

# Grado en Ingeniería Informática de Gestión y Sistemas de Información

-Departamento de Ingeniería de Sistemas y Automática -

Trabajo de Fin de Grado

## **DATALAICE**

From intelligent DATAbank generation to medical AI-based CAD  
system creation



---

**Amaia Perez Gamboa**

**Julio 2022**

**Dirección:** Teodoro Rico Pastrana



# Resumen

---

La evolución tecnológica permite al campo de la medicina mas que avances en tratamientos y medicamentos. El desarrollo de nuevas herramientas y el creciente uso de las computadoras en aplicaciones clínicas generaron la necesidad de un método estándar de almacenamiento y transferencia de información entre los dispositivos.

Por consiguiente, para que la comunicación no sea dañada por el uso de diferentes formatos, se desarrolló el estándar DICOM (Digital Imaging and Communications in Medicine). Surge como consecuencia de la heterogeneidad de equipos y es una estandarización. Fue creado en 1983 por un comité norteamericano formado por el American Collage Of Radiology (ACR) y por la National Eletrical Manufacturers Association (NEMA). El actual estándar DICOM está en la tercera versión, que sigue en un constante desarrollo.

Mediante el desarrollo de DICOM, se estandarizan los protocolos de comunicación, los formatos de información y a su vez el almacenamiento de imágenes. No solo se trata de un fichero, sino que también establece reglas de comunicación, cuando dos sistemas quieren transmitir la imagen de un sitio a otro.

El uso de DICOM proporciona nuevas oportunidades en diversos campos como en la telemedicina ya que permite acceder de forma remota a las imágenes desde diferentes hospitales.

El presente proyecto tiene como objetivo llevar a cabo una investigación de formatos de imagen medica DICOM y realizar un estudio sobre las funcionalidades que proporciona. A su vez, para el almacenamiento y gestión de estas, se realiza una configuración y un despliegue de Orthanc DICOM – Server, cumpliendo la función de un sistema PACS. Por último, se lleva a cabo el desarrollo de una aplicación web sobre Orthanc para posibilitar búsquedas, consultas y tratamiento de datos, como la anonimización o seudonimización de los datos proporcionados por las imágenes. A su vez, esta aplicación provee la integración con otros PACS hospitalarios.

El desarrollo de esta memoria esta limitado debido a la confidencialidad de los datos que se tratan. El código de la aplicación contiene datos propios de la empresa y el acceso a los servidores conectados a ella también tienen un acceso limitado. Es por ello por lo que, el funcionamiento y la exposición de ciertos datos será expuesto con mas detalle en la defensa de este Trabajo Fin de Grado.

# Índice general

|   |              |
|---|--------------|
| <b>RESUMEN.....</b>                                       | <b>3</b>     |
| <b>INDICE GENERAL.....</b>                                | <b>4</b>     |
| <b>INDICE DE FIGURAS Y TABLAS.....</b>                    | <b>5-7</b>   |
| <b>I. INTRODUCCIÓN.....</b>                               | <b>9-13</b>  |
| a. Contextualización.....                                 | 9-10         |
| b. Descripción del proyecto.....                          | 11-12        |
| c. Organización del documento.....                        | 13           |
| <b>II. ALCANCE DEL PROYECTO.....</b>                      | <b>15-20</b> |
| a. Propuesta y objetivo del proyecto.....                 | 15-16        |
| b. Planificación de tiempo.....                           | 17-19        |
| i. Desglose de tareas.....                                | 17-18        |
| ii. Diagrama de Gantt.....                                | 18-19        |
| c. Presupuesto del proyecto.....                          | 20           |
| <b>III. CONCEPTOS TEORICOS.....</b>                       | <b>22-39</b> |
| a. El estándar DICOM.....                                 | 22-30        |
| i. Introducción.....                                      | 22           |
| ii. Estructura en memoria de un archivo DICOM.....        | 22-23        |
| iii. Estructura física de un archivo DICOM.....           | 24-30        |
| 1. Tipos de elementos de dato.....                        | 26           |
| 2. Metadatos de un archivo DICOM.....                     | 26-30        |
| b. Sistema PACS.....                                      | 31-34        |
| i. Orthanc – DICOM Server.....                            | 33           |
| 1. Api REST de Orthanc.....                               | 33-34        |
| c. Anonimización y seudonimización de archivos DICOM..... | 35-39        |
| i. Aspectos generales.....                                | 35-36        |
| ii. Ejemplos de código.....                               | 37-39        |
| <b>IV. DATAIACE.....</b>                                  | <b>41-66</b> |
| a. Análisis y diseño de la aplicación.....                | 41-43        |
| b. Programas y herramientas.....                          | 44-45        |
| c. Base de datos.....                                     | 46-47        |
| d. Proceso de anonimización/seudonimización.....          | 48-51        |
| i. Tratamiento de datos.....                              | 48-49        |
| ii. Desarrollo del proceso.....                           | 49-51        |
| e. Funcionamiento e implementación.....                   | 52-66        |
| <b>V. CONCLUSIONES.....</b>                               | <b>68</b>    |
| <b>VI. BIBLIOGRAFÍA.....</b>                              | <b>70-71</b> |

# Índice de figuras y tablas

## INDICE DE FIGURAS

|   |    |
|---|----|
| <b>I. INTRODUCCIÓN</b>  |    |
| 1.1.- Esquema del generador de banco de datos.....                        | 11 |
| <b>II. ALCANCE DEL PROYECTO</b>   |    |
| 2.1.- Diagrama de Gantt.....  | 19 |
| <b>III. CONCEPTOS TEORICOS</b>  |    |
| 3.1. - Estructuras principales del modelo de información DICOM.....       | 23 |
| 3.2. - Relación entre los IODs y atributos .....                          | 23 |
| 3.3.- Estructura de un archivo DICOM.....                                 | 24 |
| 3.4. - Ejemplos de etiquetas a nivel de paciente.....                     | 27 |
| 3.5.- Ejemplos de etiquetas a nivel de estudio.....                       | 27 |
| 3.6.- Ejemplos de etiquetas a nivel de serie.....                         | 28 |
| 3.7.- Ejemplos de etiquetas a nivel de instancia.....                     | 29 |
| 3.8.- Diagrama de entidad-relación.....                                   | 30 |
| 3.9.- Logo de Orthanc – Dicom Server.....                                 | 33 |
| <b>VII. DATALAICE</b>   |    |
| 4.1.- Esquema del módulo de ingesta de datos.....                         | 41 |
| 4.2.- Archivos de Angular.....  | 45 |
| 4.3.- Esquema de la base de datos.....                                    | 46 |
| 4.4.- Pantalla de ‘Extractor de estudios’.....                            | 52 |
| 4.5.- Pantalla de ‘Pacientes’.....  | 53 |
| 4.6.- Filtro de búsqueda ‘Nombre de paciente’.....                        | 53 |
| 4.7.- Filtros de búsqueda ‘ID’ y ‘Nombre de paciente’.....                | 53 |
| 4.8.- Pantalla de ‘Estudios’.....   | 54 |
| 4.9.- Filtros de búsqueda ‘Nombre del paciente’ y ‘Edad’ en estudios..... | 54 |

|   |    |
|---|----|
| 4.10.- Filtros de búsqueda 'Modalidad' y 'Fecha desde' en estudios.....                           | 55 |
| 4.11.- Filtros de búsqueda 'Fabricante' y 'Fecha hasta' en estudios.....                          | 55 |
| 4.12.- Filtros de búsqueda 'ID del paciente', 'Sexo' y 'Descripción del estudio' en estudios..... | 55 |
| 4.13.- Referencia al símbolo de detalles en la pantalla 'Estudios'.....                           | 56 |
| 4.14.- Ventana emergente 'Detalles de estudio'.....   | 56 |
| 4.15.- Botón para anonimizar/seudonimizar estudios.....   | 56 |
| 4.16.- Pantalla para seleccionar los estudios.....  | 57 |
| 4.17.- Ventana emergente para anonimizar/seudonimizar.....  | 57 |
| 4.18.- Ventana emergente para crear un proyecto.....  | 57 |
| 4.19.- Ventana emergente con los proyectos existentes.....  | 58 |
| 4.20.- Pantalla de 'Estudios' con el nuevo estudio anonimizado.....                               | 58 |
| 4.21.- Pantalla de 'Proyectos' con un nuevo proyecto.....   | 58 |
| 4.22.- Pantalla de 'Series' correspondiente a un estudio concreto.....                            | 59 |
| 4.23.- Filtros de búsqueda 'Modalidad' y 'Fecha desde' en series.....                             | 59 |
| 4.24.- Ventana emergente 'Detalles de serie' de un estudio.....                                   | 60 |
| 4.25.- Botón para anonimizar/seudonimizar series de un estudio.....                               | 60 |
| 4.26.- Pantalla para seleccionar las series de un estudio.....                                    | 60 |
| 4.27.- Pantalla de 'Series'.....  | 61 |
| 4.28.- Filtros de búsqueda 'Modalidad' y 'Fecha desde' en series.....                             | 61 |
| 4.29.- Ventana emergente 'Detalles de serie'.....   | 62 |
| 4.30.- Botón para anonimizar/seudonimizar series.....   | 62 |
| 4.31.- Pantalla para seleccionar las series.....  | 63 |
| 4.32.- Ventana emergente para crear un nuevo proyecto.....  | 63 |
| 4.33.- Pantalla de 'Estudios' con el nuevo estudio.....   | 63 |
| 4.34.- Pantalla de 'Proyectos' con un proyecto nuevo.....   | 64 |
| 4.35.- Ventana emergente para asignar un estudio a un proyecto existente.....                     | 64 |
| 4.36.- Ventana emergente de error.....  | 64 |
| 4.37.- Filtros de búsqueda 'ID del proyecto' y 'Nombre del proyecto' en proyectos....             | 64 |
| 4.38.- Referencia al símbolo para editar la descripción en la pantalla 'Proyectos'.....           | 65 |
| 4.39.- Ventana emergente 'Editar descripción'.....  | 65 |
| 4.40.- Referencia al botón push.....  | 65 |

|   |    |
|---|----|
| 4.41.- Pantalla de proyectos modificada.....                        | 66 |
| 4.42.- Pantalla de proyectos con modificaciones en un proyecto..... | 66 |

## INDICE DE TABLAS

### II. ALCANCE DEL PROYECTO

|   |    |
|---|----|
| 2.1.- Tabla con las tareas y fechas del proyecto..... | 18 |
| 2.2.- Tabla de costes.....                            | 20 |

### III. CONCEPTOS TEORICOS

|   |    |
|---|----|
| 3.1.- Representaciones de valores DICOM.....            | 25 |
| 3.2.- Resumen de la estructura de un archivo DICOM..... | 25 |

# Capítulo 1

## Introducción

### Contenido

---

- 1.1. Contextualización
  - 1.2. Descripción del proyecto
  - 1.3. Organización del documento
-



## 1.1. CONTEXTUALIZACIÓN

---

DATALAICE es un proyecto que actualmente está en vigor en la empresa Deusto Seidor, empresa donde he desarrollado mis prácticas voluntarias. Se trata de un proyecto conjunto con Ingecontrol, empresa tecnológica que proporciona productos y soluciones en los sectores de las TIC y los equipos electrónicos de medida.

Deusto Sistemas es una empresa de consultoría, desarrollo e implementación de servicios TIC creada en 2002 como un grupo de empresas tecnológicas, con el objetivo de liderar el mercado de soluciones informáticas y de telecomunicaciones. En cuanto a Seidor, se trata de una consultoría tecnológica que ofrece una gran variedad de soluciones y servicios de innovación. La unión de las dos entidades viene a raíz de que Seidor adquirió el 60% de la empresa Deusto Sistemas, por lo que, a partir de ese momento, Deusto Sistemas opera bajo el nombre de Deusto Seidor con el objetivo de ampliar el portfolio de servicios y de reforzar el posicionamiento como uno de los proyectos referentes de tecnologías de la información en Euskadi.

El consorcio formado por Deusto Sistemas e Ingecontrol, con el centro tecnológico Vicomtech como subcontratado, cuenta con la experiencia suficiente para el desarrollo del proyecto y con una estrategia de comercialización que fomenta que ambas empresas se posicionen a la vanguardia en la creación de CADs. El proyecto permite la generación de dos subproductos - generador de bancos de datos de imagen médica y generador de CADs-, que pueden ser comercializados de manera conjunta o incluso independiente ampliando así las oportunidades de mercado de ambas empresas.

DATALAICE se enmarca en el grupo de Proyectos de Investigación y Desarrollo de carácter competitivo de Hazitek y está orientado al desarrollo de dos subproductos que implican un alto nivel de conocimiento científico y tecnológico a fin de dar mejor satisfacción a las demandas del mercado y de la sociedad. El proyecto trata de impulsar el nicho de oportunidad de la salud personalizada y la dimensión de la prioridad estratégica de Biociencias/Salud, ayudando a potenciar las capacidades empresariales y mejorar la competitividad de Deusto Sistemas e Ingecontrol.

En la era actual, se requiere de una mayor personalización de la práctica asistencial mediante la adopción de tratamientos personalizados, asociados muchas veces a nuevas tecnologías de diagnóstico en los que la imagen médica tiene un papel muy relevante. Se pretende que esta tecnología repercuta en una mejor gestión de los pacientes, especialmente en aquellas patologías en las que los tratamientos de medicina personalizada pueden aportar una mayor calidad en la atención. Así, se destaca la importancia de desarrollar sistemas de ayuda a la detección y diagnóstico de patologías en base a estudios de imagen, que permitan reducir el coste actual de las pruebas de diagnóstico y aumentar la eficacia de los tratamientos a través de una mejor estratificación de los pacientes para mejorar su calidad de vida.

Los productos desarrollados en DATALAICE, el generador de bancos de datos y el generador de CADs, se diseñarán de tal manera que proporcionen grandes ventajas respecto a la forma existente, ya que serán productos con funcionalidades únicas en el mercado que abarquen todo el ciclo de creación de CADs de manera ágil, adaptable e integrada, con un claro enfoque en la futura certificación y comercialización de los CADs generados con las soluciones de DATALAICE.

## 1.2. DESCRIPCIÓN DEL PROYECTO

En el proyecto Datalaice se desarrollan 2 soluciones, que pueden comercializarse de forma conjunta o independiente. Un generador de bancos de datos de imagen medica que se concibe como un conjunto de módulos y servicios que permiten seleccionar, recuperar, almacenar y anotar estudios de imagen medica y un generador de soluciones CAD que se considera como el conjunto de herramientas, servicios y aplicaciones que permiten entrenar modelos que provean información (predicción, clasificación, detección) y posteriormente desplegarlos para que se pueda realizar inferencia sobre nuevos datos.

El presente proyecto se centra en el desarrollo de la solución de un generador de bancos de datos, concretamente en el módulo de ingesta de datos. El generador de bancos de datos tiene como objetivo principal la realización de una ingesta de datos desde PACS utilizando protocolos estándar y desde bases de datos, anonimizar o seudonimizar metadatos sensibles y proveer acceso a los bancos de datos generados. A su vez, se proyecta una gestión de un sistema PACS en la nube, donde se faciliten ficheros DICOM y se puedan realizar anotación de los datos y su almacenamiento.

