **“tenderers”**, contiene la información relativa a todas las entidades que hayan realizado una oferta durante la licitación.
- **“documents”**, contiene todos los ficheros y documentos asociados a la licitación.
- **“milestones”**, contiene un listado de los hitos relativos a la licitación.

- “**amendments**”, contiene la información relativa a la enmienda o cambio en las condiciones de la licitación y por consiguiente su nueva publicación.

```

"tender": {
  "id": "ocds-213czf-000-00015-01-tender",
  "title": "Planificación de mejoras para líneas de bicicletas",
  "description": "Licitaciones acerca de la mejora de las vías de
circulación de bicicletas dentro de la ciudad",
  "status": "active",
  "items": [
    {
      "id": "0015",
      "description": "string",
      "classification": {
        "scheme": "CPV",
        "id": "45233130",
        "description": "Construcciones para vías de servicio públicas.",
        "uri": "http://cpv.data.ac.uk/code-45233130"
      },
      "additionalClassifications": [
        {
          "scheme": "CPV",
          "id": "45233162-2",
          "description": "Contrucción de vías para bicicletas.",
          "uri": "http://cpv.data.ac.uk/code-45233162.html"
        }
      ],
      "quantity": 8,
      "unit": {
        "name": "Millas",
        "value": {
          "amount": 120000,
          "currency": "GBP"
        }
      }
    }
  ],
  "minValue": {
    "amount": 600000,
    "currency": "GBP"
  },
  "value": {
    "amount": 1100000,
    "currency": "GBP"
  },
  "procurementMethod": "open",
  "procurementMethodRationale": "Una licitación abierta se requiere por
parte del gobierno de la UE",
  "awardCriteria": "bestProposal",
  "awardCriteriaDetails": "La mejor oferta en cuanto a importe recibirá
la adjudicación de la oferta.",
  "submissionMethod": [

```

```

        "electronicSubmission"
    ],
    "submissionMethodDetails": "Enviar por el portal al enlace:
http://example.com/submissions/ocds-213czf-000-00001-01/",
    "enquiryPeriod": {
        "startDate": "2021-04-01T09:00:00Z",
        "endDate": "2021-04-14T17:30:00Z"
    },
    },
    "hasEnquiries": false,
    "tenderPeriod": {
        "startDate": "2021-02-01T09:00:00Z",
        "endDate": "2021-04-01T18:00:00Z"
    },
    },
    "awardPeriod": {
        "startDate": "2021-06-01T00:00:00Z",
        "endDate": "2021-08-01T23:59:59Z"
    },
    },
    "procuringEntity": {
        "id": "GB-LAC-E09000003",
        "name": "London Borough of Barnet"
    },
    },
    "documents": [
    {
        "id": "0005",
        "documentType": "notice",
        "title": "Tender Notice",
        "description": "Licitación oficial publicada.",
        "url": "http://example.com/tender-notices/ocds-213czf-000-00001-01.html",
        "datePublished": "2021-03-01T09:00:00Z",
        "format": "text/html",
        "language": "es"
    }
    ]
}

```

#### 2.2.1.2.4 “awards”

En esta sección se describen los datos relativos a las adjudicaciones, paso subsiguiente al proceso de licitación, mediante el que se anuncia el ganador de la licitación.

- “**id**”, contiene el código identificador único relativo a la adjudicación.
- “**title**”, contiene el título descriptivo de la adjudicación.
- “**description**”, contiene la descripción relativa a la adjudicación.
- “**status**”, contiene el registro que define el estado en el que se encuentra la adjudicación, pudiendo ser este:
  - *Pending*, define un contrato propuesto que todavía no ha entrado en vigor.
  - *Active*, define un contrato firmado por la totalidad de las entidades involucradas en el mismo y que por lo tanto está en desarrollo.

- *Cancelled*, define un contrato cancelado antes de que pudiera ser firmado por alguna de las entidades involucradas.
- *Terminated*, define un contrato cerca de su finalización, ya sea por motivos de terminación de proyecto o por imposibilidad de terminación.
- **“date”**, contiene registrada la fecha en la que el contrato de adjudica a la entidad ganadora.
- **“value”**, contiene el registro del importe total relativo a la adjudicación del contrato.
- **“suppliers”**, contiene información relativa a los proveedores adjudicados al contrato.
- **“items”**, contiene un listado con la información relativa a los bienes y servicios adjudicados al contrato.
- **“contractPeriod”**, contiene la información acerca del intervalo de tiempo por el que se tiene adjudicado el contrato.
- **“documents”**, contiene adjuntos los ficheros y documentos relativos a la adjudicación.
- **“amendments”**, contiene la información relativa a la enmienda o cambio en las condiciones de la adjudicación y por consiguiente su nueva publicación.

```

"awards": [
{
  "id": "ocds-213czf-000-00015-01-tender",
  "title": "Planificación de mejoras para líneas de bicibletas",
  "description": "Licitaciones acerca de las mejoras de las vías de
circulación de bicicletas dentro de la ciudad",
  "status": "pending",
  "date": "2010-05-10T09:30:00Z",
  "value": {
    "amount": 11000000,
    "currency": "GBP"
  },
  "suppliers": [
    {
      "id": "GB-COH-1234561124",
      "name": "Bicicletas Decathlon"
    }
  ],
  "items": [
    {
      "id": "0015",
      "description": "string",
      "classification": {},
      "additionalClassifications": [],
      "quantity": 8,
      "unit": {
        "name": "Millas",
        "value": {
          "amount": 137000,
          "currency": "GBP"
        }
      }
    }
  ]
}

```

```

    }
  },
  ],
  "contractPeriod": {
    "startDate": "2021-07-01T00:00:00Z",
    "endDate": "2021-08-01T23:59:00Z"
  },
  documents: [
    {
      "id": "0006",
      "documentType": "notice",
      "title": "Award notice",
      "description": "Award of contract to build new cycle lanes in the
      centre of town to AnyCorp Ltd.",
      "url": "http://example.com/tender-notices/ocds-213czf-000-00001-
      04.html",
      "datePublished": "2021-05-10T09:30:00Z",
      "format": "text/html",
      "language": "en"
    }
  ]
]

```

#### 2.2.1.2.5 “contracts”

- “id”, contiene el código identificador único del contrato.
- “awardID”, contiene el código identificador de la adjudicación relativa del contrato.
- “title”, contiene el título descriptivo del contrato.
- “description”, contiene la información relativa a la descripción del contrato.
- “status”, especifica el estado en el que se encuentra el contrato.
- “period”, contiene el registro de las fechas de comienzo y finalización del contrato.
- “value”, contiene el importe total para el contrato.
- “items”, contiene la información relativa a los bienes y servicios del contrato, con la recomendación por parte del estándar OCDS de no repetirlos en caso de ser iguales que los descritos en la adjudicación.
- “dateSigned”, contiene el registro con la fecha en la que se ha realizado la firma del contrato por la totalidad de las entidades involucradas.
- “documents”, contiene los ficheros y documentos relativos al contrato.
- “implementation”, contiene información relativa al estado del desarrollo del contrato.
- “relatedProcesses”, contiene información sobre procesos de contratación relacionados al contrato.
- “milestones”, contiene un listado con los hitos establecidos para el contrato.
- “amendments”, contiene la información relativa a la enmienda o cambio en las condiciones del contrato y por consiguiente una nueva publicación de la entrega o una nueva licitación.

```

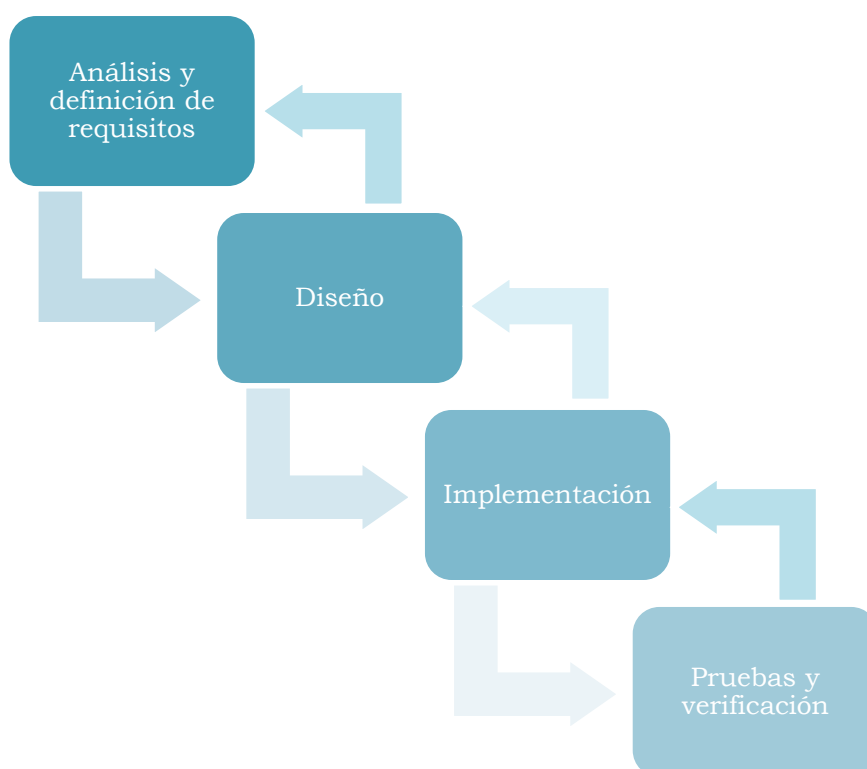
"contracts": [
  {
    "id": "ocds-213czf-000-00015-contract-01",
    "awardID": "ocds-213czf-000-00015-award-01",
    "title": "Contrato de mejoras para líneas de bicicletas.",
    "description": "Un contrato ha sido firmado para la construcción
de nuevas vías para bicicletas en la ciudad.",
    "status": "active",
    "period": {
      "startDate": "2021-07-01T00:00:00Z",
      "endDate": "2021-08-01T23:59:00Z"
    },
    "value": {
      "amount": 11000000,
      "currency": "GBP"
    },
    "items": [
      {
        "id": "0001",
        "description": "string",
        "classification": {
          "scheme": "CPV",
          "id": "45231230",
          "description": "Construcciones para vías de servicio
públicas.",
          "uri": "http://cpv.data.ac.uk/code-45233130"
        },
        "quantity": 8,
        "unit": {
          "name": "Millas",
          "value": {
            "amount": 137000,
            "currency": "GBP"
          }
        }
      }
    ]
  },
  "dateSigned": "2021-06-10T14:23:12Z",
  "documents": [
    {
      "id": "0015",
      "documentType": "contractSigned",
      "title": "Contrato firmado",
      "description": "El contrato firmado para la construcción
de vías para bicicletas",
      "url": "http://example.com/contracts/ocds-213czf-000-
00001",
      "datePublished": "2015-06-10T16:43:12Z",
      "format": "application/pdf",
      "language": "es"
    }
  ]
}
]

```

### 3 Desarrollo

En este capítulo se procede a explicar las tecnologías, programas y librerías empleadas en el desarrollo del software de extracción de los datos de la Plataforma de Contratación Pública del Estado. El objetivo principal es otorgar a los lectores una visión general de los pasos establecidos para la transformación de los datos abiertos, los resultados obtenidos, y las posibles mejoras que pueda tener el software. Cabe destacar que la totalidad del código está situado en el repositorio personal de GitHub del autor del proyecto, y se encuentra totalmente disponible para su visualización o reutilización.

La idea con la que comienza este proyecto de fin de Grado es el análisis y categorización de los datos abiertos de la Plataforma de Contratación Pública del Estado. Lo que se busca con esto es dar una visión más asequible y organizada de los datos abiertos para su posterior análisis y uso para diferentes fines. Para ello se propone un modelo de desarrollo en Cascada con el que se busca la comprobación de los objetivos detallados en la fase anterior con cada nueva iteración. Este modelo se muestra en el siguiente gráfico y a continuación se procede a explicar en detalle todos los pasos definidos en cada una de las fases de desarrollo.



*Ilustración 15 Modelo de Desarrollo en Cascada*

## **3.1 Análisis**

En esta sección se explica de forma detallada la planificación del desarrollo, el análisis previo de los ficheros a tratar, la definición de los requisitos que debe de cumplir el software final y restricciones o limitaciones que el software pueda tener, facilitando así las tareas de desarrollo y pruebas del mismo.

### **3.1.1 Estructura de los ficheros de Contratación Pública del Estado**

Para poder definir correctamente los requisitos que debe de tener el software, es necesario un previo análisis de los datos abiertos elegidos para este trabajo. Estos datos se encuentran disponibles en el Portal Institucional del Ministerio de Hacienda [12], en forma de ficheros .zip agrupados por meses, que a su vez están organizados en su interior mediante ficheros .atom. Dichos ficheros están escritos en lenguaje XML, y cada proyecto de contratación pública viene definido dentro de la etiqueta **<entry>**. En el supuesto de acceder a uno de estos ficheros, se debe tener de referencia esta etiqueta, ya que tiene como objetivo separar los distintos contratos de los procesos de Licitación Pública.

Al comienzo del fichero .atom que contiene los contratos agrupados, se muestran los datos relativos al autor del fichero, que en la correcta generación de los archivos .csv carecen de valor debido que todos los ficheros son publicados por la Plataforma de Contratación Pública del Estado y, además, los contratos internamente describen las fechas de creación y publicación. Con la intención de no repetir información y no aumentar la longitud del proyecto, los lectores pueden encontrar la descripción de la estructura interna de cada contrato en la sección 2.1.2. de este documento, debido a que se ha utilizado un extracto de los datos abiertos para los ejemplos.

Una vez que se tiene realizado un análisis previo de la estructura interna de los ficheros de Contratación Pública, se puede dar paso a la definición de los requisitos necesarios para el correcto desarrollo del software.

### **3.1.2 Restricciones y limitaciones de software**

El software está pensado y desarrollado para tratar datos de Contratación Pública del Estado. En este sentido, y dado que los datos abiertos están publicados en formato XML, únicamente se pueden leer ficheros en este formato, ya que del contrario el flujo y extracción de datos generaría numerosos errores que no han sido tratados. Siendo esto así, y aunque en un principio se realizó un diseño capaz de leer cualquier estructura XML, este software no puede llevar a cabo dicha labor por errores de parseo de documentos generados por la API DOM más adelante explicada. Es por esto, que se recomienda el uso exclusivo de los ficheros de la Plataforma Oficial del Estado.

### 3.1.3 Requisitos de software

En esta subsección se definen los requisitos funcionales y de calidad de software junto con una descripción detallada de cada uno de estos.

Requisitos	Descripción
R1	Se debe permitir la lectura de ficheros con cualquier extensión y desde cualquier ruta de memoria.
R2	Debe poder generar los ficheros en cualquier ruta de memoria.
R3	Debe generar un fichero con tantos registros como contratos tenga el fichero original.
R4	Debe mostrar los nombres de los nodos extraídos como nombres de cada columna del fichero correspondiente.
R5	Debe generar tantos ficheros como “id” o “code” (sin repetir el nodo padre de los mismos) contenga el contrato.
R6	Todos los registros de los distintos ficheros deben de tener un identificador de referencia al contrato del que provienen.
R7	El nombre del fichero debe de ser igual que el nodo padre del id que se encuentra.
R8	Cada nuevo contrato extraído debe de ser guardado en los ficheros ya existentes.
R9	No se deben repetir datos de un fichero a otro.
R10	Tanto los datos como los nombres de las columnas deben ser separados por comas para generar correctamente los ficheros .csv.
R11	Se debe de poder elegir la extensión de los ficheros generados por tareas de análisis de resultados.
R12	Se debe generar un nuevo registro por cada nodo leído.
R13	Posibilitar la ejecución del código desde cualquier máquina.
R14	Permitir la edición de un fichero de configuración.

*Tabla 4 Descripción de Requisitos*

## 3.2 Diseño

En un primer intento de generar ficheros .csv a partir de los datos de Contratación Pública disponibles, se ha optado por generar un único archivo, cuyos nombres de tablas fueran los nodos de cada <entry> sin duplicados. Para ello se toma la decisión de desarrollar un bucle recursivo que fuera capaz de devolver los nombres de los nodos independientemente de la complejidad y extensión que pudiera tener cada <entry>, y el valor se inserta en un atributo que almacena una lista con dichos nombres. Esta lista una vez procesada se inserta en lo que posteriormente es la cadena separada por comas en la primera fila del archivo .csv. Cabe mencionar que con el término procesado se hace referencia a la comparación de los elementos del primer <entry> sin repeticiones con los elementos nuevos de cada <entry> que se procesa, con el objetivo de tener una lista de todos los nombres del fichero .xml sin duplicados. Una vez se ha recorrido todo el fichero y las columnas del archivo han sido definidas, se procede a extraer los datos de cada nodo, y para ello se emplea una función de la librería DOM, explicada en las siguientes secciones, que tiene la peculiaridad de ser capaz de devolver todos los datos contenidos en cada nodo, creando así errores relacionados con la duplicación de datos.

Es por esto por lo que se ha realizado un análisis de la función recursiva para limitar la extracción de los datos de cada Licitación o <entry> únicamente a aquellos nodos que no dispongan de nodos hijos. Una vez solucionada esta casuística, los datos obtenidos se almacenan en una lista que posteriormente se inserta en el archivo a continuación de la primera línea que contiene las columnas. Al hacer esto afloran los errores de este primer abordaje de la extracción, debido a que dentro de cada <entry> se repiten nodos, algunos contratos tienen una longitud mayor o menor que el primer contrato insertado en el archivo .csv, por lo que las columnas se muestran desplazadas o en algunos casos hay más datos que columnas y se llega a la conclusión de que no es posible generar un único .csv.