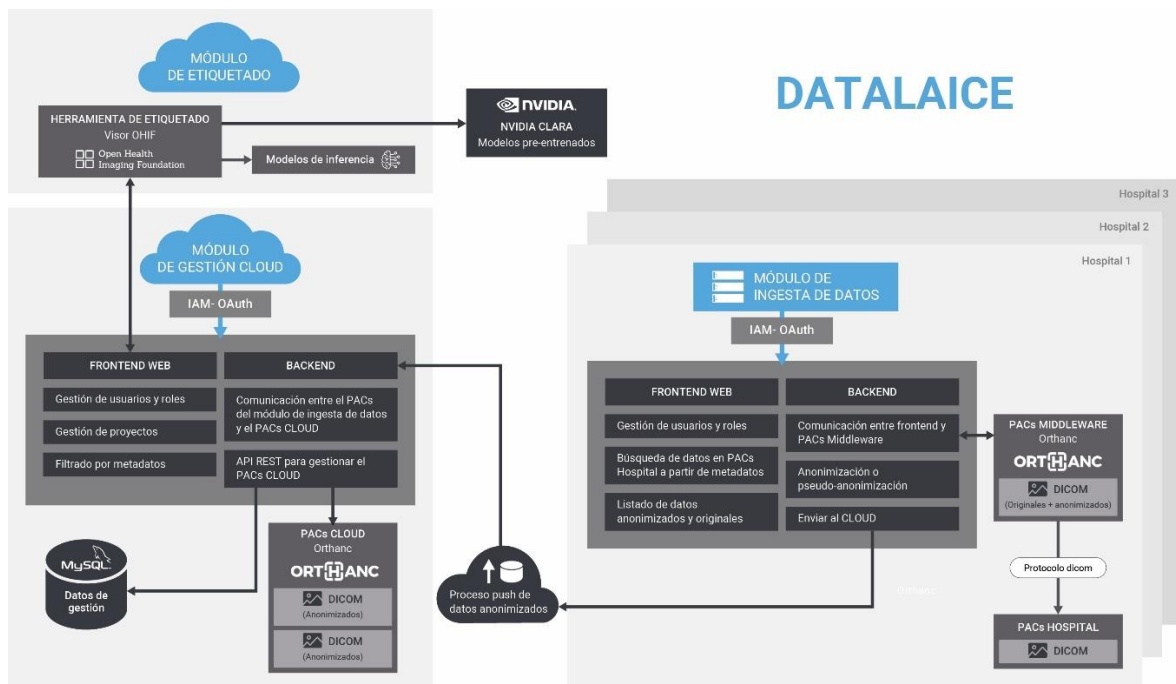


Figura 1.1 - Esquema del generador de banco de datos

Como ha sido mencionado anteriormente, la funcionalidad de este proyecto dentro del generador de banco de datos se centra concretamente en el modulo de ingesta de datos, que permite conectarse otro PACS utilizando el protocolo DICOM para buscar y recuperar estudios de imagen.

Este módulo cumple con los siguientes requerimientos:

- La provisión de una infraestructura de almacenamiento de datos de imagen y anotaciones a gran escala.
- La posibilidad de realizar peticiones y transferencias de estudios de imagen desde otros PACS.
- Búsquedas de estudios en PACS de manera sencilla e intuitiva mediante la utilización de determinados metadatos a modo de filtro.
- La provisión de herramientas de anonimización o seudonimización de estudios de imagen.
- La disposición de un proceso de volcado de datos anonimizados al módulo de gestión CLOUD.

## 1.3. ORGANIZACIÓN DEL DOCUMENTO

---

Esta es la organización de la presente memoria que documenta el Trabajo de Fin de Grado:

- *Capítulo 1 – Introducción:* introducción al presente documento, en el que se define en que consiste el proyecto.
- *Capítulo 2 – Alcance del proyecto:* definición de los objetivos que tiene el proyecto y como se ha organizado el tiempo para llevarlo a cabo. Por otro lado, se hace un presupuesto teniendo en cuenta los recursos empleados para el desarrollo de este.
- *Capítulo 3 – Contenido teórico:* investigación realizada sobre los diferentes recursos técnicos necesarios para la realización de la práctica.
- *Capítulo 4 – DataIaice:* exposición de la interfaz web, de las herramientas utilizadas y del resultado de la implementación.
- *Capítulo 5 – Conclusiones:* conclusiones arrojadas tras las anteriores fases del Trabajo Fin de Carrera.

# Capítulo 2

## Alcance del proyecto

### Contenido

---

- 2.1 Propuesta y objetivo del proyecto
  - 2.2 Planificación de tiempo
    - 2.2.1 Desglose de tareas
    - 2.2.2 Diagrama de Gantt
  - 2.3 Presupuesto del proyecto
-

## 2.1. PROPUESTA Y OBJETIVO DEL PROYECTO

---

Durante las últimas décadas, el creciente desarrollo de tecnologías no invasivas de imagen está habilitando nuevos horizontes para analizar y estudiar la anatomía del cuerpo humano y sus funciones. Esto se debe en gran medida a las imágenes de resonancia magnética (MRI) que proporcionan grandes cantidades de datos con un nivel de calidad muy elevado.

Así, el objetivo general del proyecto DATALAICE consiste en la creación y demostración de una solución tecnológica junto a la infraestructura necesaria para generar nuevas aplicaciones de soporte a la detección y diagnóstico asistido por ordenador de diversas patologías basado en técnicas de inteligencia artificial. Actualmente, en la creación y puesta en marcha de un nuevo sistema CAD, se implementan estas herramientas de forma adaptada al proyecto, aun cuando suelen compartir funcionalidades similares. La presente propuesta proporcionará un conjunto de herramientas genérico que venga a eliminar la necesidad de implementar dichas herramientas una y otra vez, reduciendo de forma considerable los costes y riesgos tecnológicos asociados a la creación y despliegue de nuevos sistemas CAD basados en imagen.

Para cumplir con los mencionados objetivos generales, DATALAICE debe cumplir con los siguientes objetivos basados en esta práctica. En función a los objetivos tecnológicos, este proyecto debe llevar a cabo el desarrollo de un extractor de estudios de PACS hospitalarios basado en el estándar DICOM, que permita realizar la selección y recuperación de casos según una serie de criterios de inclusión/exclusión. En cuanto a los objetivos socioeconómicos, el proyecto está diseñado para agilizar el proceso de extracción y anotación de datos requeridos para el desarrollo de nuevos productos CAD o ensayos clínicos, reduciendo los costes asociados a dichas tareas.

En relación con dichos objetivos, el módulo de ingesta de datos cumple con las siguientes funcionalidades:

- La infraestructura de almacenamiento se basa en el despliegue de un PACS de código abierto, Orthanc, en concreto.
- Se implementa una interfaz web para realizar búsquedas en diferentes PACS hospitalarios.
- Se implementa una aplicación *backend* para almacenar las referencias (rutas, conexiones) de los diferentes PACS. Este *backend* comunica la interfaz web con el PACS Middleware de este módulo y con los PACS hospitalarios.
- El módulo de ingesta debe contar con una herramienta de anonimización o pseudonimización de los datos en función de las necesidades del usuario.

- El módulo proporciona funcionalidades *push* para enviar ficheros dicom al módulo de gestión de anotaciones CLOUD.



## 2.2. PLANIFICACIÓN DE TIEMPOS

---

### 2.2.3. DESGLOSE DE TAREAS

En este apartado se especifica el desglose de tareas necesarias para el desarrollo de la practica:

- Investigación y formación
  - El formato DICOM. Búsqueda de la definición del estándar, de los diversos metadatos que contiene y estudio de la estructura interna que tiene un archivo DICOM.
  - Proceso de anonimización/seudonimización: Búsqueda de información sobre la anonimización yseudonimización y sobre la relevancia de las etiquetas DICOM en el proceso. De la misma manera, estudio de ejemplos de anonimización yseudonimización de ficheros DICOM.
  - Orthanc: Instalación de Orthanc y realización de pruebas de recuperación de ficheros DICOM
  - Grails: Seguir tutoriales de uso de Grails para entender su funcionamiento. Por otro lado, estudiar el lenguaje de programación groovy y comparar las similitudes con Java.
  - Angular: Familiarizarse con el uso de Angular.
  - Curso de Python
    - Estudio de la librería *pydicom*, la cual será usada para la gestión de lectura y escritura de los archivos DICOM en el proceso de anonimización oseudonimización.
- Realizar un pequeño ejemplo de aplicación en Visual Studio Code con Grails y Angular para entender la estructura de la aplicación DATALAICE.
- Crear las pantallas en la interfaz web con los datos de los pacientes, estudios y series que se recuperan desde Orthanc.
- Establecer un filtro de búsqueda en base a las etiquetas DICOM para pacientes, estudios y series.
- Integrar una base de datos para gestionar la creación de los nuevos proyectos con los DICOM anonimizados y sus respectivos datos.
  - Definir y crear el modelo de datos de la BD
- Proceso de anonimización/seudonimización
  - Estudio de las etiquetas de un DICOM original vs un DICOM anonimizado: clasificar las etiquetas DICOM en base a la importancia que tienen en un archivo anonimizado.
  - Implementación del proceso
  - Realización de pruebas para verificar su funcionamiento
  - Insertar los datos en la BD

- Gestión de proyectos
  - Diseñar y crear la pantalla de proyectos en la interfaz web
  - Insertar los datos de los proyectos en la BD y gestionar las relaciones de estos con los datos de los DICOM anonimizados.
- Gestión del proceso *push*
  - Recuperar cada DICOM anonimizado desde Orthanc y mandarlo a un nuevo Orthanc conectado al módulo CLOUD, mencionado anteriormente.
  - Gestión de los ficheros DICOM que se mandan al módulo CLOUD.
- Pruebas y correcciones
- Reuniones: dos reuniones semanales con el director del proyecto en la empresa.

### 2.2.3. DIAGRAMA DE GANTT

A continuación, se desarrolla un diagrama de Gantt para llevar a cabo un mejor proceso de gestión del proyecto, aportando claridad, una vista general simplificada, una mejor gestión del tiempo y una mayor flexibilidad. De esta manera, identificamos cuales son los componentes básicos del proyecto para poder organizarlo en tareas mas gestionables.

En la siguiente tabla se especifica una estimación de los periodos de tiempo que han sido dedicados a cada tarea en base a fechas.

| #         | TAREA  | INICIO     | FIN        |
|-----------|--|------------|------------|
| <b>1</b>  | <b>Investigación y formación</b>                   |            |            |
| 1.1       | El formato DICOM                                   | 11/01/2022 | 17/02/2022 |
| 1.2       | Anonimización/seudonimización                      | 11/01/2022 | 15/04/2022 |
| 1.3       | Orthanc  | 20/01/2022 | 27/01/2022 |
| 1.4       | Grails   | 24/01/2022 | 28/01/2022 |
| 1.5       | Angular  | 24/01/2022 | 28/01/2022 |
| 1.6       | Curso de Python                                    | 28/02/2022 | 08/03/2022 |
| 1.6.1     | Librería <i>pydicom</i>                            | 08/03/2022 | 01/04/2022 |
| <b>2</b>  | <b>Crear ejemplo de aplicación</b>                 | 27/01/2022 | 03/02/2022 |
| <b>3</b>  | <b>Crear pantallas de la interfaz web</b>          |            |            |
| 3.1       | Pantalla de los pacientes                          | 02/02/2022 | 08/02/2022 |
| 3.2       | Pantalla de los estudios                           | 02/02/2022 | 10/02/2022 |
| 3.3       | Pantalla de las series                             | 14/02/2022 | 16/02/2022 |
| <b>4</b>  | <b>Establecer filtro de búsqueda</b>               |            |            |
| 4.1       | Filtro para los pacientes                          | 27/04/2022 | 29/04/2022 |
| 4.2       | Filtro para los estudios                           | 10/02/2022 | 18/02/2022 |
| 4.3       | Filtro para las series                             | 21/02/2022 | 24/02/2022 |
| <b>5</b>  | <b>Integrar una base de datos</b>                  | 19/04/2022 | 22/04/2022 |
| <b>6</b>  | <b>Proceso de anonimización/seudoanonimización</b> |            |            |
| 6.1       | Estudiar las etiquetas DICOM                       | 08/03/2022 | 24/03/2022 |
| 6.2       | Implementar el proceso                             | 10/03/2022 | 14/04/2022 |
| 6.3       | Insertar los datos en la BD                        | 25/04/2022 | 03/05/2022 |
| 6.4       | Pruebas de verificación                            | 14/04/2022 | 24/05/2022 |
| <b>7</b>  | <b>Gestión de proyectos</b>                        |            |            |
| 7.1       | Diseñar y crear la pantalla                        | 29/04/2022 | 04/05/2022 |
| 7.2       | Insertar los datos en la BD                        | 05/05/2022 | 13/05/2022 |
| <b>8</b>  | <b>Gestión del proceso <i>push</i></b>             |            |            |
| <b>9</b>  | <b>Pruebas y correcciones</b>                      | 13/05/2022 | 10/06/2022 |
| <b>10</b> | <b>Reuniones</b>                                   | 11/01/2022 | 10/06/2022 |

Tabla 2.1. - Tabla con las tareas y fechas del proyecto

Por otro lado, se ha realizado un diagrama mas concreto, por cada semana, en base a las fechas indicadas en cada tarea. Cabe mencionar, que las fechas especificadas son una estimación que por factores internos o externos han podido ser ligeramente modificadas.

### Diagrama de Gantt

| Nombre del proyecto |  | DATAIACE   |            |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
|---------------------|--|------------|------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Fecha de Inicio     |  | 11/01/2021 |            |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| #                   | TAREA                                    | INICIO     | FIN        | Semana 1 | Semana 2 | Semana 3 | Semana 4 | Semana 5 | Semana 6 | Semana 7 | Semana 8 | Semana 9 | Semana 10 | Semana 11 | Semana 12 | Semana 13 | Semana 14 | Semana 15 | Semana 16 | Semana 17 | Semana 18 | Semana 19 | Semana 20 | Semana 21 | Semana 22 |
| 1                   | Investigación y formación                |            |            |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 1.1                 | El formato DICOM                         | 11/01/2022 | 17/02/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 1.2                 | Anonimización/seudonimización            | 11/01/2022 | 15/04/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 1.3                 | Orthanc                                  | 20/01/2022 | 27/01/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 1.4                 | Grails                                   | 24/01/2022 | 28/01/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 1.5                 | Angular                                  | 24/01/2022 | 28/01/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 1.6                 | Curso de Python                          | 28/02/2022 | 08/03/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 1.6.1               | Librería pydicom                         | 08/03/2022 | 01/04/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 2                   | Crear ejemplo de aplicación              | 27/01/2022 | 03/02/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 3                   | Crear pantallas de la interfaz web       |            |            |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 3.1                 | Pantalla de los pacientes                | 02/02/2022 | 08/02/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 3.2                 | Pantalla de los estudios                 | 02/02/2022 | 10/02/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 3.3                 | Pantalla de las series                   | 14/02/2022 | 16/02/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 4                   | Establecer filtro de búsqueda            |            |            |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 4.1                 | Filtro para los pacientes                | 27/04/2022 | 29/04/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 4.2                 | Filtro para los estudios                 | 10/02/2022 | 18/02/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 4.3                 | Filtro para las series                   | 21/02/2022 | 24/02/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 5                   | Integrar una base de datos               | 19/04/2022 | 22/04/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 6                   | Proceso de anonimización/seudonimización |            |            |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 6.1                 | Estudiar las etiquetas DICOM             | 08/03/2022 | 24/03/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 6.2                 | Implementar el proceso                   | 10/03/2022 | 14/04/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 6.3                 | Insertar los datos en la BD              | 25/04/2022 | 03/05/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 6.4                 | Pruebas de verificación                  | 14/04/2022 | 24/05/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 7                   | Gestión de proyectos                     |            |            |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 7.1                 | Diseñar y crear la pantalla              | 29/04/2022 | 04/05/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 7.2                 | Insertar los datos en la BD              | 05/05/2022 | 13/05/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 8                   | Gestión del proceso push                 |            |            |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 9                   | Pruebas y correcciones                   | 13/05/2022 | 10/06/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |
| 10                  | Reuniones                                | 11/01/2022 | 10/06/2022 |          |          |          |          |          |          |          |          |          |           |           |           |           |           |           |           |           |           |           |           |           |           |

Figura 2.1. - Diagrama de Gantt

## 2.3. PRESUPUESTO DEL PROYECTO

En este apartado se ha realizado un análisis de los costes que supone el desarrollo de este proyecto en función de los recursos necesarios. Los costes que se han identificado son los siguientes:

### Coste del personal (mano de obra)

Para la realización de esta practica se han empleado un total de 750 horas en la empresa. Este tiempo ha sido principalmente en formación, ya que se trata de un tema completamente nuevo para mí, en reuniones que tenían como objetivo realizar un control y dirigir el desarrollo del proyecto y en la ejecución del proyecto.

### Material para la implementación

Debido a las necesidades para llevar a cabo la implementación del proceso de anonimización, se ha visto la necesidad de adquirir conocimientos sobre el lenguaje de programación Python. Durante una semana, realice un curso sobre Python telemáticamente.

### Licencias Software

Se han hecho uso de varias plataformas, de las cuales una ha requerido de una licencia. En concreto, para poder acceder, editar y eliminar datos de una base de datos de esta práctica desde el programa Visual Studio Code, denominado ApexSQL. Se tiene en cuenta que 10 % del coste total es empleado en este trabajo.

### Equipamiento

Para la realización de la practica se ha hecho uso de un ordenador portátil *Lenovo*, con su respectivo ratón inalámbrico y unos cascos con micrófono para poder realizar las reuniones telemáticamente. Se considera que del rendimiento total del equipamiento se emplea un 15 % para la realización de la práctica.

| Tipo de coste                     | Coste total (€)        |
|-----------------------------------|------------------------|
| Coste del personal (mano de obra) | 2450 €                 |
| Material para la implementación   | 35 €                   |
| Licencias Software                | 20 € x 0,10 = 2 €      |
| Equipamiento                      | 850 € x 0,15 = 127,5 € |
| Beneficio industrial (%15)        | +392,175 €             |
| IVA (%21)                         | +631,40 €              |
| <b>TOTAL</b>                      | <b>3638,07 €</b>       |

Tabla 2.2. – Tabla de costes

# Capítulo 3

## Conceptos teóricos

### Contenido

---

- 3.1. El estándar DICOM
    - 3.1.1. Introducción
    - 3.1.2. Estructura en memoria de un archivo DICOM
    - 3.1.3. Estructura física de un archivo DICOM
      - 3.1.3.1. Tipos de elementos de dato
      - 3.1.3.2. Metadatos de un archivo DICOM
  - 3.2. Sistema PACS
    - 3.2.1. Orthanc – DICOM Server
      - 3.2.1.1. Api REST de Orthanc
  - 3.3. Anonimización y seudonimización de archivos DICOM
    - 3.3.1. Aspectos generales
    - 3.3.2. Ejemplos de código
-

## 3.1. EL ESTANDAR DICOM

---

### 3.1.1. INTRODUCCIÓN

La creciente tendencia hacia la digitalización de los hospitales y a la utilización de sistemas de adquisición y tratamiento digital de imágenes médicas es una constante hoy en día, sobre todo en el ámbito del diagnóstico por imagen. El uso de las imágenes médicas se caracteriza por la importancia que requiere para los diagnósticos de pacientes, tratamientos terapéuticos y evaluación de resultados, reduciendo así las dosis de radiación a los pacientes y los cortes anatómicos, por ejemplo.

Debido a los avances tecnológicos, el uso de la imagen digital ha aumentado de manera notoria debido a su gran calidad y a la posibilidad de transmitirla a distintos puntos de manera inmediata, haciendo necesaria la adopción de un estándar que facilite su uso.

DICOM es un estándar en comunicación e imágenes en medicina, que proporciona una transmisión segura de los archivos médicos entre hospitales, centros de investigación y departamentos, facilitando así el manejo sencillo de la información médica. El estándar ha sido desarrollado en función de las necesidades de fabricantes y usuarios con el equipamiento de imagen médica para la interconexión de dispositivos sobre redes estándares.

Por lo que se podría decir que el estándar Digital Imaging and Communications in Medicine (DICOM), es un protocolo de comunicación estándar de carácter internacional de los sistemas de información que unifica los formatos de almacenamiento de las imágenes médicas y la transmisión de estos de un hardware a otro.

### 3.1.2. ESTRUCTURA EN MEMORIA DE UN ARCHIVO DICOM

El estándar Digital Imaging and Communications in Medicine (DICOM), es un protocolo de comunicación estándar de carácter internacional de los sistemas de información que unifica los formatos de almacenamiento de las imágenes médicas y la transmisión de estos de un hardware a otro.

La estructura de datos se basa en un diseño orientado a objetos. En los modelos son llamados entidades y los atributos cumplen la función de describir las características de un objeto. Cuando se define un objeto, dando valores a los atributos, se denomina instancia del objeto.

Los objetos son agrupados en clases de objetos denominadas IOD (Information Object Definition) que están formadas por la imagen y su respectiva información. Dichas clases se comunican entre si mediante mensajes llamados DIMSE (DICOM Message Service Element), que contiene operaciones que pueden realizarse sobre un objeto. De esta manera, IOD y DIMSE forman SOP (Service-Object Pair), la unidad funcional de DICOM.

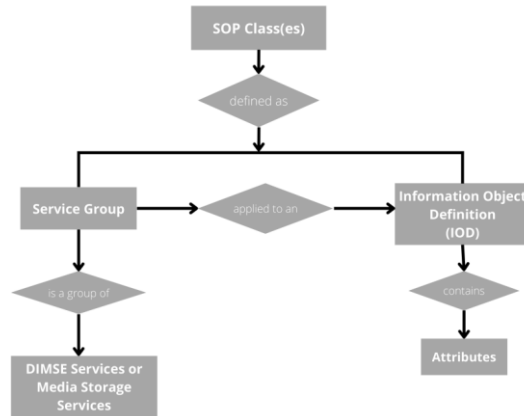


Figura 3.1. - Estructuras principales del modelo de información DICOM

Un IDO es un modelo de datos abstracto que se utiliza para especificar información sobre objetos del mundo real, es decir, un IOD no representa una instancia específica de un objeto del mundo real, sino una clase de objetos del mundo real que comparten las mismas propiedades.

Se compone de partes de información relacionada, agrupadas en entidades de información. Cada entidad contiene información sobre un único objeto como un paciente, un estudio, un equipo, una imagen, etc. Un IOD utilizado para representar una sola clase de objetos del mundo real se denomina objeto de información normalizada. Un IOD que incluye información sobre objetos del mundo real relacionados se denomina objeto de información compuesta.

Las entidades de información están formadas por atributos que describen una única parte de la información. Cada atributo está definido por un nombre, una etiqueta, un tipo de dato y una descripción. Los atributos que están relacionados se agrupan en módulos de información de objetos o IOMs (Information Object Modules) los cuales pueden ser usados en mas de un IOD.

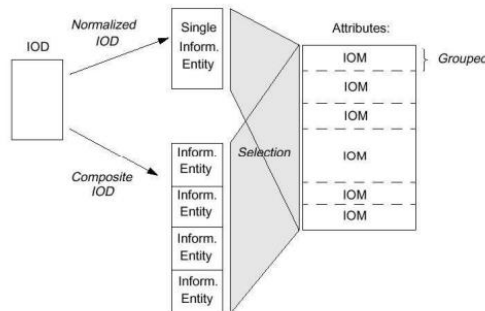


Figura 3.2. - Relación entre los IODs y atributos [20]

### 3.1.3. ESTRUCTURA FÍSICA DE UN ARCHIVO DICOM

Se establece una jerarquía entre los datos que permite realizar una clasificación según el contenido de la información por grupos, facilitando la identificación, el acceso a las variables y los parámetros de interés de un mismo archivo.

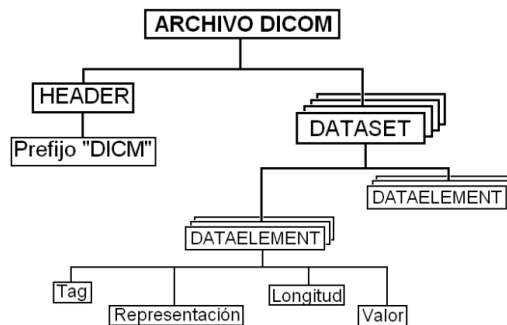


Figura 3.3.- Estructura de un archivo DICOM [5]

El elemento que diferencia un archivo DICOM es denominado HEADER o cabecera que consta de 128 bytes de archivos de preámbulo y 4 bytes de prefijo “DICM”. El preámbulo puede estar compuesto por información necesaria sobre la aplicación principal para su ejecución.

Por otro lado, el cuerpo del archivo esta forma por una secuencia de *DataSets* que contienen información sobre objetos del mundo real y que a su vez están constituidos por *DataElements*, que son valores codificados de los atributos del objeto, identificados y clasificados por un TAG o etiqueta.

Cada etiqueta identifica unívocamente un elemento de datos. Se trata de un par ordenado de números de 16 bits codificado en hexadecimal compuesto por dos partes. DICOM utiliza la siguiente anotación para representar una etiqueta (gggg,eeee), siendo el primer grupo de números el valor que representa al grupo de elementos, mientras que el segundo hace referencia al elemento en concreto.

Los elementos de datos estándares definidos por DICOM tienen el número de grupo par, excepto los elementos con el número de grupo (0000,eeee), (0002,eeee), (0004,eeee) o (0006,eeee) que identifican comandos y elementos de formato de los ficheros DICOM. Por otro lado, los elementos de datos que no son definidos por DICOM (datos privados) tienen un número de grupo impar diferente a (0001,eeee), (0003,eeee), (0005,eeee), (0007,eeee) o (FFFF,eeee).

Otro de los valores que compone un elemento de datos es el valor de representación (VR), el cual es una cadena de caracteres de dos bytes que define el tipo y el formato del valor o valores del elemento. La presencia del valor depende de la sintaxis de transferencia, la cual recoge un conjunto de reglas que declaran las técnicas de codificación que las aplicaciones son capaces de soportar.



| REPRESENTACIÓN DE VALOR | SIGNIFICADO                    |
|-------------------------|--------------------------------|
| AE                      | Entidad de aplicación          |
| AS                      | Cadena de edad                 |
| AT                      | Etiqueta de atributo           |
| CS                      | Cadena de código               |
| DA                      | Fecha                          |
| DS                      | Cadena decimal                 |
| DT                      | Fecha y hora                   |
| FL                      | Punto flotante único (4 bytes) |
| FD                      | Punto flotante doble (8 bytes) |
| IS                      | Cadena entera                  |
| LO                      | Cadena larga                   |
| LT                      | Texto largo                    |
| OB                      | Otra cadena de bytes           |
| OF                      | Otra cadena flotante           |
| OW                      | Otra palabra                   |
| PN                      | Nombre de la persona           |
| SH                      | Cadena corta                   |
| SL                      | Firmado largo                  |
| SQ                      | Secuencia de elementos         |
| SS                      | Firmado corto                  |
| ST                      | Texto corto                    |
| TM                      | Hora                           |
| UI                      | Identificador único (UID)      |
| UL                      | Largo sin firmar               |
| UN                      | Desconocido                    |
| US                      | Corto sin firmar               |
| UT                      | Texto ilimitado                |

Tabla 3.1.- Representaciones de valores DICOM

El campo 'Valor' está compuesto por un número par de bytes que contienen el valor del dato almacenado. La longitud del valor especificado se encuentra almacenado en el campo 'Longitud', que contiene un entero sin signo 16 o 32 bits. Este número no incluye la longitud de la etiqueta, del valor de representación ni del propio campo de longitud.

|                                     |  |
|-------------------------------------|--|
| <b>Preámbulo</b>                    | Tamaño: 128 bytes (puestos al valor 00h).  |
|                                     | Prefijo: 4 bytes con la cadena de caracteres "DICM". Permite diferenciar si un fichero es DICOM o no.                |
| <b>Etiqueta</b>                     | N.º de grupo: identifica el elemento dentro del fichero DICOM. Pueden existir más de un elemento con el mismo valor. |
|                                     | N.º de elemento: identifica el elemento dentro del grupo   |
| <b>VR (Valor de Representación)</b> | Valor de representación: la forma en la que está almacenado el contenido del elemento.                               |
| <b>Longitud</b>                     | 16 o 32 bits enteros, indicando la longitud del valor almacenado.  |
| <b>Valor</b>                        | Valor del elemento de datos, puede contener información vinculada al fichero (información sobre el paciente).        |

Tabla 3.2.- Resumen de la estructura de un archivo DICOM

### 3.1.3.1. TIPOS DE ELEMENTOS DE DATO

Un atributo puede ser necesario o no para el manejo y la clasificación de un DICOM y por ello, se establecen diferentes tipos de atributos en función de su obligatoriedad:

- **Tipo 1:** Son elementos de dato obligatorios, es decir, el campo 'Valor' debe tener un dato valido y el valor del campo 'Longitud' no podrá ser cero. La ausencia de un valor valido, se considera una violación del protocolo.
- **Tipo 1C:** Son elementos obligatorios bajo ciertas condiciones específicas. En caso de que estas condiciones se cumplan, estos elementos tendrán los mismos requisitos que los elementos de datos de tipo 1
- **Tipo 2:** Son elementos de dato obligatorios. No obstante, el valor del elemento puede ser desconocido, la longitud puede ser cero y el campo 'Valor' puede estar vacío, teniendo en cuenta, que la etiqueta de este elemento si debe de estar presente.
- **Tipo 2C:** En caso de que se cumplan ciertos requisitos tendrán la misma función que los elementos de datos de tipo 2.
- **Tipo 3:** Estos elementos de dato son optativos. Pueden estar ausentes en un set de datos y no influyen en la dinámica del protocolo.

Se realiza un estudio más concreto de los tipos de elementos de datos en el apartado de anonimización y seudonimización de los archivos DICOM, ya que hay que tener en cuenta que etiquetas son de qué tipo para el reemplazo y la eliminar los datos

### 3.1.3.2. METADATOS DE UN ARCHIVO DICOM

Los archivos DICOM contienen metadatos que proporcionan información sobre los datos de imagen, como el tamaño, las dimensiones, la profundidad de bits, la modalidad utilizada para crear los datos y la configuración del equipo utilizado para capturar la imagen. Cada elemento de los datos clínicos tiene unas características específicas que pueden etiquetarse con metadatos para que la información sea más accesible, útil y completa. Es decir, todo detalle que pueda ser útil para la búsqueda y comprensión de los datos clínicos puede codificarse en un metadato.

De esta forma, establece una jerarquía entre los datos que permite realizar una clasificación según el contenido de la información por grupos, facilitando la identificación, el acceso a las variables y los parámetros de interés dentro de un mismo archivo.