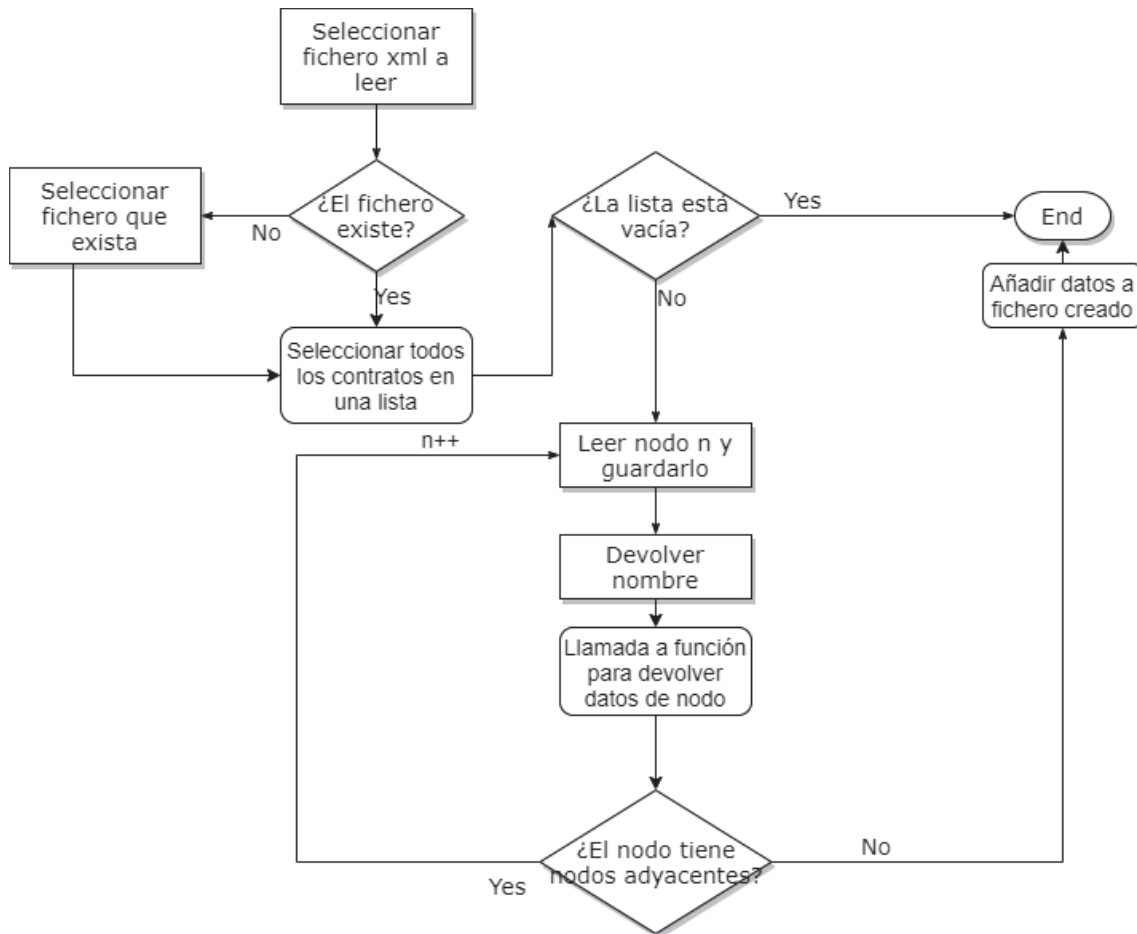
En un segundo análisis del generado de ficheros .csv se opta por generar varios ficheros por cada <entry>, dependiendo de los nodos que contenga. Para identificar los nodos que pueden generar un fichero nuevo, se lleva a cabo el análisis del documento, y se establece que cada vez que se encuentre la secuencia de caracteres “id” o “code”, se debe de generar un nuevo fichero .csv. Esta idea surge a raíz de que los ficheros .xml no son planares, sino que son dinámicos, y por lo tanto no pueden ser transferidos manteniendo la organización de los datos en un único fichero .csv. Lo que se quiere conseguir de esta forma es poder obtener los datos de tal forma que puedan ser “limpiados” desde el punto de vista de tratamiento de datos.

Se espera obtener un número amplio de ficheros generados. Todos estos ficheros deberán contener las columnas relativas a cada nodo que se define como nuevo archivo, ya que se estima que los nodos van a tener el mismo número de subnodos hijos, es decir, que, al avanzar en el procesado del fichero de Licitación Pública, por cada <entry> recorrido se espera obtener el nodo <Address> con sus respectivos atributos como pueden ser “id” o “Country”. Por tanto, es necesario guardar el contrato del que se extraen los mismos nodos, y

para ello se toma la decisión de almacenar el id del <entry> del que se extrae el nodo en cada nuevo registro que se inserta en el fichero, en concreto un subString del id completo de cada contrato, que puede ser utilizado como identificador, no único, para los registros.

### 3.2.1 Diagramas de flujo de software

Para el desarrollo del software, y posteriormente al análisis de los ficheros, se pretende dividir el flujo del programa en dos procedimientos, uno que sea capaz de devolver la información de aquellos nodos que dan información a modo de resumen de los contratos, y otro que devuelve los datos concretos contenidos en todos los nodos. Por ello, se definen los siguientes diagramas de flujo.



*Ilustración 16 Diagrama de flujo para datos de resumen de contratos*

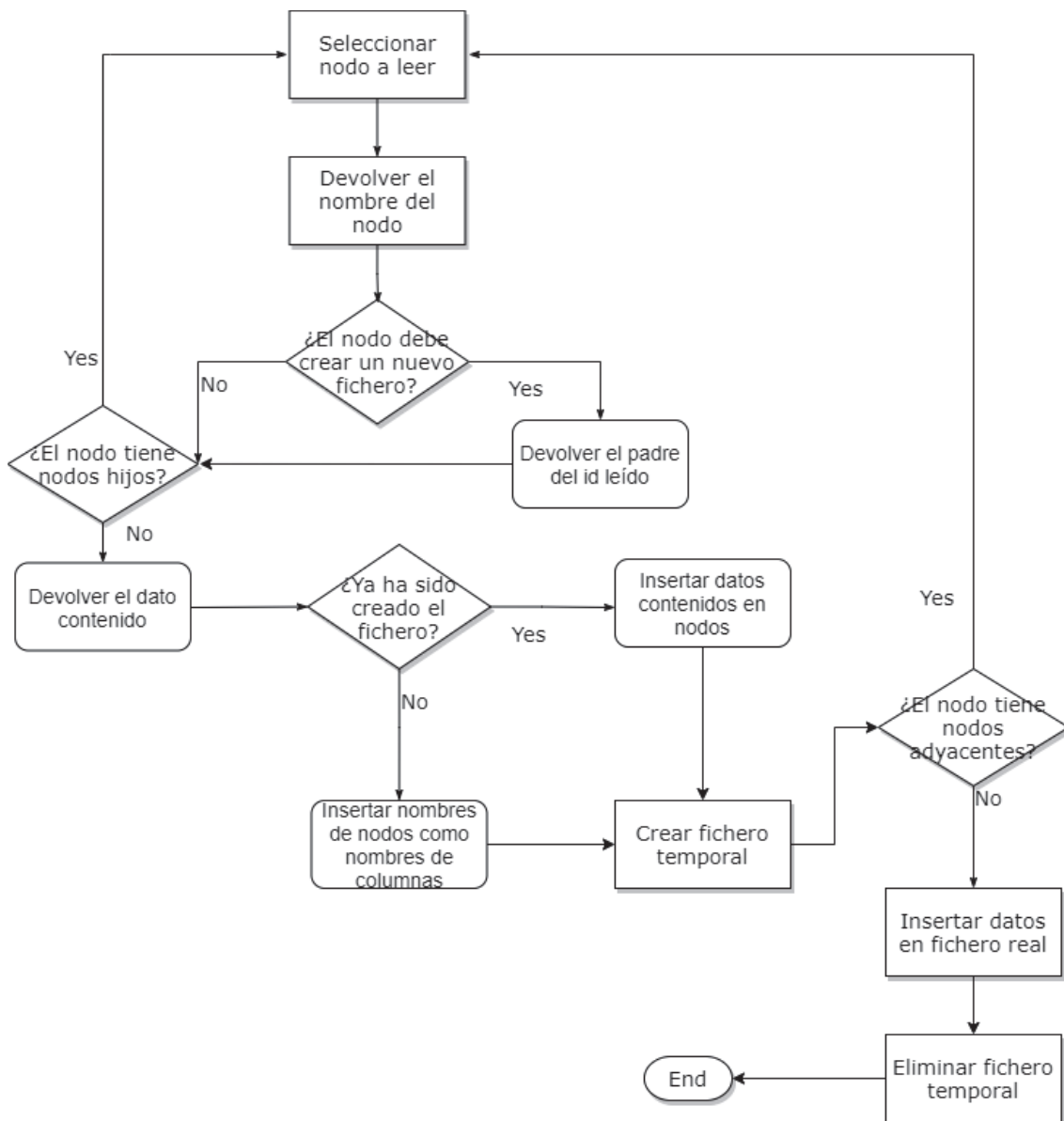


Ilustración 17 Diagrama de flujo para datos de contratos

Cabe destacar que el procedimiento mostrado con el primer diagrama de flujo debe realizar llamadas al procedimiento explicado en el segundo diagrama, debido a que este último contiene la funcionalidad de leer y extraer los nodos y datos separados de cada contrato procesado gracias al bucle recursivo mencionado en la introducción de esta sección. En concreto, el primer procedimiento ejecuta la orden de tratamiento de cada contrato según se vaya avanzando en la lectura del fichero de datos abiertos relativos a la Contratación Pública.