Se pueden identificar tres niveles principales de metadatos necesarios para aumentar la eficiencia de un sistema de almacenamiento que permita a todas las partes involucradas acceder y colaborar:

- **Metadatos a nivel de paciente**

Son metadatos que relacionan a los pacientes con cada archivo de datos clínicos. Tener toda la información demográfica asociada a cada expediente clínico hace que sea mucho más significativo para los usuarios durante todo el proceso de atención, desde las compañías de seguros hasta los investigadores, los profesionales médicos y los administradores.

| Grupo (número de grupo) | Elemento (número de elemento) | Tag Description      | Explicación china                | VR |
|-------------------------|-------------------------------|----------------------|----------------------------------|----|
| 0010                    | 0010                          | Patient's Name       | Nombre del paciente              | PN |
| 0010                    | 0020                          | Patient              | ID ID del paciente               | LO |
| 0010                    | 0030                          | Patient's Birth Date | Fecha de nacimiento del paciente | DA |
| 0010                    | 0032                          | Patient's Birth Time | Hora de nacimiento del paciente  | TM |
| 0010                    | 0040                          | Patient's Sex        | Sexo del paciente                | CS |
| 0010                    | 1030                          | Patient's Weight     | Peso del paciente                | DS |
| 0010                    | 21C0                          | Pregnancy Status     | Estado de embarazo               | US |

Figura 3.4. - Ejemplos de etiquetas a nivel de paciente

#### Metadatos a nivel de examen

El siguiente nivel de metadatos recopila información relacionada con cada visita del paciente. Los metadatos adquiridos en el momento de la solicitud o la cita, junto con los metadatos del paciente, se pueden usar para definir y administrar los flujos de trabajo departamentales. Además, cada examen puede contener múltiples objetos de datos, como imágenes, vídeos y documentos, adquiridos utilizando diferentes métodos y dispositivos.

Los metadatos a nivel de examen pueden incluir información general, como el nombre del paciente, evento médico o número de registro, el médico principal, departamento y procedimiento, así como datos adicionales como el tipo de estudio, número de orden, modalidad y ubicación, descripción del examen, etc.

Si bien los departamentos que trabajan con imágenes para diagnóstico recogen diferentes cantidades de metadatos a nivel de examen, nunca está de más volver a evaluar el tipo de metadatos que se necesitaría al realizar la exportación a un sistema de archivo y recuperación central para toda la organización.

| Group | Elemen | Tag Description   | Explicación china  | VR |
|-------|--------|---|--|----|
| 0008  | 0050   | Accession Number: A RIS generated number that identifies the order for the Study. | Número de verificación: el número de serie de RIS para identificar el orden de verificación.   | SH |
| 0020  | 0010   | Study ID  | Verifique la identificación.   | SH |
| 0020  | 000D   | Study Instance UID: Unique identifier for the Study.                              | Número de instancia de inspección: marque de forma única el número de inspecciones diferentes. | UI |
| 0008  | 0020   | Study Date: Date the Study started.   | Fecha de inspección: la fecha en que comenzó la inspección.                                    | DA |
| 0008  | 0030   | Study Time: Time the Study started.   | Hora de inspección: la hora en que comenzó la inspección.                                      | TM |
| 0008  | 0061   | Modalities in Study   | Los diferentes tipos de inspección contenidos en una inspección.                               | CS |
| 0008  | 0015   | Body Part Examined  | El sitio que se va a comprobar.  | CS |
| 0008  | 1030   | Study Description   | La descripción del cheque.   | LO |
| 0010  | 1010   | Patient's Age   | La edad del paciente en el momento del examen, no la edad real del paciente en ese momento.    | AS |

Figura 3.5.- Ejemplos de etiquetas a nivel de estudio

| Group | Element | Tag Description  | Explicación china   | VR |
|-------|---------|--|---|----|
| 0020  | 0011    | Series Number:A number that identifies this Series.  | Número de serie: el número para identificar diferentes inspecciones.  | IS |
| 0020  | 000E    | Series Instance UID:Unique identifier for the Series.  | Número de instancia de serie: un número que marca de forma única diferentes secuencias.   | UI |
| 0008  | 0060    | Modality   | Modo de inspección (MRI / CT / CR / DR)   | CS |
| 0008  | 103E    | Series Description   | Ver descripción e instrucciones   | LO |
| 0008  | 0021    | Series Date  | Comprobar Fecha   | DA |
| 0008  | 0031    | Series Time  | mira la hora  | TM |
| 0020  | 0032    | Image Position (Patient): The x, y and z coordinates of the upper left hand corner of the image, in mm.              | Posición de la imagen: las coordenadas x, y, z de la esquina superior izquierda de la imagen en el sistema de coordenadas espaciales, en milímetros. Si está bajo inspección, se refiere a las coordenadas de la esquina superior izquierda de la primera imagen de la secuencia. | DS |
| 0020  | 0037    | Image Orientation (Patient):The direction cosines of the first row and the first column with respect to the patient. | Orientación de la imagen  | DS |
| 0018  | 0050    | Slice Thickness:Nominal slice thickness, in mm.  | Grosor de la capa.  | DS |
| 0018  | 0088    | Spacing Between Slices   | La distancia entre capas, en mm   | DS |
| 0020  | 1041    | Slice Location: Relative position of exposure expressed in mm.   | La posición relativa real en mm.  | DS |
| 0018  | 0023    | MR Acquisition   | Adquiridor  | CS |
| 0018  | 0015    | Body Part Examined   | partes del cuerpo.  | CS |

Figura 3.6.- Ejemplos de etiquetas a nivel de serie

- **Metadatos a nivel de objeto**

El tercer nivel de metadatos identifica y describe los objetos específicos creados durante la visita de un paciente.

Cualquier examen puede producir diferentes objetos como imágenes, videos y documentos. Todos estos objetos se pueden asociar con pacientes y procedimientos individuales, pero cada objeto tiene características únicas que se pueden identificar a través de metadatos a nivel de objeto para una mayor búsqueda, visualización y análisis.

Por ejemplo, los metadatos a nivel de objeto pueden incluir información como el dispositivo usado para capturar la imagen, la configuración del dispositivo, la posición de la cámara, la iluminación, la parte del cuerpo que se muestra, lado derecho o izquierdo, ángulo de visión, y otra información útil que podría no ser tan obvia en la imagen.

| Group | Element | Tag Description   | Explicación china   | VR   |
|-------|---------|---|---|------|
| 0008  | 0008    | Image Type:Image identification characteristics.  | .   | CS   |
| 0008  | 0018    | SOP Instance UID  | Ejemplos de POE   | UID. |
| 0008  | 0023    | Content Date: The date the image pixel data creation started.   | La fecha en que se tomó la imagen.  | DA   |
| 0008  | 0033    | Content Time  | La hora a la que se tomó la imagen.   | TM   |
| 0020  | 0013    | Image/Instance Number:A number that identifies this image.  | Código de imagen: el número para identificar la imagen.   | IS   |
| 0028  | 0002    | Samples Per Pixel:Number of samples (planes) in this image.   | La frecuencia de muestreo de la imagen.   | US   |
| 0028  | 0004    | Photometric Interpretation:Specifies the intended interpretation of the pixel data.   | La explicación del fotómetro, para imágenes de TC, usa dos valores enumerados MONOCHROME1, MONOCHROME2 Se usa para juzgar si la imagen es en color, MONOCHROME1 / 2 es una imagen en escala de grises, RGB es una imagen en color verdadero y otros.      | CS   |
| 0028  | 0010    | Rows: Number of rows in the image.  | El número total de líneas de la imagen, resolución de línea.  | US   |
| 0028  | 0011    | Columns: Number of columns in the image.  | El número total de columnas de imagen, resolución de columna.   | US   |
| 0028  | 0030    | Pixel Spacing:Physical distance in the patient between the center of each pixel.  | Tamaño de píxel: la distancia física entre los centros de píxeles.  | DS   |
| 0028  | 0100    | Bits Allocated:Number of bits allocated for each pixel sample. Each sample shall have the same number of bits allocated.  | Número de bits asignados: el número de bits asignados al almacenar cada valor de píxel. Cada muestra debe tener el mismo valor.   | US   |
| 0028  | 0101    | Bits Stored:Number of bits stored for each pixel sample. Each sample shall have the same number of bits stored  | Número de bits almacenados: hay valores enumerados de 12 a 16. El número de bits utilizados para almacenar cada píxel, cada muestra debe tener el mismo valor.  | US   |
| 0028  | 0102    | High Bit:Most significant bit for pixel sample data. Each sample shall have the same high bit.  | Alto.   | US   |
| 0028  | 0103    | Pixel Representation:Data representation of the pixel samples. Each sample shall have the same pixel representation.Enum: 0000H=unsigned integer,0001H=2's complement.  | El tipo de representación de datos de píxeles: este es un valor enumerado, respectivamente números hexadecimales 0000 y 0001. 0000H = entero sin signo, 0001H = complemento a 2.  | US   |
| 0028  | 1050    | Window Center   | Nivel de ventana.   | DS   |
| 0028  | 1051    | Window Width  | Ancho de ventana.   | DS   |
| 0028  | 1052    | Rescale Intercept:The value b in relationship between stored values (SV) and the output units. Output units = m*SV + b. Required if Modality LUT Sequence (0028, 0030) is not present.  | Intercepción: Si muestra que la tabla de correspondencia de colores LUT de diferentes modos no existe, use la ecuación Unidades = m * SV + b para calcular el valor de píxel real para el valor de píxel mostrado. Donde este valor es b en la expresión. | DS   |
| 0028  | 1053    | Rescale Slope:m in the equation specified by Rescale Intercept (0028,1052). Required if Rescale Intercept is present.   | Pendiente: este valor es m en la expresión.   | DS   |
| 0028  | 1054    | Rescale Type:Specifies the output units of Rescale Slope (0028,1053) and Rescale Intercept (0028,1052). Enum: US=Unspecified Required if Photometric Interpretation is MONOCHROME2, and Bits Stored is greater than 1.This specifies an identity Modality LUT transformation. | La unidad del valor de salida. Este es un valor enumerado,  | LO   |

Figura 3.7.- Ejemplos de etiquetas a nivel de instancia

A todo esto, cabe mencionar la existencia de metadatos DICOM privados. La especificación DICOM define muchos de los campos de metadatos mencionados, pero los archivos pueden contener campos adicionales, denominados metadatos privados. Estos metadatos privados suelen ser definidos por los proveedores de equipos para proporcionar información adicional sobre los datos que proporcionan.

Para el desarrollo de este proyecto, se han identificado 4 niveles jerárquicos en cuestión a los metadatos que presenta un archivo DICOM: paciente, estudio, serie e instancia.

Se interpreta al paciente como la persona que recibe el examen y a su vez, al estudio como el examen que se realiza a un paciente. Se trata de un procedimiento de imágenes que se realiza en un hospital, en una fecha y hora determinadas. Por otro lado, identificamos varias series por cada estudio realizado. Una serie puede representar al paciente que ha sido escaneado físicamente varias veces en un mismo estudio o puede representar al paciente que ha sido escaneado una única vez y los datos se han reconstruido de diferentes maneras. Por último, las series están formadas por un número de instancias, que representan cada segmento de una imagen. En este contexto, una instancia es sinónimo del propio archivo DICOM.

En definitiva, un paciente puede contener uno o mas estudios, que pueden estar formados por una o mas series y cada una de esas series esta compuesta por una imagen o un conjunto de imágenes.

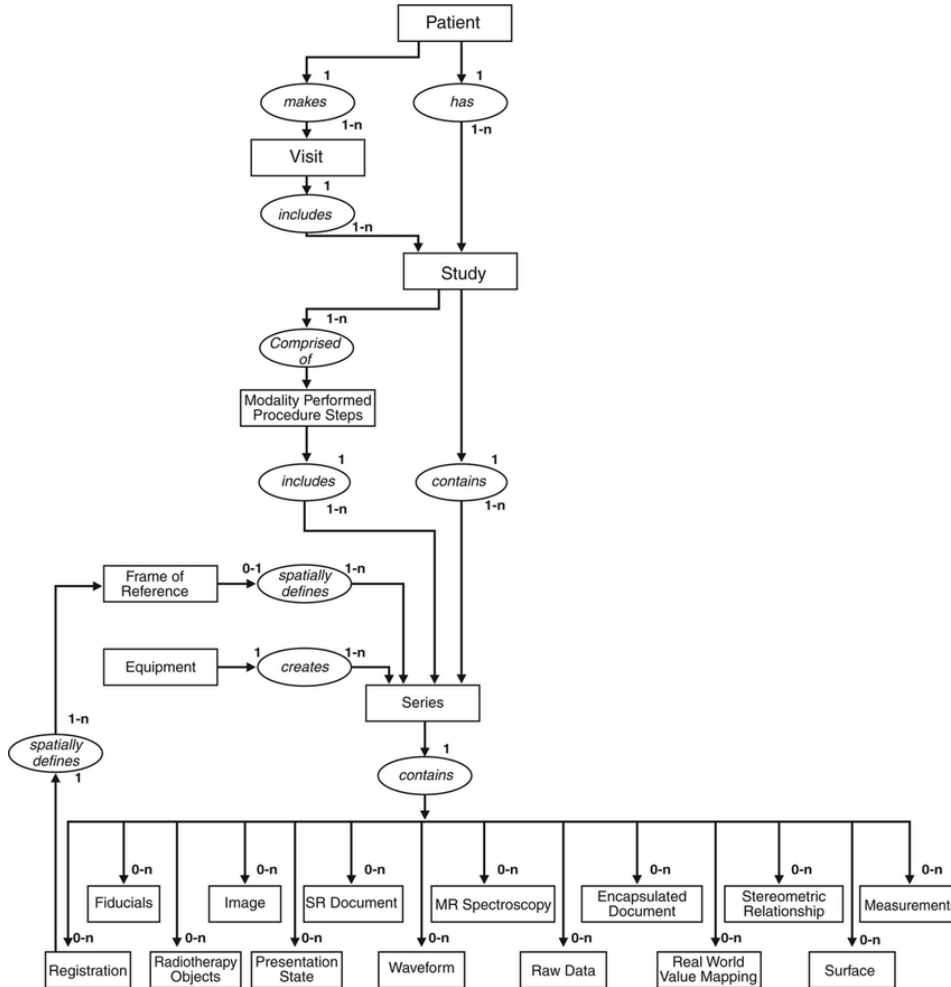


Figura 3.8.- Diagrama de entidad-relación. [7]

## 3.2. SISTEMA PACS

---

Una parte de los objetivos del proyecto consiste en implementar un sistema de procesamiento que permita realizar peticiones y transferencias de archivos médicos desde otros PACS y búsquedas de estudios en PACS. Para cumplir con ello es de vital importancia que se comprenda la funcionalidad de un sistema PACS.

Un sistema PACS combina hardware y software dedicado al almacenamiento, recuperación, gestión, distribución y presentación digital de imágenes médicas incluido la transmisión de estas a estaciones de visualización dedicadas, a través de una red informática. Una red PACS consiste en un servidor central con una base de datos que contiene las imágenes en formato DICOM. Todos estos procesos tienen lugar mediante el uso de elementos de servicio (SOP) como C-STORE, C-FIND, C-MOVE, entre otros. Cada estación en la red se identifica por su dirección de red (dirección IP), un puerto de comunicación (TCP / IP) y un nombre de entidad (AETitle), parámetros con la información necesaria para distinguir cada nodo de un sistema PACS.

Los 4 usos principales que caracterizan la implementación de sistemas PACS son:

- Reemplaza el archivado físico: los PACS ofrecen un costo menor en comparación con los sistemas de archivado.
- Acceso remoto: mediante el uso de los PACS se amplían las posibilidades de los sistemas convencionales, proporcionando capacidades de visualización y presentación de informes de forma remota (telemedicina).
- Plataforma de integración de imágenes electrónicas: los PACS proporcionan la plataforma electrónica para la interfaz de imágenes médicas con otros sistemas como el sistema de información hospitalaria (HIS), Electrónica de Sistemas de Registros Médicos (EMR), Sistemas de Gestión de Prácticas y Sistemas de Información de Radiología (RIS).
- Administración de flujo de datos: debido a que la funcionalidad de los PACS es gestionar y manejar estudios de pacientes, requieren de una eficiente administración de flujo de datos.

Como ha sido mencionado anteriormente, entre los diferentes tipos de componentes software y hardware que son usados por los sistemas PACS destacan:

- Adquisición de imágenes:

La adquisición de imágenes se basa en dos modalidades: por un lado, está el caso en el que exista una gran cantidad de tipos de imagen que ya se encuentran en formato digital debido a su naturaleza y, por otro lado, el caso de las imágenes disponibles únicamente en placa, las cuales se requieren una digitalización manual mediante un digitalizador de placas (scanner) para que puedan ser almacenadas en un sistema PACS.

- Una red de comunicación

Es un componente imprescindible de los sistemas PACS. Por lo general, se cuenta con una serie de elementos de distintas velocidades de acceso, aunque también puede ser una red tipo Ethernet. La instalación de una red de alta velocidad depende directamente de la demanda de ancho de banda, debido a que la mayor parte del tráfico de información está dentro de una misma unidad de imagenología.

- Bases de datos

Principalmente el buen funcionamiento de un sistema PACS está basado en el diseño y la implementación de un sistema de bases de datos. Las consultas a las imágenes son más frecuentes en las horas siguientes a la adquisición de esta y es por ello por lo que el almacenamiento a corto plazo (pocas horas) debe hacerse en los sistemas locales (memoria y disco). En cambio, a largo plazo puede hacerse ya sea en unidades de disco óptico o en cinta magnética. La base de datos ofrece la posibilidad de almacenar tanto imágenes como voz y texto.

- Estaciones de diagnóstico y visualización

Un sistema PACS presenta la información visual mediante estaciones de diagnóstico y visualización cumpliendo con los requerimientos de calidad adecuados. Por otro lado, requieren una incorporación de funciones básicas de procesamiento como el cambio de contraste, acercamientos, ecualización de histogramas, despliegue en 3D, registro y filtrado.

- Sistemas de almacenamiento

Los sistemas de almacenamiento de imágenes deben estar basadas en una estructura jerárquica dependiente de la demanda de la imagen.

El almacenamiento a corto plazo tiene los siguientes requisitos: gran cantidad de GB, transferencia de alrededor de 50 imágenes por minuto, 1-15 días de almacenamiento.

El almacenamiento a largo plazo cumple con las siguientes características: capacidad de varios Terabytes, empleo de robots o "jukeboxes" de discos ópticos, capacidad mínima de almacenamiento de dos años, empleo de cinta e imágenes comprimidas.

- Impresión

Es necesario contar con al menos una terminal de impresión para realizar copias radiográficas y copias en papel. A menudo, es necesario imprimir estudios de radiología y eso conlleva a que se debe tener el modo de ordenador la impresión.



### 3.2.1. ORTHANC – DICOM SERVER

La infraestructura de almacenamiento del proyecto está basada en el despliegue de un sistema PACS, para el almacenamiento ordenado de datos DICOM y la comunicación con otros PACS. Orthanc DICOM Server es el sistema empleado, ya que se trata de un servidor de código abierto y está diseñado en beneficio de hospitales e investigadores. Mediante Orthanc, realizamos la extracción de estudios a la aplicación desde un PACS hospitalario, la que a su vez realiza llamadas a otro Orthanc intermediario para una búsqueda y una manipulación de datos.



Figura 3.9.- Logo de Orthanc – Dicom Server [21]

Se podría definir Orthanc como un servidor DICOM, de código abierto y liviano para investigación y atención médica. Está diseñado principalmente para facilitar secuencias de comandos DICOM para la rutina clínica, para conseguir que los flujos DICOM sean mejores y para la gestión de datos para la rutina clínica y la investigación médica (mini-PACS). Otra de sus funciones es facilitar el análisis automatizado de imágenes médicas, llevando imágenes DICOM a la comunidad de Computer Vision. Para cumplir con los objetivos mencionados, Orthanc es ligero, rápido, escrito en C++ e independiente, ya que las dependencias se pueden vincular estáticamente.

Es una multiplataforma que cumple con el estándar DICOM, ya que está construido sobre DCMTK, colección de bibliotecas y aplicaciones que implementan gran parte del estándar DICOM. Tiene la incorporación de un software para examinar, construir y convertir archivos de imágenes DICOM, manejar medios fuera de línea, enviar y recibir imágenes a través de una conexión de red, así como servidores de listas de trabajo y almacenamiento de imágenes demostrativas. Por todo ello, por ejemplo, llegamos a la conclusión de que un archivo de formato NIFTI no tiene compatibilidad de uso con Orthanc, ya que carece de cierta información que necesita DICOM.

#### 3.2.1.1. API REST de Orthanc

Lo que hace único a Orthanc es el hecho de que cuenta con una API RESTful. Gracias a esta característica principal, se proporciona la gran ventaja de poder manejar Orthanc desde cualquier lenguaje de programación y desde aplicaciones externas. La API REST de Orthanc ofrece un acceso programático completo a todas las características principales de Orthanc. Orthanc Explorer, interfaz administrativa integrada de Orthanc, recurre completamente a esta API REST para todas sus funcionalidades. Esto implica que todo lo que se hace a través de Orthanc Explorer, es posible también hacerlo a través de consultas REST.

Entre las funcionales que ofrece la API, es destacable la posibilidad de transferir datos médicos a través de Internet utilizando HTTP. La aplicación web hace uso de estas llamadas para varias funcionalidades principales:

- Las etiquetas DICOM de las imágenes médicas almacenadas se descargan en formato de archivo JSON.
- Crear de forma asíncrona un archivo ZIP que contiene el estudio DICOM cuyo identificador Orthanc se proporciona en la URL. La carpeta ZIP esta organizada en subcarpetas en función del estudio y de la/s series. Se obtiene las instancias de cada estudio organizadas por series, en formato DICOM con extensión '.dcm'.
- Cargar instancias DICOM

Se estudiará con más exactitud el uso de las funcionalidades empleadas en el código de la aplicación mas adelante, empleando diferentes ejemplos.

## 3.3. ANONIMIZACIÓN Y SEUDONIMIZACIÓN DE UN ARCHIVO DICOM

---

El módulo de ingesta de datos provee varias funcionalidades de las cuales destaca la anonimización o seudonimización de un estudio de imágenes. Además de las imágenes, el formato DICOM incluye metainformación identificativa relativa al paciente como el nombre y apellidos o el número del historial clínico del paciente. De esta manera, conseguimos mantener el anonimato de la información eliminando y reemplazando dicha información previamente. Esta característica abre multitud de posibilidades relacionadas con la importación de información desde ficheros generados por cualquier equipo o dispositivo médico.

La anonimización y la seudonimización son medidas de seguridad que garantizan la protección de datos de carácter personal, especialmente cuando hablamos de datos sensibles, como por ejemplo, los relativos a la salud.

### 3.3.1. ASPECTOS GENERALES

La función que cumple un proceso de anonimización es la de alterar de forma irreversible una serie de datos, desvinculando los datos personales de los datos identificativos. De esta manera, se genera un conjunto de datos completamente disociado del paciente al que pertenecen.

Por lo tanto, mediante la anonimización se debe producir una ruptura de la cadena de identificación. Esta cadena está compuesta por datos de identificación directa y de datos de identificación indirecta. Mientras que los microdatos identifican a un paciente de manera directa, los datos de identificación indirecta son datos cruzados que provienen de la misma o de diferentes fuentes que permiten una posible reidentificación de las personas, como por ejemplo desde la información de otras bases de datos, redes sociales, etc. De acuerdo con lo estipulado por la AEPD, Agencia Española de Protección de Datos, se deben de considerar varios principios:

- Principio proactivo: Se deben de tener en cuenta las medidas imprescindibles que garanticen la privacidad de las personas desde el diseño del sistema de información.
- Principio de privacidad por defecto: Se debe de tener en cuenta el grado de detalle final que deben tener los datos anonimizados.
- Principio de privacidad objetiva: El riesgo que existe tras el proceso de anonimización ha de ser conocido y asumido por el responsable a la hora de diseñarlo.
- Principio de plena funcionalidad: Se debe de tener en cuenta la utilidad final de los datos que han sido modificados, de manera que se garantice lo máximo posible la inexistencia de distorsión.

- Principio de información y formación: Es de vital importancia que la persona responsable de llevar a cabo el proceso sea plenamente consciente de las implicaciones de mismo. Se deben de contemplar las necesidades formativas desde el diseño conceptual.

Por otro lado, en cuanto a la seudonimización, se trata de un proceso que realiza un cambio de datos original por un alias o seudónimo y conserva de forma segura un conjunto de correspondencias entre la identidad real y la identidad seudónima del paciente. Por ejemplo, cambiar el nombre y apellidos de un paciente por una clave numérica. Solo quien posee la información adicional que establece el vínculo entre los nuevos datos y los originales, puede conseguir identificar a la persona. Este proceso fue introducido por el RGPD y su adaptación en LOPDGDD como una medida válida para la protección de datos.

Existen varias técnicas para la seudonimización y entre las más habituales se encuentran:

- Cifrado con clave secreta: se utiliza una clave secreta para cifrar un conjunto de datos, de forma que solo quien posee dicha clave pueda identificar a los pacientes descifrando la base de datos.
- Función con clave almacenada: Consiste en emplear una función hash que utiliza una clave secreta como valor de entrada, de manera que para acceder a los datos se debe reproducir la ejecución de la función con el atributo y la clave secreta.
- Cifrado determinista o función hash con clave de borrado de clave: Se genera un número aleatorio para usar como seudónimo de cada dato personal y posteriormente borra la tabla de correspondencia. Esto conlleva a reducir los riesgos de vincular los datos personales con los conjuntos de datos.
- Descomposición en tokens: Esta técnica esta basada en las anteriores y consiste en utilizar mecanismos de cifrado unidireccionales o en asignar un número de secuencia o número generado aleatoriamente y que no derive matemáticamente de los datos originales.

Aunque se utilicen con una finalidad parecida, la anonimización y la seudoanonimización son dos procesos diferentes. La gran diferencia recae en el vínculo que se guarda con la identidad original del paciente. Como ha sido mencionado anteriormente, la anonimización destruye cualquier vínculo con los datos originales del paciente mientras que en la seudonimización conserva de forma segura un conjunto de correspondencias entre la identidad real y la seudónima del paciente. Son ambas las posibilidades que se ofrecen en la aplicación, donde será el cliente quien elija si desea anonimizar o seudonimizar un estudio o una serie de estudios.

Cabe mencionar la existencia de HIPAA, Ley de Portabilidad y Responsabilidad de Seguros de Salud de los EE. UU, que obliga a las entidades que trabajan en el manejo de información médica protegida (PHI) a implementar medidas de seguridad físicas, garantizando el derecho a la privacidad y confidencialidad de la información de salud. Esta ley por ejemplo, prescribe dos opciones para la desidentificación de datos médicos: la eliminación de 18 identificadores de puerto o la certificación por parte de un experto de que el riesgo de reidentificación a partir de los datos es bajo.

## 3.3.2. EJEMPLOS DE CÓDIGO

### a. Ejemplo de anonimización en Orthanc

Orthanc 0.5.0 introduce la anonimización de los recursos DICOM.

- De una sola instancia [15]

Orthanc ofrece la posibilidad de anonimizar una única instancia DICOM y posteriormente poder descargar el archivo DICOM anónimo resultante. Sin embargo, no se almacena la instancia DICOM anonimizada:

```
$ curl http://localhost:8042/instances/6e67da51-d119d6ae-c5667437-87b9a8a5-0f07c49f/anonymize -X POST -d '{} ' > Anonimizado.dcm
```

Se controla el resultado de la anonimización especificando un cuerpo JSON.

```
$ curl http://localhost:8042/instances/6e67da51-d119d6ae-c5667437-87b9a8a5-0f07c49f/anonymize -X POST \
--data '{
    "Replace": { //asociar un nuevo valor a las etiquetas DICOM
        "PatientName": "Hello",
        "0010-1001": "World"
    },
    "Keep": [ //etiquetas que se quieren mantener
        "StudyDescription",
        "SeriesDescription"
    ],
    "KeepPrivateTags": true,
    "DicomVersion" : "2017c" //versión del estándar DICOM
}' > Anonymized.dcm
```

Todas las etiquetas que se anonimizarán deberán eliminarse antes de aplicar la superposición. Para asociar una etiqueta DICOM con su nuevo valor o agregar una nueva etiqueta al archivo, se utiliza la matriz asociativa *REPLACE*. Las etiquetas DICOM se pueden especificar por su nombre o identificador hexadecimal (PatientName, 0010-1001).

Con *Keep*, se especifica una lista de etiquetas para evitar la anonimización total, es decir, se indican las etiquetas que no se quieren ni cambiar ni eliminar. Para que las etiquetas específicas del fabricante, es decir, las etiquetas privadas, no se eliminen mediante el proceso de anonimización, se configura "*KeepPrivateTags*": *true*. El comportamiento predeterminado es eliminar las etiquetas privadas, ya que las etiquetas privadas pueden contener información específica del paciente.

*DicomVersion* especifica la versión del estándar DICOM utilizado para la anonimización. Los valores permitidos son 2008, 2017c o 2021b. Si el parámetro no está presente, se utiliza la versión más alta compatible con Orthanc.

*Remove* también se puede usar para proporcionar una lista de etiquetas para eliminar manualmente.

### b. Ejemplo de anonimización en Python

Mediante el siguiente código, se editan las etiquetas que contienen los ficheros DICOM mediante el lenguaje de programación Python y de una librería denominada *pydicom*. Es decir, se trata de editar los tags que contienen información personal del sujeto.

```

#importamos librerias necesarias
import matplotlib.pyplot as plt
from matplotlib import pylab
import pydicom

#indicamos el fichero con el que vamos a trabajar
filename = "21560.dcm"

#Leemos fichero dicom
file = pydicom.dcmread(filename)

#editamos las etiquetas del fichero
file.PatientName = "anonymized"
file.PatientID = ""
file.PatientBirthDate = ""
file.PatientSex = ""
file.PatientAge = ""
file.PatientWeight = ""
file.PatientAddress = ""

# guardamos los cambios realizados en el fichero
file.save_as(filename)
  
```

### c. Ejemplo de anonimización en Matlab

La sintaxis que utiliza Matlab para anonimizar un archivo DICOM sería la siguiente:

- **[dicomanon\(file in,file out\)](#)**: elimina los datos médicos confidenciales del archivo DICOM especificado en [file in](#) y crea un nuevo archivo [file out](#) con la información modificada, dejando el resto de información (datos de imagen y otros atributos) sin cambios.
- **[dicomanon\(\\_\\_,'keep',fields\)](#)**: modifica todos los datos confidenciales, excepto los enumerados en la matriz de [fields](#), que es un array de celdas con nombres de campo. Esta sintaxis es útil para mantener metadatos que no identifican de forma exclusiva a un paciente pero que son útiles para el diagnóstico. Tenga en cuenta que mantener ciertos campos puede comprometer la confidencialidad del paciente.
- **[dicomanon\(\\_\\_,'update',attributes\)](#)**: modifica los datos confidenciales manteniendo ciertos campos y actualizando otros campos confidenciales particulares. [Attributes](#) es la estructura cuyos campos son los nombres de

los atributos a conservar. Usar esta sintaxis para mantener Estudios/Series/Imágenes o para reemplazar un valor específico con una propiedad más genérica (p.ej., eliminar la fecha de nacimiento del paciente (*PatientBirthDate*) manteniendo su edad (*PatientAge*).

- **`dicomanon( __,Name,Value)`**: proporciona opciones adicionales al analizador mediante pares de nombre-valor.