## 3.2.2 Herramientas y lenguaje de desarrollo

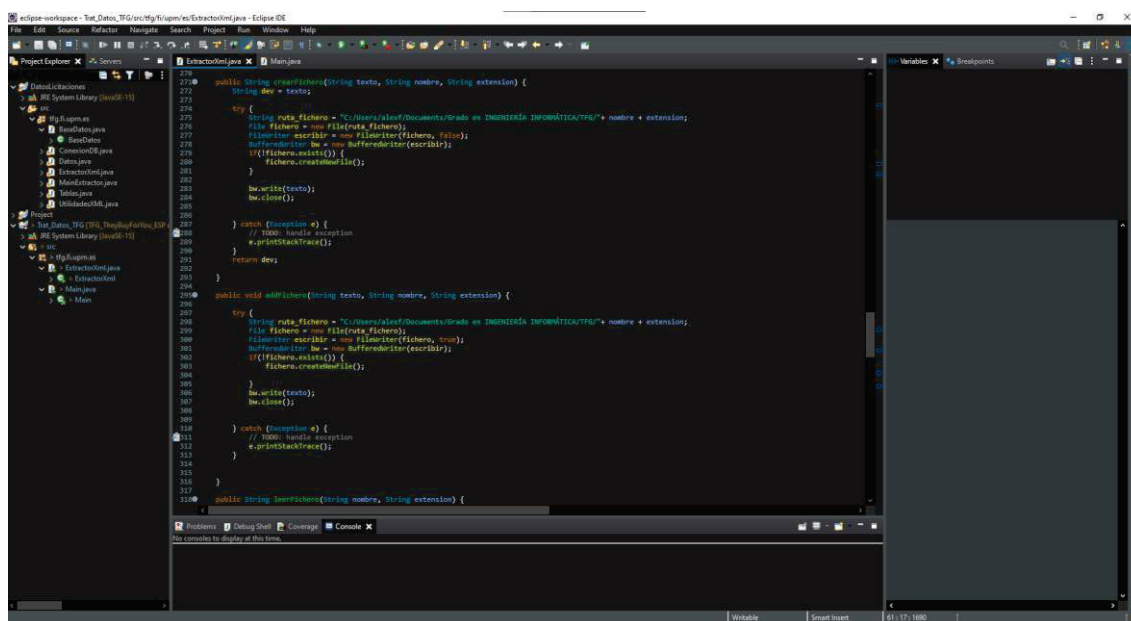
Para el desarrollo del software se propone el uso principal de una plataforma de desarrollo y dos editores de código.

### 3.2.2.1 Eclipse

Eclipse es una plataforma de desarrollo software para entornos de desarrollo integrado (IDE), y es popularmente conocida por los usuarios para realizar desarrollos en lenguaje Java, gracias a su IDE denominado Java Development Toolkit (JDT) y su compilador Eclipse Compiler for Java (ECJ). Este software ha sido creado por IBM, pero actualmente se encuentra respaldado por la Fundación Eclipse. Se puede entender a Eclipse como una aplicación que proporciona herramientas para desarrollar, compilar y ejecutar software, realizar tareas de depuración de código, y que además permite la interacción y gestión de entornos de trabajo. A continuación, se procede a explicar brevemente los principales componentes para dar un mayor entendimiento a los lectores.

- **Ventanas, editor y perspectivas.** En el caso de Eclipse, el uso de la aplicación y el sistema de trabajo se rige por la distribución de perspectivas, siendo esto una configuración predefinida, dividiendo así la pantalla de la aplicación en varias ventanas y dejando como principal la ventana de edición de código.
- **Creación y gestión de proyectos.** Gracias a la distribución mediante ventanas y a que los proyectos se crean en paquetes y carpetas, la aplicación permite la gestión de los ficheros dentro de las carpetas ubicadas en la memoria de la máquina empleada para realizar el desarrollo.
- **Depuración de código.** Eclipse contiene un depurador que permite a los usuarios ejecutar su código línea a línea y ver el contenido de las variables, gracias una vez más a la organización por perspectivas, en concreto a las ventanas abiertas para el proceso de depuración, que muestran todas las variables definidas con sus respectivas posiciones de memoria y los valores que toman en cada momento.

Para finalizar la breve explicación, cabe destacar que Eclipse permite el enlazado del proyecto desarrollado con el controlador de versiones GitHub, por lo que permite tener el código fuente actualizado desde cualquier ordenador desde el que se trabaje.



*Ilustración 18 IDE de desarrollo Eclipse*

### 3.2.2.2 Sublime Text y Notepad++

Sublime Text es un visor y editor de texto y código fuente que soporta la gran mayoría de lenguajes de programación y lenguajes de etiquetas. Originalmente fue desarrollado como la continuación del clásico editor de texto Vim, aunque actualmente ha construido su propia identidad y la única equivalencia que se encuentra con Vim es un modo de edición de texto vintage. Este editor de texto se ha empleado para el análisis de los ficheros de datos referentes a la Plataforma de Contratación Pública del Estado.

NotePad++ es un visor y editor de texto y código fuente libre desarrollado para Windows inicialmente. Al igual que Sublime Text, admite multitud de lenguajes y formatos, y en este proyecto se ha empleado para el análisis de los ficheros .csv creados a partir de la extracción de los datos de la Plataforma de Contratación Pública del Estado.



*Ilustración 16 Sublime Text*



*Ilustración 17 Notepad++*

### 3.2.2.3 Java

Java es un lenguaje de programación y a su vez un plataforma que permite el funcionamiento de numerosos portales y aplicaciones web. Su primera publicación se realizó en el año 1995, y en el año 2012 se convierte en uno de los lenguajes más populares para el desarrollo de software y en especial para el desarrollo de aplicaciones web del tipo cliente-servidor.

Aunque es un lenguaje de programación nuevo e integrado en la plataforma de Java (entorno de desarrollo creado por Sun Microsystems capaz de ejecutar aplicaciones, compuesto por un máquina virtual y bibliotecas de Java), la mayor parte del lenguaje tiene sus raíces en C y C++, aunque Java sea considerado un lenguaje de mayor nivel. Java es un lenguaje de programación que sigue la metodología de programación orientada a objetos, que lo que pretende es tener en un mismo conjunto los datos y el código fuente, dando una visión de estos como un paquete contenedor del funcionamiento del código fuente y el estado o situación de los datos.



*Ilustración 18 Java*

## 3.3 Implementación de la Extracción de los Datos

En esta subsección se procede a explicar en detalle la implementación realizada, las librerías empleadas en la lectura de los ficheros con formato XML, la estructura del proyecto y la organización de las versiones de desarrollo mediante la plataforma GitHub.

Para dar una visión detalla a los lectores del flujo de generación de los ficheros con los datos extraídos, se divide el mismo en varias etapas definidas a continuación:

1. **Selección del fichero.** En esta fase se debe de seleccionar el fichero con formato XML del que se desean extraer los datos. Para ello basta con insertar la ruta de memoria del fichero.
2. **Parseo del fichero.** Como el desarrollo del código se realiza en lenguaje Java, para la correcta lectura del fichero se requiere la importación en la clase ExtractorXml, de las librerías y APIs:
  - a. `javax.xml.parsers.DocumentBuilder;`
  - b. `javax.xml.parsers.DocumentBuilderFactory;`
  - c. `org.w3c.dom.NodeList;`
  - d. `org.w3c.dom.Document;`

e. `org.w3c.dom.Node;`

El motivo de esto es que Java no es capaz de comprender la estructura en árbol definida en el fichero de partida, por lo que es necesario realizar un parseo del fichero para su posterior conversión al modelo definido por el estándar w3c Document explicado en esta sección.