### i. Ejemplos de uso [17]

1.- Eliminar todos los metadatos confidenciales de un archivo DICOM y crear una versión de un archivo DICOM con toda la información personal.

```
dicomanon('US-PAL-8-10x-echo.dcm', 'US-PAL-anonymized.dcm');
```

2.- Eliminar todos los datos confidenciales de un archivo DICOM y crear una nueva versión con información personal, manteniendo ciertos campos que podrían ser útiles para el entrenamiento.

```
dicomanon('US-PAL-8-10x-echo.dcm', 'US-PAL-anonymized.dcm', 'keep', ...
{'PatientAge', 'PatientSex', 'StudyDescription'})
```

3.- Anonimiza una serie de imágenes, manteniendo la jerarquía.

```
values.StudyInstanceUID = dicomuid;
values.SeriesInstanceUID = dicomuid;
d = dir('*.dcm');
for p = 1:numel(d)
    dicomanon(d(p).name, sprintf('anon%d.dcm', p), ... 'update',
values)
end
```

# Capítulo 4

## DATALAICE

### Contenido

---

- 4.1. Análisis y diseño de la aplicación
  - 4.2. Programas y herramientas
  - 4.3. Base de datos
  - 4.4. Proceso de anonimización/seudonimización
    - 4.4.1. Tratamiento de datos
    - 4.4.2. Desarrollo del proceso
  - 4.5. Funcionamiento e implementación
-



## 4.1. ANALISIS Y DISEÑO

La interfaz web DATALAICE que está basada en el módulo de ingesta de datos tiene como funcionalidad principal la conexión con PACS mediante el protocolo DICOM con el propósito de crear conjuntos de datos sobre los que llevar a cabo análisis o estudios.

El funcionamiento de la aplicación comienza con la extracción de estudios desde un PACS Hospitalario, que en este caso corresponde al de Vicomtech. El conjunto de imágenes medicas son seleccionadas mediante un filtro y la extracción se realiza en base al protocolo DICOM. Estas imágenes se almacenan en un PACS de Orthanc denominado como Middleware, desde el cual recuperamos las imágenes en la aplicación web.

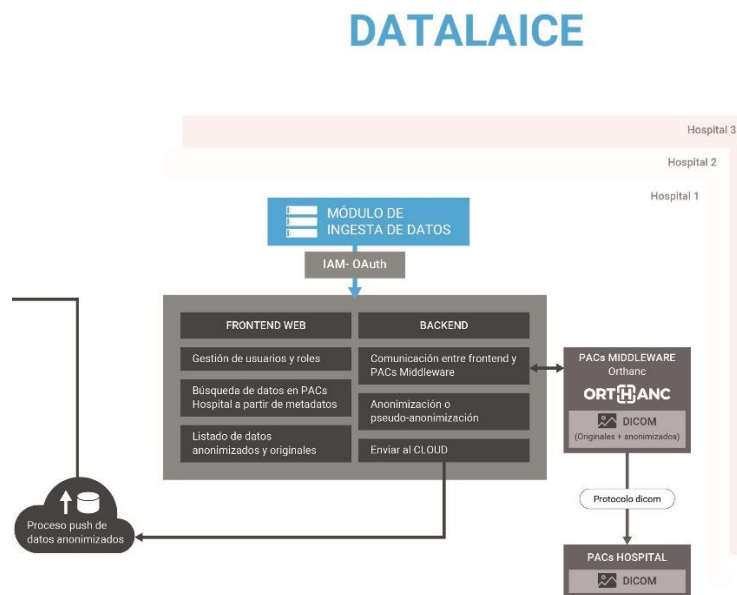


Figura 4.1.- Esquema del módulo de ingesta de datos

En los PACS existen pacientes, para los que se almacenan estudios, definidos como las series adquiridas de una sesión determinada. Cada serie se compone de una o mas instancias, en las que se almacena una imagen con unas propiedades espaciotemporales dadas. Esta jerarquía resulta muy eficiente para el almacenamiento y recuperación de los estudios de imagen médica, puesto que basta con saber el identificador de un paciente concreto para consultar los estudios disponibles en el sistema. En la aplicación web de DATALAICE, contamos con 3 pantallas diferentes en las que se puede diferenciar la búsqueda por pacientes, por estudios y por series.

Principalmente, DATALAICE se compone por estas funcionalidades:

## Extracción de estudios

La extracción de estudios consiste en conectar la aplicación a un PACS externo, en este caso al Orthanc de Vicomtech, y mediante un filtro de búsqueda recuperar estudios almacenados en él. Cuando tengamos los estudios deseados en nuestra interfaz, se podrán seleccionar los necesarios y descargarlos a nuestro PACS Orthanc Middleware para su gestión.

Es decir, la extracción se basa en la comunicación con otros PACS mediante estándares (DICOMWeb, HTTP) para realizar búsquedas y recuperación de estudios.

## Filtro de búsqueda

La interfaz web cuenta con un filtro de búsqueda en base a los pacientes, los estudios y las series. La comunicación con el PACS permite recuperar los estudios mediante llamadas HTTP, pudiendo acceder a los metadatos de cada DICOM y las búsquedas se realizan a partir de los siguientes metadatos:

- De paciente: ID, nombre y edad del paciente
- De estudio: ID, nombre, sexo y edad del paciente, descripción del estudio, modalidad, fabricante, fecha desde y fecha hasta
- De serie: ID de la serie, modalidad, fabricante, fecha desde y fecha hasta.

## Anonimización o seudonimización

El modulo de ingesta de datos tiene una herramienta de anonimización y seudonimización de los datos en función de las necesidades del usuario. La anonimización encripta o elimina todos los datos e identificadores únicos que pueden identificar unívocamente al paciente, siguiendo el formato definido por el estándar DICOM. Es de vital importancia que los datos encriptados no impidan el almacenamiento en la estructura, es decir que deben seguir cumpliendo el estándar DICOM.

Se clasifican los metadatos en función de su obligatoriedad, es decir, se tienen en cuenta cuales son los metadatos que deben ser encriptados por un nuevo valor, o cuales pueden dejarse vacíos o ser eliminados. La implementación se realiza en base a una librería de Python de código abierto denominada *pydicom*.

Mediante esta herramienta se anonimizan los datos recibidos de un PACS y se depositan en la infraestructura de almacenamiento, teniendo en cuenta que nunca se podrán almacenar datos no anonimizados.

## Gestión de proyectos

Los estudios y las series que se quieren anonimizar se almacenan en diferentes proyectos, controlando que un mismo DICOM no puede ser anonimizado dos veces para el mismo proyecto, pero si en diferentes proyectos. En caso de que un fichero DICOM haya sido anonimizado dos veces, pero estén almacenados en diferentes proyectos, estarán encriptados con valores nuevos diferentes, por lo que una vez que este anonimizado, se almacenaran como dos ficheros DICOM totalmente diferentes.

Cada proyecto es distinguido por la acción de anonimizar o seudonimizar. Es decir, cada proyecto será de un tipo, de anonimizar o seudonimizar, por lo que, por ejemplo, solo se podrán almacenar estudios anonimizados en un proyecto de tipo anonimizar.

### **Funcionalidad push**

Por otro lado, el modulo proporciona funcionalidades *push*, que sirven para enviar ficheros DICOM anonimizados al módulo de gestión de anotaciones CLOUD. Los ficheros DICOM que se desean anonimizar están organizados en proyectos, por lo que la funcionalidad *push* se hace por cada proyecto, envían los DICOM pertenecientes a él.

## 4.2. PROGRAMAS Y HERRAMIENTAS

---

Las empresas requieren de páginas webs rápidas, eficientes y con un diseño atractivo que se acomoden a las necesidades de los usuarios y clientes potenciales. Es por ello por lo que es de vital importancia saber aplicar las mejores herramientas de desarrollo web.

El desarrollo web consiste en crear, diseñar y mantener páginas y aplicaciones web y esto se lleva a cabo a través de distintos programas y softwares. Existen varios tipos de desarrollo web para tener en cuenta, como son el front-end, el back-end y el full-stack. Este proyecto esta basado en un tipo de desarrollo full-stack, donde se ejecutan las tareas de un front-end y un back-end en su totalidad.

Para el desarrollo de la aplicación web, se han utilizados diferentes programas y herramientas:

### **Visual Studio Code**

Como editor de código fuente independientes, se ha utilizado el programa Visual Studio Code. Es una plataforma lanzada por primera vez el 18 de noviembre de 2015 bajo la licencia MIT y su código fuente fue publicado en GitHub. Cuenta con soporte para operaciones de desarrollo como la depuración, la ejecución de tareas, control integrado de Git, refactorización de código y el control de versiones. Está programado en TypeScript, JavaScript y CSS, aunque es compatible con varios lenguajes de programación. Entre sus numerosas funcionalidades, destacan las siguientes:

- Ver y editar archivos locales. Incluso con en una maquina restringida donde no se puede instalar Visual Studio Code completo, se puede usar *vscode.dev* para ver y editar los archivos locales.
- Crear aplicaciones HTML, JavaScript y CSS del cliente y herramientas del navegador para la depuración.

### **Angular**

Para la tecnología front-end, se ha hecho uso de la plataforma de desarrollo Angular. Se trata de un framework de código abierto orientado principalmente a la creación de SPA (aplicaciones de una sola página) utilizando la arquitectura Modelo – Vista – Controlador. El desarrollo se hace mediante TypeScript, un superset de JavaScript que ofrece herramientas adicionales al lenguaje, como el tipado estático o los decoradores.

Para la construcción de aplicaciones Angular son necesarios plantillas HTML que contienen etiquetas adicionales especiales de Angular, componente de clase que gestionan las plantillas, servicios que encapsulas lógica de la aplicación y módulos que organizan los componentes y servicios.

En el caso de la aplicación web de DATALAICE, tenemos los componentes *extractor*, *studies*, *series*, *series-c*, *patients* y *projets*. Cada uno de ellos este compuesto por un archivo HTML, un CSS que define el estilo de la plantilla HTML y un archivo TypeScript entre otros y definen las pantallas y el funcionamiento y la gestión de estas.

Por otro lado, tenemos el servicio principal *api-http.service*, en el cual se realizan las llamadas a la parte del servidor de la aplicación.

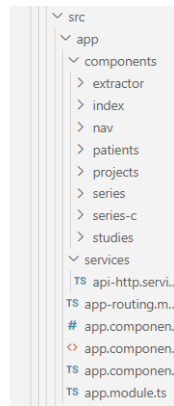


Figura 4.2.- Archivos de Angular

## Grails

Como back-end se hace uso del framework Grails, que esta basado en el lenguaje de programación Groovy, que a su vez esta basada la Plataforma Java. Emplea el patrón Modelo – Vista – Controlador. Además, hace uso de un lenguaje dinámico como Groovy para que los desarrolladores se centren únicamente en escribir código, actualizar, depurar y testear fallos y para reducir los tiempos de desarrollo.

La arquitectura de Grails esta definida por cuatro capas que están claramente diferenciadas. En la base de la arquitectura se encuentra una maquina virtual de java y por encima de ella, están los lenguajes de programación en los que esta basado Grails, Java y Groovy. Encima de estos, tenemos el propio framework Grails, con la opción abierta de poder incluir otras librerías externas. Por último, la ultima capa esta formada por la propia aplicación, en base al patrón modelo vista controlador.

En cuanto a Groovy, que es lenguaje que ha sido utilizado para la realización del proyecto, se trata de la parte fundamental en la que se basa Grails. Es un lenguaje de programación potente y flexible, que hace que con su integración con Java sea un lenguaje dinámico y su sintaxis sea sencilla.

## MySQL

Para el almacenamiento de los datos que se generan con uso de la aplicación, se hace uso de una base de datos relacional. En la presente práctica, se emplea el sistema de gestión de base de datos de código abierto MySQL como modo de recopilar la información distribuida y organizada en forma de tablas. Además, estas tablas guardan las relaciones entre ellas mediante claves y organizando la información real en filas y columnas.

Al estar basado en código abierto, permite disponer de una solución fiable y estandarizada para las aplicaciones.

## 4.3. BASE DE DATOS

Como ha sido comentado antes, el framework Grails está basado en el patrón Modelo – Vista – Controlador, donde los *modelos* se conocen como *clases de dominios* que permiten mostrar los datos mediante el uso de la *vista*. Una de las ventajas de este framework y de dichas clases de dominio, es la posibilidad que ofrecen de generar el esquema de la base de datos. Aprovechando esta funcionalidad, el modelo de la base de datos se crea desde las clases de dominio de Grails en esta práctica, siguiendo el siguiente esquema:

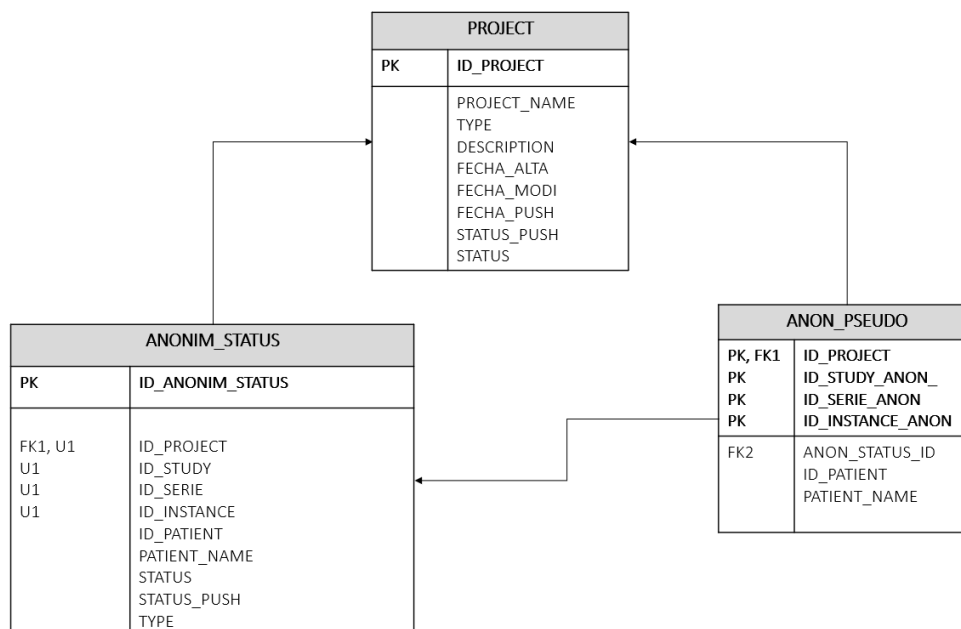


Figura 4.3.- Esquema de la base de datos

### PROJECT

La tabla PROJECT se encarga de albergar los diferentes proyectos que se crean con su respectiva información. Por cada proyecto se almacena un id, un nombre de proyecto y una descripción. La pantalla de proyectos en la aplicación ofrece la posibilidad de modificar la descripción, por lo que se guarda la fecha en la que ese proyecto ha sido modificado por última vez. Además, también se guarda la fecha en la que se ha dado de alta cada proyecto.

Por otro lado, cada proyecto es de un tipo concreto, que puede ser el valor ‘A’, que significa que las instancias de ese proyecto han sido anonimizadas, o el valor ‘P’, que significa que los estudios han sido seudonimizados.

Respecto a la herramienta *push*, cada proyecto tiene un *status\_push* el cual especifica si todas las instancias de ese proyecto han sido lanzadas al módulo CLOUD o no, junto con la fecha del lanzamiento, con los valores ‘P’ de ‘Pending’ y ‘E’ de ‘Ended’. A su vez, guardamos otro valor denominado *status*, que indica si todas las instancias de ese proyecto han sido

anonimizadas o seudonimizadas, con el valor 'E' de 'Ended', o no, con el valor 'P' de 'Pending'. Esto conlleva a que si el *status* contiene el valor 'P' posteriormente no se pueda hacer *push*.