3. **Identificación de todos los contratos.** Una vez se tiene parseado el fichero, se proceden a organizar en una lista todos los contratos, ya que lo que se pretende, según se ha explicado en la sección de Diseño, es generar un nuevo registro por cada contrato existente. Para realiza una búsqueda mediante el nodo `<entry>`.
4. **Extracción de los datos de resumen.** Cada contrato contiene unos primeros nodos relativos a los datos de resumen, fecha de publicación, título e id del contrato, en base a la estructura ya explicada en el apartado CODICE de este documento. Para extraer estos datos, se selecciona el primer contrato de la lista definida en el paso anterior, y se excluyen aquellos nodos vacíos que la API DOM genera por defecto. En este paso también se extrae el identificador que sirve a modo de identificador de los registros generados por cada contrato en los distintos ficheros. Como estos primeros nodos no tienen nodos hijos, realizan la llamada a la función recursiva en la que se itera una sola vez por cada nodo para devolver los datos contenidos.
5. **Extracción de los datos internos de cada nodo.** Una vez han sido extraídos y guardados en un atributo los valores de los nombres y datos leídos de los nodos resumen, se procede a leer cada uno de los nodos del contrato. En la fase de diseño se define que se debe generar un nuevo fichero .csv por cada nodo que contenga la secuencia de caracteres “id” o “code”. Esta especificación ha sido muy laboriosa de cumplir, ya que, para generar el nuevo fichero, se emplea el nombre del nodo padre máximo distinto del nodo usado para definir los contratos. Para una mejor visualización de esto, se muestra el siguiente ejemplo:

```
<entry>
  <cac-place-ext:ContractFolderStatus>
    <cbc:ContractFolderID>4270012027200</cbc:ContractFolderID>
    <cbc-place-ext:ContractFolderStatusCode languageID="es"
      listURI="https://contrataciondelestado.es/codice/cl/2.04/Syndica
      tionContractFolderStatusCode-2.04.gc">RES</cbc-place-
      ext:ContractFolderStatusCode>
```

El programa recursivo ha sido creado de tal forma que sea consciente en todo momento cuando debe y cuando no generar un nuevo fichero. Para ello, cuando encuentra un nodo que contenga la secuencia de caracteres “id” , siendo en este ejemplo el nodo `<cbc:ContractFolderID>`, lanza la creación de un nuevo fichero vacío con el nombre de `ContractFolderStatus`, ya que se realiza un tratamiento de los nombres de los nodos para eliminar los caracteres no deseados. El programa realiza una búsqueda del padre del nodo que lanza la creación del archivo vacío sin que llegue a salirse del contrato, es decir, siempre en los límites definidos por `<entry>`. Esto último se mantiene gracias a un contador que

almacena el número de niveles que se avanza hacia el interior de los nodos para saber así en que nodo padre debe parar.

6. **Inserción de los datos dentro del fichero temporal.** A partir del paso anterior, el programa inserta en el archivo temporal `ContractFolderStatus_1.csv` todos los nodos hijos del nodo padre empleado para dar nombre al fichero vacío. El archivo temporal en el que se insertan los datos de los nodos únicamente cambia cuando se encuentra un nuevo nodo que contiene la secuencia de caracteres “id” o “code”.

En el programa, para la correcta estructura del fichero resultante, se definen varios modos de inserción de los datos en los ficheros temporales:

- a. Inserción de datos pertenecientes al primer contrato.
- b. Inserción de datos pertenecientes al resto de contratos.
- c. Inserción de datos cuyo contrato sea distinto del primero tratado y que no tengan un fichero vacío creado.

La explicación de estos modos de inserción es la siguiente. En la lectura del primer contrato, los ficheros están vacíos, por lo que no contienen ni columnas ni registros, y por lo tanto, es necesario rellenar las columnas con los nombres de los nodos correspondientes a cada fichero, además de los registros. En la lectura del segundo contrato y sucesivos, las columnas ya están definidas, por lo que no es necesario volver a insertar la información relativa a las misma, sino que basta con insertar los registros. El último caso corresponde a aquellos contratos distintos del primer contrato leído, que además contiene nodos nuevos a este, y que por lo tanto crean nuevos ficheros no existentes en los que se debe de insertar tanto las columnas como los registros.

7. **Inserción de los datos en el fichero vacío definitivo.** Este paso es el último, en el que la información contenida en los ficheros temporales se traspassa al fichero definitivo, y los ficheros temporales se borran, dando así resultado a los ficheros `.csv`.

	A	B	C
1	id	ContractFolderID	ContractFolderStatusCode
2	789966	4270012027200	RES
3	742525	2103/0109	RES
4	789906	2103/0133	RES
5	799051	204152041512012100	RES
6	794697	2012/CTR/0900402	RES
7	790072	04401-01/2012	RES
8	742446	2021620120112	RES
9	789687	202162012016900	RES
10	799052	SE(DT)-4100	RES
11	799053	1438/11	RES
12	783360	07-15-2012	RES
13	740355	2012/00065	RES

*Ilustración 19 Ejemplo de fichero ContractFolderStatus.csv generado*

### 3.3.1 API DOM

DOM, o Modelo de Objetos del Documento, es el término empleado para describir la API que posibilita la lectura, escritura o modificación, y en este proyecto en concreto, la extracción de datos de documentos XML mediante lenguajes de programación. La estructura del Modelo es semejante a la estructura del documento que se parsea, y el adjetivo semejante es el término correcto a usar en lugar de iguales debido a que al parsear el fichero deseado, la API DOM modela una estructura en la que se añaden espacios en blanco. Omitiendo este pequeño detalle, se puede entender que mediante la API se obtiene una estructura de objetos y nodos con sus respectivas propiedades. Para ilustrar esto se muestra un ejemplo tomado de W3C Recommendation.

Dadas las siguientes líneas de código, se modela la estructura sin espacios en blanco mostradas a continuación:

```
<table>
  <tbody>
    <tr>
      <td>Shady Grove</td>
      <td>Aeolian</td>
    </tr>
    <tr>
      <td>Sobre el río, Charlie</td>
      <td>Dorian</td>
    </tr>
  </tbody>
</table>
```

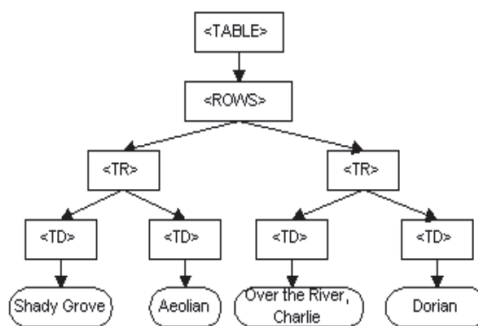


Ilustración 20 Estructura DOM [14]

Una vez se ha mostrado una vista general de la API, se procede a explicar los métodos claves en la extracción de los datos del documento XML. Entre estos métodos, cabe destacar:

- **Node.getNodeName()**, permite devolver el nombre del nodo que se está accediendo.
- **Node.getTextContent()**, permite devolver la información contenida en el nodo sobre el que se invoca, sin embargo, este método devuelve también la información contenida en los nodos hijos del nodo principal, lo que a la hora de extraer los datos ha supuesto un problema de integridad de datos, ya que puede ocasionar múltiples repeticiones de los mismo si no es invocado correctamente.
- **Node.getFirstChild()**, permite situarse en el primer nodo hijo del nodo sobre el que se invoca.
- **Node.getNextSibling()**, permite situarse en el nodo siguiente al nodo sobre el que se invoca, siguiendo la misma línea jerárquica.
- **Node.getParentNode()**, permite situarse en el nodo antecesor al nodo sobre el que se invoca, es decir una jerarquía por encima.

### 3.3.2 Estructura de ficheros de Código Fuente

En esta subsección se muestra la estructura de los ficheros del Código Fuente. Es necesario entender que como toda la lectura y extracción de los datos se realiza mediante una función recursiva, no se ha necesitado implementar un gran número de clases. Además, dada la complejidad del código, se ha dedicado una gran parte del tiempo de este trabajo a la mejora del tiempo de ejecución del software desarrollado.

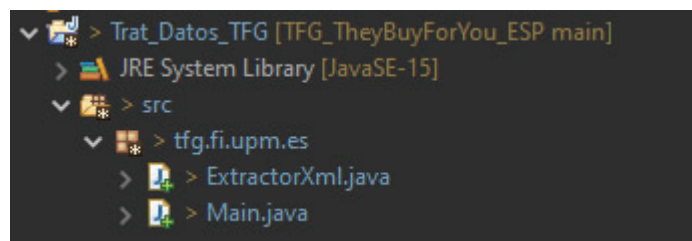


Ilustración 21 Estructura de ficheros

La estructura se compone de un paquete que contiene dos clases. La clase `ExtractorXml` define los métodos encargados del parseo de los ficheros de datos de Contratación Pública, extracción de los datos y generación de los ficheros. La clase `Main` se encarga de recibir la ruta del fichero que se desea tratar y ejecuta el proceso de generado de los ficheros `.csv`.

Adicionalmente, se ha configurado un fichero de configuración, para posibilitar la ejecución del programa desde cualquier máquina o entorno. La estructura del fichero de configuración "*config.properties*" es muy simple, ya que para la ejecución del programa se necesita únicamente la ruta del fichero que se va a procesar, y la ruta de la carpeta en la que se van a guardar los archivos `.csv` generados.

```
#definir la ruta del fichero a leer
file_path=file:///C:/Users/alex/Documents/Grado_en_Ing_Inf/TFG/licitacionesPerfilesContratanteCompleto3_
#definir la ruta de la carpeta para guardar ficheros generados
folder_path=C:/Users/alex/Documents/Grado_en_Ing_Inf/TFG/Intento_2_Completo/
```

Ilustración 22 Fichero de configuración

A raíz de este requerimiento, también se ha generado un fichero `.jar`, que contiene los ficheros mostrados en la Ilustración 21, y por lo tanto es el encargado del procesado de los ficheros.

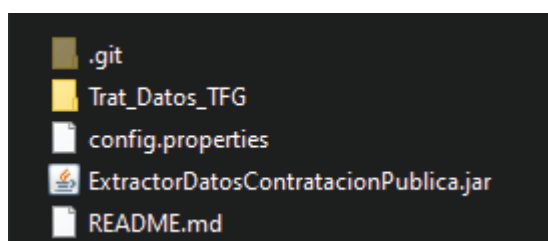


Ilustración 23 Fichero de ejecución

Para ejecutar el programa, es necesario el uso de línea de comandos, y de tener instalado Java en la máquina para poder invocar a la máquina virtual del mismo. El código de ejecución del programa, habiendo definido previamente las rutas necesarias en el fichero de configuración, es el siguiente:

```
java -jar ExtractorDatosContratacionPublica.jar
```

### 3.3.3 Control de versiones

Para tener un control de versiones, dar la posibilidad de visualizar el código fuente de este desarrollo y acceder al software para su descarga y uso, se opta por crear un repositorio en la plataforma GitHub. Este repositorio está enlazado directamente con los ficheros del entorno de desarrollo de Eclipse, tal y como se puede ver en la Ilustración 21, ([TFG\_TheyBuyForYou\_ESP main]).

Para el control de versiones, únicamente se han realizado actualizaciones de los ficheros cuando los cambios han sido significativos o cuando se han corregido errores.

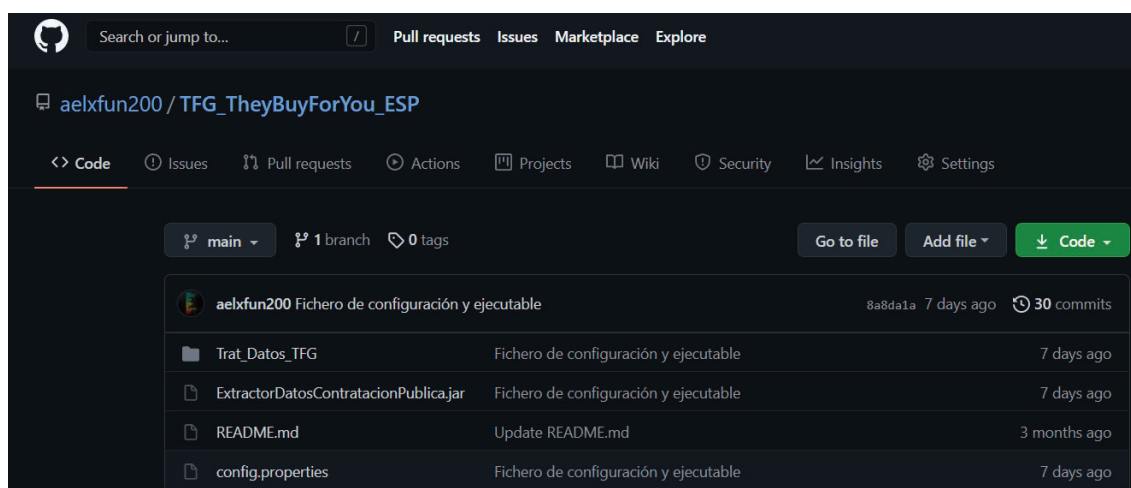


Ilustración 24 Control de versiones GitHub

En la imagen se puede ver la carpeta contenedora del código fuente del software desarrollado, el fichero de propiedades que permite la ejecución del software en cualquier sistema o máquina, y el propio archivo del software, denominado “*ExtractorDatosContratacionPublica.jar*”. La estructura de la carpeta contenedora del código fuente es idéntica a la estructura explicada en la sección anterior.

### 3.3.4 Pruebas realizadas

Las pruebas de este software desarrollado se pueden dividir en los siguientes tipos:

- Pruebas de lectura
- Pruebas de escritura
- Pruebas de generación de ficheros
- Pruebas de correcta organización de ficheros

Para las **pruebas de lectura**, durante la iteración del programa se muestra por pantalla cada uno de los nodos leídos con sus respectivos datos, y posteriormente se realiza un extrapolado de los mismo con el fichero original.

```
Se invoca a la función textoXml con el valor del contador: 3
- IssueDate
Se invoca a la función textoXml con el valor del contador: 3
2012-12-27
```

*Ilustración 25 Pruebas de lectura*

Para las **pruebas de escritura**, se ha de realizar un control de cada uno de los ficheros generados en parte y realizar un extrapolado de los datos escritos en cada uno los ficheros junto con el fichero de datos de Contratación Pública. Además, con la intención de facilitar la tarea de depuración de erros, se hace uso de la consola de Eclipse para mostrar los datos que se deben de insertar en cada fichero, como se puede ver en la Ilustración 23.

```
Se inserta en el fichero: Party_1+++++
id, Party, WebsiteURI, PartyIdentification, ID
799224, http://www.sepes.es, EA0000845
+++++
```

*Ilustración 26 Escritura de ficheros*

Para comprobar la **correcta generación de los ficheros**, se debe realizar un previo análisis de los ficheros que se deben de generar gracias a los “id” o “code” que cada contrato contenga, y después este número debe ser igual al número de ficheros generados.

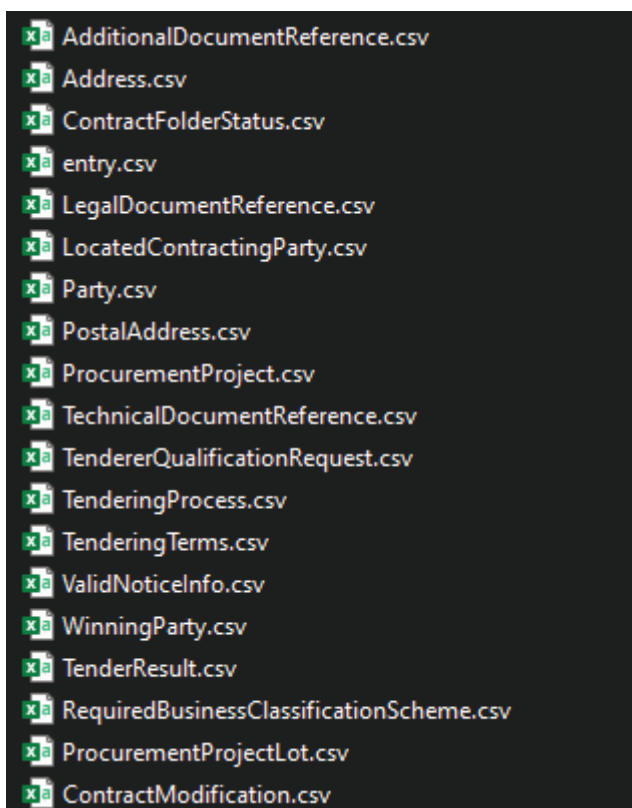
Las pruebas de la **correcta organización de los ficheros** hacen referencia a detectar los posible errores en la creación de los mismos, como pueden ser:

- La repetición de datos entre dos ficheros cuya generación se sucede linealmente.
- El desplazamiento de las columnas y registros de un fichero a otro sucesivo.
- La pérdida de datos al generar los ficheros.

## 4 Resultados y conclusiones

### 4.1 Resultados

Este trabajo ha consistido en desarrollar un programa capaz de extraer datos de los ficheros provenientes de la Plataforma de Contratación Pública del Estado. Lo que se ha buscado es la posibilidad de dar a los usuarios o personas noveles en el ámbito de los datos abiertos de Contratación Pública del Estado y el tratamiento de los ficheros mencionados durante el transcurso de este documento, un entorno más asequible o familiar para la visualización y el entendimiento de los mismos. Así mismo, proveer a los profesionales implicados en las tareas de Contratación Pública, la opción de poder realizar labores de limpieza de datos, y una base de partida para la generación de un Grafo de Conocimientos a partir de los datos extraídos. A continuación, se muestran los ficheros generados.