## ANONIM\_STATUS

La tabla ANONIM\_STATUS almacena principalmente la información sobre las instancias que pertenecen a cada proyecto. Se almacena un id por cada instancia de cada proyecto, junto con el id del estudio y el id de la serie a la que pertenece y el id de la instancia.

Al igual que en la tabla PROJECT, se gestiona el estado de las instancias, si han sido anonimizadas o seudonimizadas al completo mediante el valor *status*. En cambio, mediante el valor *status\_push* definimos si el DICOM ha sido lanzado al modulo CLOUD. Por último, se registra por cada instancia de que tipo ha sido el proceso, anonimizado o seudonimizado.

Además, se almacena también información sobre el paciente al que pertenece cada instancia para gestionar dicha información al anonimizar los datos.

## ANON\_PSEUDO

Los datos que se general al anonimizar o seudonimizar un DICOM, se almacenan en la tabla ANON\_PESUDO. Por cada proyecto, se almacena un nuevo id de estudio, un nuevo id de serie y un nuevo id de instancia.

La diferencia entre un DICOM anonimizado o seudonimizado viene dada en el valor ANON\_STATUS\_ID, el cual guarda la relación con la tabla ANONIM\_STATUS y por consiguiente la relación entre la información anonimizado y la original. En caso de que se quiera anonimizar, ese valor será nulo. En cambio, en caso de que se quiera seudonimizar, se guardara el id de ANONIM\_STATUS correspondiente.

Por último, se guarda también los nuevos datos correspondientes al paciente para la posterior gestión en Orthanc.

## 4.4. PROCESO DE ANONIMIZACIÓN/SEUDONIMIZACIÓN

---

### 4.4.1. TRATAMIENTO DE DATOS

El aspecto fundamental para tener en cuenta en primer lugar en un proceso de anonimización son los datos DICOM y la relevancia de cada uno de ellos. El cambio y la eliminación de valores de este proceso esta basado principalmente en los requerimientos que pide el API de Orthanc.

DICOM define la desidentificación como la protección de propiedades confidenciales a través del cifrado o la eliminación, siendo el objetivo encontrar un equilibrio entre la anonimización, la seudonimización, la desidentificación y la reidentificación DICOM.

La anonimización o seudonimización no supone simplemente eliminar información confidencial como las identidades de los pacientes y las fechas de los estudios, sino que supone un cuidado con la información que recopilamos. Algunas propiedades son obligatorias, por lo que, si se eliminan, el DICOM sería inutilizable.

Como ya ha sido mencionado anteriormente en esta memoria, existen varios tipos que distinguen a los elementos de datos, entre los cuales los principales son:

#### **TIPO 1**

Estos son los atributos obligatorios en la instancia SOP y deben tener un valor valido. En el caso de este proceso, se han considerado como obligatorios los siguientes elementos de dato: Graphic Annotation Sequence (0070,0010), Person Identification Code Sequence (0040,1101), Person Name (0040, A123), Verifying Observer Name (0040,A075), Verifying Observer Sequence (0040,A075). El valor de estos atributos es sustituido por el valor ‘anonymized’.

Por otro lado, tenemos las etiquetas con los valores UID, los cuales tienen que ser sustituido por otro UID valido. Las etiquetas que han sido consideradas son las siguientes: Concatenation UID, Context Group Extension Creator UID, Creator Version UID, Device UID, Dimension Organization UID, Dose Reference UID, Failed SOP Instance UID List, Fiducial UID, Frame Of Reference UID, Instance Creator UID, Irradiation Event UID, Large Palette Color Lookup Table UID, Media Storage SOP Instance UID, Palette Color Lookup Table UID, Referenced Frame Of Reference UID, Referenced General Purpose Scheduled Procedure Step Transaction UID, Referenced SOP Instance UID, Referenced SOP Instance UID In File, Related Frame Of Reference UID, Requested SOP Instance UID, Series Instance UID, SOP Instance UID, Storage Media File Set UID, Study Instance UID, Synchronization Frame Of Reference UID, Template Extension Creator UID, Template Extension Organization UID, Transaction UID, UID. El valor de estas etiquetas es sustituido por el valor que genera el método *generate\_uid()* de la biblioteca *pydicom.uid* en Python.

En Orthanc los DICOM están organizados en base a la etiqueta ‘SOP Instance UID’ en referencia a las instancias, a la etiqueta ‘Series Instance UID’ en referencia a las series y ‘Study



Instance UID' en referencia a los estudios. Por lo que, en el proceso de anonimizar y seudonimizar se gestiona el valor de cada uno de ellos para que cada instancia este agrupada en su correspondiente UID de serie y UID de estudio.

## TIPO 2

Las etiquetas de tipo 2 son las que no pueden ser eliminadas, pero pueden no contener ningún valor. Las etiquetas principales que se tienen en cuenta como de tipo 2 son: Accession Number (0008,0050), Content Creator Name (0070,0084), Patient ID (0010,0020), Patient Birth Date (0010,0030), Patient Name (0010,0010), Patient Sex (0010,0040), Referring Physician Name (0008,0090), Study Date (0008,0020), Study ID (0020,0010), Study Time (0008,0030), Verifying Observer Identification Code Sequence (0040,A088).

Al igual que los UID, Orthanc utiliza la etiqueta Patient ID para organizar los estudios, las series y las instancias por pacientes. Por lo que, durante el proceso, se hace una gestión de los pacientes y se les asigna el mismo valor a todos los estudios, las series y las instancias de un mismo paciente, al igual que, también gestionan los valores del nombre del paciente y el ID de estudio. Para generar un ID aleatorio, se hace uso del método `uuid4()` de la biblioteca `uuid`. En el caso del nombre del paciente, se sustituye por un valor que comienza con las 'AN-' y continua con los números que pertenecen a la fecha y hora del momento exacto (*Figura 4.19*). Los valores de las demás etiquetas se dejan vacíos ya que no requieren de un valor.

## TIPO 3

El número de etiquetas de tipo 3 es el más amplio. Estas etiquetas son las que no son requeridas en un DICOM anonimizado, por lo que, se eliminan. Entre todas las etiquetas de este tipo destacan: Acquisition Comments, Actual Human Performers Sequence, Additional Patient History, Curve Date, Digital Signature UID, Generator ID, Patient Comments, Patient Age, Pregnancy Status, Series Description, Verifying Organization, Modified Attributes Sequence, Interpretation Diagnosis Description, Patient Weight, Person Address, Reason For Study.

Estas etiquetas se eliminan mediante el método `delattr(dataset, tag)` de la biblioteca `pydicom`.

### 4.4.2. DESARROLLO DEL PROCESO

La implementación del proceso ha sido desarrollada en el lenguaje de programación Python, en concreto en base a la librería `pydicom`.

En el módulo de ingesta datos, los DICOM anonimizados o seudonimizados se organizan en proyectos mencionado anteriormente. Por lo que, una vez hayan sido seleccionados los archivos a anonimizar o seudonimizar, la primera gestión que se realiza es entorno a los proyectos. Se ofrecen dos opciones: crear un nuevo proyecto o agregar la colección de instancias a un proyecto existente. En el primero de los casos, se crea un nuevo proyecto con su respectivo id y los datos mencionados anteriormente en la tabla PROJECT de la base de datos. Además, se comprueba que el nombre del proyecto no este utilizado anteriormente.

En el segundo de los casos, se selecciona cual es el proyecto al que se le quieren añadir más instancias y se accede a la base de datos para saber cual es el id de dicho proyecto. Ese id se almacena para que posteriormente pueda ser adjudicado en las demás tablas.

Una vez, que se ha seleccionado o creado el proyecto, se encuentra la posibilidad de elegir entre anonimizar o seudonimizar las instancias. A su vez, como se ha mostrado previamente, la interfaz proporciona las opciones de anonimizar o seudonimizar un estudio o una serie. Por lo que todo ello conlleva a que en el proceso se hayan creado 2 ficheros diferentes con 2 funciones diferentes, en función de las opciones que el usuario puede elegir: *anonymize\_functions* y *seudonim\_functions*. Cada uno de ellos está compuesto por dos funciones, una para anonimizar un estudio y otra para una serie.

Cabe mencionar, que durante este proceso se tiene en cuenta si se ha realizado una búsqueda respecto a la modalidad. Un estudio puede contener varias series de modalidades diferentes, por lo que en caso de querer anonimizar o seudonimizar series de una modalidad concreta perteneciente a ese estudio, se tiene en cuenta en el método donde se anonimiza o seudonimiza el estudio seleccionando únicamente las instancias de esas series.

A continuación, una vez que haya sido seleccionada cual es la función que se quiere ejecutar, se hace una llamada HTTP a Orthanc mediante las siguientes URLs:

- **/studies/ID/archive:** descarga la información del estudio con ese ID en concreto.
- **/studies/ID/series:** devuelve una lista con los IDs de las series que pertenecen a ese estudio.
- **/series/ID/archive:** descarga la información de la serie con el ID pasado por parámetro.

Los datos que devuelve la llamada se almacenan en una carpeta ZIP denominada *temp.zip*, creada en el mismo directorio. Para descomprimir esa carpeta se hace uso de la función *zipfile* de la biblioteca *ZipFile* y se crea una carpeta *temp*. Esa carpeta contiene todos los archivos DICOM organizados en carpetas por series y por estudios.

La parte principal del proceso consiste en acceder a cada uno de los ficheros DICOM y a sus respectivos datos e ir intercambiando y eliminando los datos como ha sido descrito en el anterior apartado (5.1. Tratamiento de datos). Para leer cada una de las instancias, se utiliza el método *dcmread('archivodicom.dcm')* de la biblioteca *pydicom*. Toda la información se guarda en un variable *dataset* y posteriormente es a través de ella como se accede a cada una de las etiquetas. Antes de cambiar ciertos valores por unos nuevos, se accede a la base de datos y se almacenan los el ID de estudio, de serie y de instancia de cada uno de los DICOM con los demás valores pertinentes de la tabla ANONIM\_STATUS. En caso de que se quiera anonimizar, se guardara el ID\_ANONIM\_STATUS para poder relacionarlo posteriormente en la tabla ANON\_PSEUDO.

Por ejemplo, si queremos cambiar el valor de la etiqueta Patient ID se realiza de la siguiente manera:

```
dataset.PatientID = 'nuevoValor'
```

Una vez que se haya hecho el análisis de todas las etiquetas DICOM y sus respectivos cambios, se guarda el nuevo DICOM. Para ello, utilizamos la función *save\_as* de la siguiente manera:

```
dataset.save_as(directoriodelDICOM)
```

Seguidamente, cuando el archivo DICOM ha sido guardado, se accede nuevamente a la base de datos, para insertar los nuevos valores correspondientes al ID de estudio, de serie y de instancia. Además, en caso de que se trate de un proceso de seudonimización, se insertara también el ID\_ANON\_STATUS correspondiente para crear la relación. Por otro lado, cuando el DICOM ha sido anonimizado o seudonimizado correctamente, se gestiona el valor *status* de la base de datos. El campo debe de tener el valor 'P' de 'Pending', por lo que lo cambiamos a 'E' de 'Ended' para señalar que las instancias han sido anonimizadas.

El siguiente paso, cuando el archivo DICOM este guardado y los datos estén correctamente insertados en la base de datos, trata sobre lanzar la carpeta con los respectivos archivos a Orthanc. Se realiza mediante una llamada http pasando como parámetro la ruta de la carpeta *temp*, la URL del destino a donde se quieren mandar los DICOM.

Cuando los DICOM ya estén almacenados en el Orthanc, se podrá visualizar el proyecto en la pantalla de proyectos de la interfaz con su respectiva información. En ese momento, se habilita el botón *push* para lanzar los DICOM al módulo CLOUD. Una vez que esta función se haya cumplido, se deshabilita el botón.

## 4.5. FUNCIONAMIENTO E IMPLEMENTACIÓN

En seguimiento a la cronología del funcionamiento de la interfaz, la primera tarea se trata de extraer los estudios desde un PACS, que en este caso desde el Orthanc de Vicomtech. Para ello se proporciona la siguiente pantalla:



*Figura 4.4.- Pantalla de 'Extractor de estudios'*

El funcionamiento de la pantalla esta basada en la conexión que tiene con un PACS externo, en este caso con el Orthanc de Vicomtech mencionado anteriormente. Al ser una empresa externa y contener datos que no son de nuestra pertenencia, el funcionamiento de la pantalla se mostrara mas detalladamente en la defensa de este proyecto. No obstante, cabe mencionar que su funcionamiento comienza cuando añades algún parámetro en el filtro y realizas una búsqueda.

El resultado de la búsqueda es mostrado en la tabla y a raíz de eso, existe la posibilidad de seleccionar cuales son los estudios que se quieren descargar a nuestro Orthanc Middleware para su posterior gestión y manipulación.

Cuando los estudios deseados son descargados y almacenados en el Orthanc Middleware, conectado a nuestra aplicacion web, su información se clasifica en 3 pantallas diferentes:

En primer lugar, tenemos la pantalla donde se muestra la información de los pacientes que tienen algún estudio:

Pacientes

Estudios

Series

Extractor de estudios

Proyectos

Pacientes

Filtro de búsqueda

ID de paciente

Nombre de paciente

Edad

0

Limpiar

Buscar

| ID   | LastUpdate      | PatientName | PatientId | PatientBirthDate | Studies                                      |
|--|-----------------|-------------|-----------|------------------|--|
| 0948fcb6-cf12ab43-bad958c1-bf057ad5-0fc6f54c | 20220112T115150 | COMUNIX     | IYET5.0   | 19410901         | 6c65289b-db2fcb71-7eaf73f4-8e12470c-a4d8d7cf |
| 7c531731-811a61db-1fa8fe5-78aef163-6384215f  | 20220217T154814 | D1-0064     | D1-0064   |                  | 05e1f50e-8a7dff7b-fe59f97a-8f2c084c-41abb5a2 |

Items per page: 2

Primera

Anterior

1

Siguiente

Última

Figura 4.5.- Pantalla de 'Pacientes'

La información sobre el paciente esta organizada en una tabla con las columnas *ID*, que es un ID que genera Orthanc al descargar los estudios, *LastUpdate*, que muestra la fecha de la ultima vez que se ha actualizado la información sobre ese paciente, *PatientName*, que guarda el nombre del paciente, *PatientId*, que hace referencia a la etiqueta Patient ID de los DICOM, *PatientBirthDate*, con la fecha de nacimiento del pacientes y por último, *Studies*, que muestra los ID de los estudios de cada paciente generados por Orthanc.