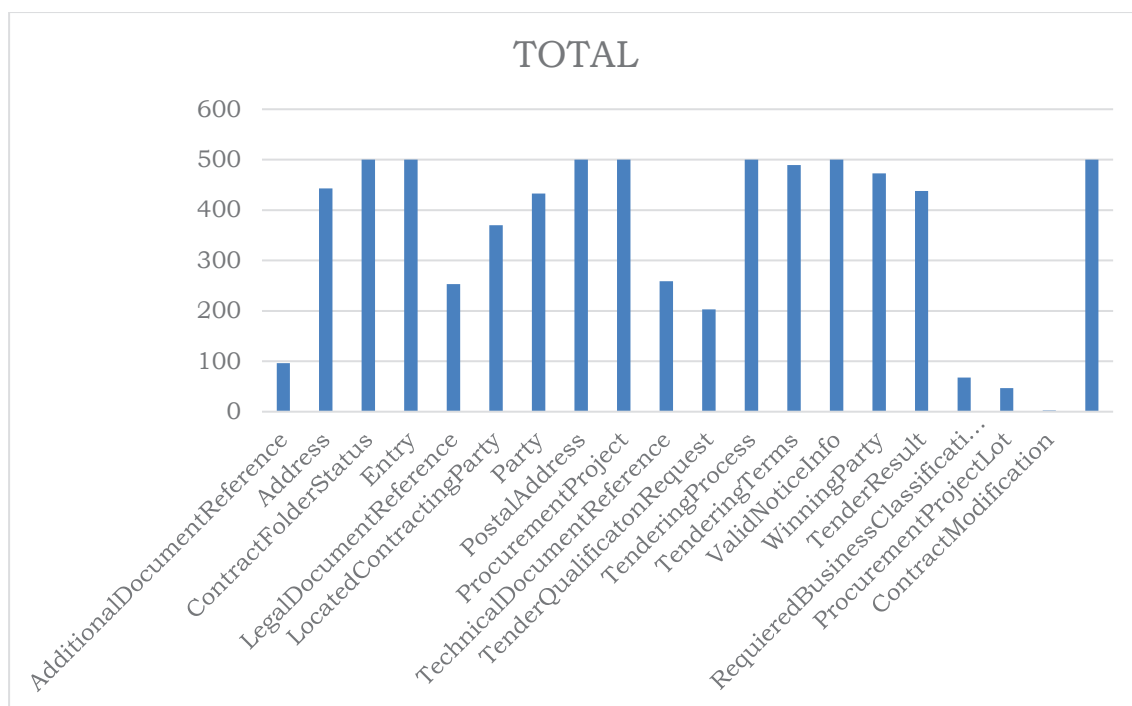


*Ilustración 27 Ficheros generados a partir de archivo de Contratación Pública*

La estructura de los archivos generados a partir de los datos extraídos ha sido diseñada de tal manera que acepta la adición de datos de distintos ficheros de Contratación Pública, lo que posibilita tener datos agrupados de múltiples fechas de publicación para el análisis y tratamiento de los mismos. Es importante mencionar que algunos de los ficheros generados contienen errores en su estructura de registros causados por la introducción errónea de distintos

datos en el archivo de Contratación Pública. En un futuro desarrollo, una línea a seguir estaría definida por la implementación de una función capaz de entender, clasificar y tratar los errores de formato antes de su inserción en los archivos .csv.

A partir de estos ficheros generados, se puede realizar un estudio para determinar aquellos elementos que son empleados mayoritariamente en los ficheros de Contratación Pública y aquellos que son secundarios.



*Ilustración 28 Gráfica de apariciones en contratos*

Como se puede observar en el gráfico obtenido a partir del análisis del fichero de licitaciones “licitacionesPerfilesContratanteCompleto3.atom”, última vez actualizado en mayo de 2016, en base a los ficheros obtenidos se puede establecer que los ficheros cuyos valores no alcanzan las 100 apariciones tienden a ser menos relevantes para el correcto desarrollo de los Contratos de Licitación Pública del Estado.

## 4.2 Conclusiones

Gracias a este proyecto he ampliado mis conocimientos acerca de los procesos llevados a cabo en la adjudicación de una actividad requerida en materias de bienes o servicios públicos. Los estudios previos para el desarrollo del programa encargado de las extracción de los datos de los ficheros de Contratación Pública me han servido para adentrarme en el ámbito del tratamiento de datos y en aumentar mis conocimientos sobre los ficheros XML y las posibilidades que

existen para el tratamiento de los mismos. Además, ha añadido valor a mi experiencia como programador, debido al desafío que ha supuesto el desarrollo de un programa recursivo capaz de analizar el documento para extraer sus datos y generar automáticamente los ficheros en función de los requisitos predefinidos.

Por último, este proyecto me ha enriquecido intelectualmente con los distintos estándares que existen en el entorno de los datos generados por las Contrataciones Públicas, con el entendimiento de qué son los datos abiertos y las distintas licencias que se pueden aplicar a los mismos, ya que, si bien es verdad que siempre he escuchado el término "*Creative Commons*", no ha sido hasta el desarrollo de este proyecto cuando de verdad lo he entendido.

Con este proyecto en el que he puesto toda mi dedicación, todas mis ganas y todo mi sudor y esfuerzos, siento que puedo dar por terminada una etapa de mi vida, y dar paso a una nueva, con el objetivo de convertirme en una persona mejor en los desempeños que tiene este oficio y en el trabajo del día a día. Una vez más, se ha demostrado que aquel que lucha y es constante logra todos sus objetivos, por muy difícil que parezca el camino.

## **5 Análisis de Impacto**

En el capítulo final de este proyecto de fin de grado, se procede a explicar el impacto que el mismo ha tenido y tendrá en un futuro en distintos contextos. Para ello se selecciona cada contexto en parte y se analiza.

### **5.1 Impacto personal**

Desde el punto de vista del ámbito personal, este proyecto ha tenido impacto sobre todo en términos de conocimientos relativos a los Procesos de Contratación, del que el autor no tenía conocimiento alguno. También conlleva el entendimiento de los datos abiertos y el inicio de la primera publicación de un fichero de este tipo, en concreto el fichero contenedor del código fuente del software, siendo este disponible en la Plataforma GitHub, y accesible por todos los usuarios deseosos de entender el desarrollo y usar el software para distintos fines. Además, ha hecho al autor cuestionarse la visión que tenía acerca del desarrollo de software en términos de beneficios, ya que siempre se ha tenido la visión de desarrollar una determinada aplicación para poder publicarla o venderla y a raíz de esto obtener el mayor beneficio económico posible, pero después de analizar los datos abiertos y explicar qué son y para qué se utilizan, se han sentado las bases de una nueva visión relativa a la ayuda que los mismos pueden proveer al resto de usuarios, ya que después de pensar acerca de todos los años pasados en la carrera, se ha llegado a la conclusión de que nadie de los presentes en el último año estaría donde está sin la ayuda de otros, siendo yo el autor el primero en reconocer y valorar esto.

Por último, ha obligado al autor a establecer las bases de un proyecto de desarrollo con las planificaciones y análisis que esto conlleva, y siendo esto de mucha ayuda para el futuro profesional del mismo.

### **5.2 Impacto empresarial**

Para el ámbito empresarial, este proyecto supone un avance por cómo está pensado y desarrollado el programa, ya que se basa en un proceso recursivo que es capaz de entender y generar los ficheros en formato .csv relativos a los datos abiertos. En este ámbito se espera dar apoyo a los profesionales de las Administraciones Públicas para un proceso más sencillo de análisis, procesado y limpieza de los datos relativos a las Contrataciones Públicas. Debido a que el software se encuentra disponible en forma de datos abiertos, las distintas entidades pueden acceder al software e incluso usarlo como referencia o base para el desarrollo de un software más potente y capaz de generar ficheros desde cualquier fichero con formato XML sin que contenga este datos pertenecientes a Contratación Pública. Este último argumento se extrae de la propia experiencia del autor en entornos laborales, en los que la mayoría de datos

sirven para generar informes, y estos informes generalmente se generan a partir de tablas con formato .csv.

### **5.3 Impacto social**

Este proyecto de fin de grado surge como solución a la necesidad de poder tener una mejor visión de cada uno de los Procesos de Contratación Pública y dar a los profesionales de las Administraciones Públicas una herramienta que les permita identificar los errores por medio de la división de los datos en distintos ficheros, a los que se pueden aplicar reglas de limpieza o tratamiento de datos, y garantizar así una mayor integridad de los mismos. Además, podría ser empleado en un futuro por la propia Plataforma de Contratación Pública como idea o base para la sección de datos abiertos que actualmente contiene un enlace roto.

A parte de esto, el proyecto también busca acercar a aquellos usuarios noveles o deseosos de ampliar sus conocimientos en el campo de las Contrataciones y Licitaciones Públicas, a los datos abiertos, a qué son los datos abiertos y a los procesos llevados a cabo por las Administraciones Públicas. Es importante remarcar el hecho de que este proyecto genera un software libre, siendo esto idéntico a los datos abiertos, cuyos principios son permitir la reutilización y redistribución a todos los usuarios sin distinciones relativas a etnias o culturas.

### **5.4 Impacto económico**

El impacto económico de este proyecto es nulo, ya que para su desarrollo se han empleado datos abiertos de la Contratación Pública del Estado, y el resultado del mismo es un software libre que se encuentra disponible en el repositorio GitHub del autor.

### **5.5 Impacto medioambiental**

En cuanto al impacto medioambiental que puede tener este proyecto, se puede decir que para su desarrollo se han empleado un portátil y un monitor, por lo que desde el punto de vista de la generación e implementación de este documento y del software libre resultante no ha habido un gasto energético inapropiado o fuera de lo normal. En cuanto a los resultados obtenidos, se calcula un impacto mínimo o incluso nulo, ya que el software se guarda, y a su vez se posibilita el acceso al mismo, en la plataforma GitHub.

## **5.6 Impacto cultural**

Debido a que los datos tratados pertenecen al Gobierno de España, no se espera tener un gran impacto cultural ya que se explica y se propone la implementación de un herramienta para el uso de los mismos. Aun siendo esto así, es interesante valorar la posibilidad que otorga este trabajo para acercar a usuarios de otros países o culturas a aprender acerca de cómo se desarrollan y gestionan los Procesos de Licitación y Contratación Pública en España.


## 6 Bibliografía

- [1] «Wikipedia,» 1 agosto 2020. [En línea]. Available: [https://es.wikipedia.org/wiki/Operador\\_econ%C3%B3mico\\_autorizado#Ventajas\\_en\\_Espa%C3%B1a](https://es.wikipedia.org/wiki/Operador_econ%C3%B3mico_autorizado#Ventajas_en_Espa%C3%B1a). [Último acceso: 20 abril 2021].
- [2] M. D. HACIENDA, «PLATAFORMA DE CONTRATACIÓN DEL SECTOR PÚBLICO,» GOBIERNO DE ESPAÑA, [En línea]. Available: <https://contrataciondelestado.es/wps/portal/plataforma>. [Último acceso: 20 abril 2021].
- [3] G. d. España, «Portal Institucional del Ministerio de Hacienda,» 05 Abril 2021. [En línea]. Available: [https://www.hacienda.gob.es/ES/ES/GobiernoAbierto/Datos%20Abiertos/Paginas/licitaciones\\_plataforma\\_contratacion.aspx](https://www.hacienda.gob.es/ES/ES/GobiernoAbierto/Datos%20Abiertos/Paginas/licitaciones_plataforma_contratacion.aspx). [Último acceso: 2021 Abril 2021].
- [4] A. P. y. Estado, «administración.gob.es,» [En línea]. Available: [https://administracion.gob.es/pag\\_Home/espanaAdmon/Transparencia-y-datos-abiertos-2/datos\\_abiertos.html](https://administracion.gob.es/pag_Home/espanaAdmon/Transparencia-y-datos-abiertos-2/datos_abiertos.html). [Último acceso: 24 abril 2021].
- [5] O. K. Foundation, «Open Knowledge Foundation,» [En línea]. Available: <https://okfn.org/opendata/>. [Último acceso: 24 abril 2021].
- [6] A. Cerrillo, «Los principios de los datos abiertos en la legislación Española,» *IDP*, pp. 62-77, 2014.
- [7] C. E. p. A. L. y. e. Caribe, «Biblioguías - Biblioteca de la CEPAL,» 18 Diciembre 2020. [En línea]. Available: <https://biblioguias.cepal.org/gestion-de-datos-de-investigacion>. [Último acceso: 30 Abril 2021].
- [8] D. G. d. P. d. Estado, «Plataforma de Contratación del Estado,» 14 Septiembre 2015. [En línea]. Available: [https://contrataciondelestado.es/codice/2.0/doc/CODICE\\_2\\_GuiaImplementacion\\_v1.3.pdf](https://contrataciondelestado.es/codice/2.0/doc/CODICE_2_GuiaImplementacion_v1.3.pdf). [Último acceso: 12 mayo 2021].
- [9] G. d. España, «Plataforma de Contratación Pública del Estado,» 15 Septiembre 2015. [En línea]. Available: [https://contrataciondelestado.es/codice/extension/doc/CODICE-PLACE-EXT\\_Guia\\_de\\_Implementacion.v.1.0.pdf](https://contrataciondelestado.es/codice/extension/doc/CODICE-PLACE-EXT_Guia_de_Implementacion.v.1.0.pdf). [Último acceso: 09 Mayo 2021].
- [10] G. d. España, «Plataforma de Contratación Pública del Estado,» 14 Septiembre 2015. [En línea]. Available: [https://contrataciondelestado.es/codice/2.0/doc/CODICE\\_2\\_GuiaImplementacion\\_v1.3.pdf](https://contrataciondelestado.es/codice/2.0/doc/CODICE_2_GuiaImplementacion_v1.3.pdf). [Último acceso: 10 Mayo 2021].
- [11] M. d. E. y. Hacienda, «Plataforma de Contratación del Sector Público,» Diciembre 2006. [En línea]. Available: <https://www.hacienda.gob.es/Documentacion/Publico/D.G.%20PATRI>

MONIO/SG%20Clasificacion%20Contratistas%20y%20Registro%20Contratos/ProcesosDeNegocioDic06.pdf. [Último acceso: 10 Mayo 2021].

- [12] G. d. E. |. M. d. Hacienda, «Portal Institucional del Ministerio de Hacienda,» 05 Enero 2021. [En línea]. Available: [https://www.hacienda.gob.es/es-ES/GobiernoAbierto/Datos%20Abiertos/Paginas/licitaciones\\_plataforma\\_contratacion.aspx](https://www.hacienda.gob.es/es-ES/GobiernoAbierto/Datos%20Abiertos/Paginas/licitaciones_plataforma_contratacion.aspx). [Último acceso: 27 Abril 2021].
- [13] D. O. d. l. U. Europea, «Portal Institucional del Ministerio de Hacienda,» 15 Marzo 2008. [En línea]. Available: <https://www.hacienda.gob.es/Documentacion/Publico/D.G.%20PATRIMONIO/Junta%20Consultiva/Reglamento%20CPV/cpv%20principal.pdf>. [Último acceso: 9 Mayo 2021].
- [14] T. R. Jonathan Robie, «World Wide Web Consortium (W3C),» [En línea]. Available: <https://www.w3.org/TR/WD-DOM/introduction.html>. [Último acceso: 09 Marzo 2021].
- [15] c. commons, «creative commons,» 7 Noviembre 2017. [En línea]. Available: <https://creativecommons.org/licenses/?lang=es>. [Último acceso: 05 Mayo 2021].
- [16] O. C. E. d. Datos, «Open Contracting Estándar de Datos,» [En línea]. Available: [https://standard.open-contracting.org/latest/es/getting\\_started/contracting\\_process/](https://standard.open-contracting.org/latest/es/getting_started/contracting_process/). [Último acceso: 05 Mayo 2021].
- [17] «Wikipedia,» 04 Noviembre 2020. [En línea]. Available: [https://es.wikipedia.org/wiki/Sublime\\_Text](https://es.wikipedia.org/wiki/Sublime_Text). [Último acceso: 10 Mayo 2021].
- [18] «Wikipedia,» 06 Abril 2021. [En línea]. Available: [https://es.wikipedia.org/wiki/Eclipse\\_\(software\)](https://es.wikipedia.org/wiki/Eclipse_(software)). [Último acceso: 15 Mayo 2021].
- [19] ORACLE, «java,» [En línea]. Available: [https://www.java.com/es/download/help/whatis\\_java.html](https://www.java.com/es/download/help/whatis_java.html). [Último acceso: 15 Mayo 2021].
- [20] «Wikipedia,» 29 Abril 2021. [En línea]. Available: [https://es.wikipedia.org/wiki/Plataforma\\_Java](https://es.wikipedia.org/wiki/Plataforma_Java). [Último acceso: 15 Mayo 2021].

Este documento esta firmado por

	<b>Firmante</b>	CN=tfgm.fi.upm.es, OU=CCFI, O=Facultad de Informatica - UPM, C=ES
	<b>Fecha/Hora</b>	Fri May 28 09:20:19 CEST 2021
	<b>Emisor del Certificado</b>	EMAILADDRESS=camanager@fi.upm.es, CN=CA Facultad de Informatica, O=Facultad de Informatica - UPM, C=ES
	<b>Numero de Serie</b>	630
	<b>Metodo</b>	urn:adobe.com:Adobe.PPKLite:adbe.pkcs7.sha1 (Adobe Signature)