La única función disponible en esta pantalla es la del filtro de búsqueda, en la cual dependiendo de los parámetros que introduzca, busca al paciente. Para la demostración, se realizan dos pruebas: introduciendo el nombre del paciente e introduciendo una parte del ID del paciente junto con el nombre:

Pacientes

Estudios

Series

Extractor de estudios

Proyectos

Pacientes

Filtro de búsqueda

ID de paciente

Nombre de paciente

Edad

0

CO

Limpiar

Buscar

| ID   | LastUpdate      | PatientName | PatientId | PatientBirthDate | Studies                                      |
|--|-----------------|-------------|-----------|------------------|--|
| 0948fcb6-cf12ab43-bad958c1-bf057ad5-0fc6f54c | 20220112T115150 | COMUNIX     | IYET5.0   | 19410901         | 6c65289b-db2fcb71-7eaf73f4-8e12470c-a4d8d7cf |

Items per page: 2

Primera

Anterior

1

Siguiente

Última

Figura 4.6.- Filtro de búsqueda 'Nombre de paciente'

Pacientes

Estudios

Series

Extractor de estudios

Proyectos

Pacientes

Filtro de búsqueda

ID de paciente

Nombre de paciente

Edad

0

7c531731

D1

Limpiar

Buscar

| ID  | LastUpdate      | PatientName | PatientId | PatientBirthDate | Studies                                      |
|---|-----------------|-------------|-----------|------------------|--|
| 7c531731-811a61db-1fa8fe5-78aef163-6384215f | 20220217T154814 | D1-0064     | D1-0064   |                  | 05e1f50e-8a7dff7b-fe59f97a-8f2c084c-41abb5a2 |

Items per page: 2

Primera

Anterior

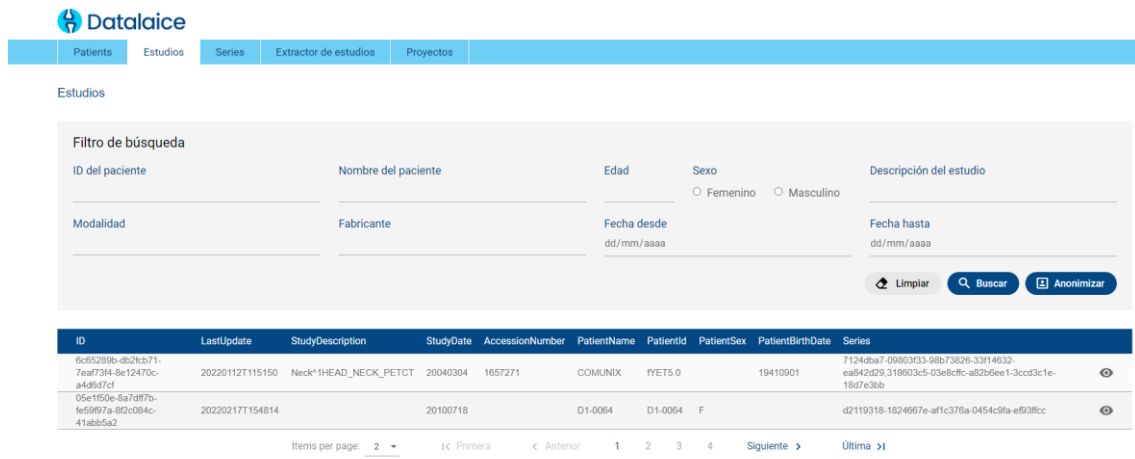
1

Siguiente

Última

Filtro 4.7.- Filtros de búsqueda 'ID' y 'Nombre de paciente'

Por otro lado, la información de los estudios también se almacena y se muestra en la pantalla de 'Estudios', que cuenta con un filtro de búsqueda más amplio:



**Datalaice**

Patients Estudios Series Extractor de estudios Proyectos

Estudios

**Filtro de búsqueda**

ID del paciente: \_\_\_\_\_ Nombre del paciente: \_\_\_\_\_ Edad: \_\_\_\_\_ Sexo:  Femenino  Masculino Descripción del estudio: \_\_\_\_\_

Modalidad: \_\_\_\_\_ Fabricante: \_\_\_\_\_ Fecha desde: dd/mm/aaaa Fecha hasta: dd/mm/aaaa

[Limpiar](#) [Buscar](#) [Anonimizar](#)

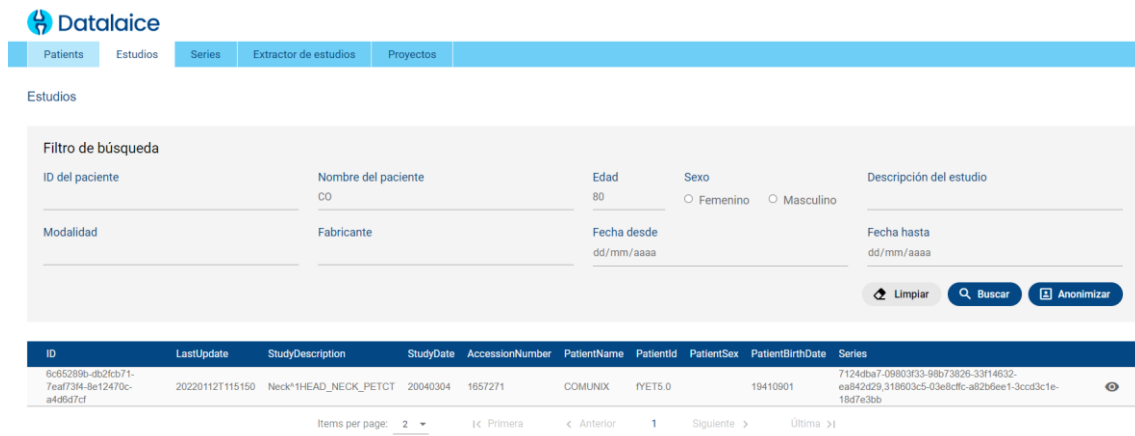
| ID   | LastUpdate       | StudyDescription      | StudyDate | AccessionNumber | PatientName | PatientId | PatientSex | PatientBirthDate | Series  |
|--|------------------|-----------------------|-----------|-----------------|-------------|-----------|------------|------------------|---|
| 6c65289b-d82fcb71-7eaf73f4-8e12470c-a4d9d7cf | 20220112T115150  | Neck*1HEAD_NECK_PETCT | 20040304  | 1657271         | COMUNIX     | 1YET5.0   |            | 19410901         | 7124dba7-09803f33-98b73826-33f14632-<br>ea842d29-318603c5-03e8c8fc-a82b6ee1-3cc03c1e-<br>18d7e3bb |
| 05e1f50e-8a7dff7d-fe59f7a-82c084c-41ab65a2   | 20220217T1154814 |                       | 20100718  |                 | D1-0064     | D1-0064   | F          |                  | d2119318-1824987e-af1c378a-0454c9fa-e9f3ffcc  |

Items per page: 2 | < Primera < Anterior 1 2 3 4 Siguinte > Última >

Filtro 4.8.- Pantalla de 'Estudios'

La información esta dividida en varias columnas: en primer lugar, tenemos el ID generado por Orthanc, el dato *LastUpdate* con la fecha en la que se ha modificado por última vez el estudio, la columna *StudyDescription* con la descripción sobre el estudio, *StudyDate* con la fecha del estudio, *AccessionNumber* con el dato de la etiqueta de los DICOM y la columna *Series* con los ID de las series que pertenecen a este estudio generados por Orthanc. Por otro lado, tenemos la información sobre el paciente al igual que en la pantalla 'Pacientes'.

El filtro de búsqueda para estudios cuenta con más parámetros de búsqueda. Se realizan varias demostraciones con diferentes parámetros:



**Datalaice**

Patients Estudios Series Extractor de estudios Proyectos

Estudios

**Filtro de búsqueda**

ID del paciente: \_\_\_\_\_ Nombre del paciente: CO Edad: 80 Sexo:  Femenino  Masculino Descripción del estudio: \_\_\_\_\_

Modalidad: \_\_\_\_\_ Fabricante: \_\_\_\_\_ Fecha desde: dd/mm/aaaa Fecha hasta: dd/mm/aaaa

[Limpiar](#) [Buscar](#) [Anonimizar](#)

| ID   | LastUpdate      | StudyDescription      | StudyDate | AccessionNumber | PatientName | PatientId | PatientSex | PatientBirthDate | Series  |
|--|-----------------|-----------------------|-----------|-----------------|-------------|-----------|------------|------------------|---|
| 6c65289b-d82fcb71-7eaf73f4-8e12470c-a4d9d7cf | 20220112T115150 | Neck*1HEAD_NECK_PETCT | 20040304  | 1657271         | COMUNIX     | 1YET5.0   |            | 19410901         | 7124dba7-09803f33-98b73826-33f14632-<br>ea842d29-318603c5-03e8c8fc-a82b6ee1-3cc03c1e-<br>18d7e3bb |

Items per page: 2 | < Primera < Anterior 1 Siguinte > Última >


Filtro 4.9.- Filtros de búsqueda 'Nombre del paciente' y 'Edad' en estudios


Estudios

**Filtro de búsqueda**

ID del paciente: \_\_\_\_\_ Nombre del paciente: \_\_\_\_\_ Edad: \_\_\_\_\_ Sexo:  Femenino  Masculino Descripción del estudio: \_\_\_\_\_

Modalidad: CT Fabricante: \_\_\_\_\_ Fecha desde: 01/01/1990 Fecha hasta: dd/mm/aaaa





| ID   | LastUpdate      | StudyDescription      | StudyDate | AccessionNumber | PatientName | PatientId      | PatientSex | PatientBirthDate | Series   |
|--|-----------------|-----------------------|-----------|-----------------|-------------|----------------|------------|------------------|--|
| 038210a9-bf8e5cf0-eb5722bf-8a699934-761a4427 | 20220630T102650 | CT THORAX W/CONTRAST  | 20000101  |                 |             | LIDC-IDRI-0290 | F          |                  | 585f94fe-fd8aadbb-1a1dca8c-dec85ac5-1c315987   |
| 6c652998-b22cb71-7eaf734-8e12470c-a449d7cf   | 20220112T115150 | Neck*1HEAD_NECK_PETCT | 20040304  | 1657271         | COMLJNX     | IYETS.0        |            | 19410901         | 7124baa7-0980333-98b73826-33f14632-ea942d29-318603c-5-03e8cfc-a82b6ee1-3ccdf3ce-18d7e3bb |

Items per page: 2 | < Primera < Anterior 1 2 Siguiete > Última >


Figura 4.10.- Filtros de búsqueda 'Modalidad' y 'Fecha desde' en estudios


Estudios

**Filtro de búsqueda**

ID del paciente: \_\_\_\_\_ Nombre del paciente: \_\_\_\_\_ Edad: \_\_\_\_\_ Sexo:  Femenino  Masculino Descripción del estudio: \_\_\_\_\_

Modalidad: \_\_\_\_\_ Fabricante: GE MEDICAL Fecha desde: dd/mm/aaaa Fecha hasta: 20/06/2022





| ID   | LastUpdate      | StudyDescription     | StudyDate | AccessionNumber | PatientName | PatientId      | PatientSex | PatientBirthDate | Series                                       |
|--|-----------------|----------------------|-----------|-----------------|-------------|----------------|------------|------------------|--|
| bad95b7b-d2245eb3-4a017410-b4415ae2-e59253a6 | 20220630T102648 |                      | 20000101  |                 |             | LIDC-IDRI-0085 | M          |                  | e2f89cc4-050f128c-b5248108-9e720310-8afa3fb4 |
| f0b59f7a-29943cb2-d947f9c2-4a78e778-99033655 | 20220630T102649 | CHEST, PA & LATERAL  | 20000101  |                 |             | LIDC-IDRI-0135 | M          |                  | 3b64c0cf-5de895f1-4fa4f38c-27d70211-1ec5ce0a |
| 038210a9-bf8e5cf0-eb5722bf-8a699934-761a4427 | 20220630T102650 | CT THORAX W/CONTRAST | 20000101  |                 |             | LIDC-IDRI-0290 | F          |                  | 585f94fe-fd8aadbb-1a1dca8c-dec85ac5-1c315987 |

Items per page: 4 | < Primera < Anterior 1 Siguiete > Última >



Figura 4.11.- Filtros de búsqueda 'Fabricante' y 'Fecha hasta' en estudios

Estudios

**Filtro de búsqueda**

ID del paciente: LIDC Nombre del paciente: \_\_\_\_\_ Edad: \_\_\_\_\_ Sexo:  Femenino  Masculino Descripción del estudio: CHEST

Modalidad: \_\_\_\_\_ Fabricante: \_\_\_\_\_ Fecha desde: dd/mm/aaaa Fecha hasta: dd/mm/aaaa


| ID   | LastUpdate      | StudyDescription    | StudyDate | AccessionNumber | PatientName | PatientId      | PatientSex | PatientBirthDate | Series                                       |
|--|-----------------|---------------------|-----------|-----------------|-------------|----------------|------------|------------------|--|
| f0b59f7a-29943cb2-d947f9c2-4a78e778-99033655 | 20220630T102649 | CHEST, PA & LATERAL | 20000101  |                 |             | LIDC-IDRI-0135 | M          |                  | 3b64c0cf-5de895f1-4fa4f38c-27d70211-1ec5ce0a |

Items per page: 4 | < Primera < Anterior 1 Siguiete > Última >

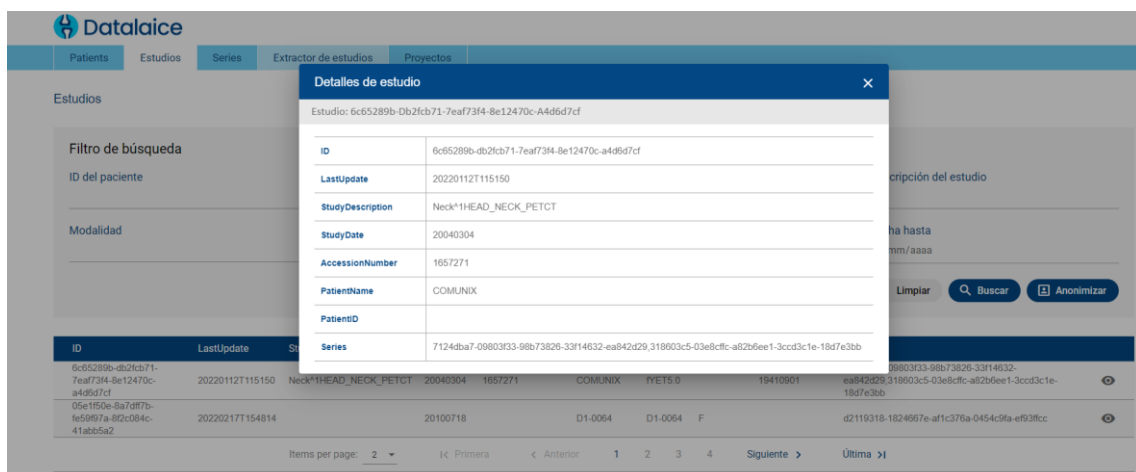
Figura 4.12.- Filtros de búsqueda 'ID del paciente', 'Sexo' y 'Descripción del estudio' en estudios

Por otro lado, la pantalla de 'Estudios' ofrece otra funcionalidad que trata sobre poder consultar mas detalles de los estudios. Cada fila de la tabla contiene un símbolo de un ojo, desde el cual se abre una ventana emergente con los datos del estudio:



| ID   | LastUpdate      | StudyDescription      | StudyDate | AccessionNumber | PatientName | PatientId | PatientSex | PatientBirthDate | Series   |
|--|-----------------|-----------------------|-----------|-----------------|-------------|-----------|------------|------------------|--|
| 6c65289b-db2fcb71-7eaf73f4-8e12470c-a4d6d7cf | 20220112T115150 | Neck*1HEAD_NECK_PETCT | 20040304  | 1657271         | COMUNIX     | fYET5.0   |            | 19410901         | 7124dba7-09803f33-98b73826-33f14632-ea842d29-318603c5-03e8cfc-a82b6ee1-3ccd3c1e-18d7e3bb |
| 05e1f50e-8a7d7fb-fe59f97a-8f2c084c-41abb5a2  | 20220217T154814 |                       | 20100718  |                 | D1-0064     | D1-0064   | F          |                  | d2119318-1824667e-af1c376a-0454c9fa-ef93fcc  |

Figura 4.13.- Referencia al símbolo de detalles en la pantalla 'Estudios'



**Detalles de estudio**

Estudio: 6c65289b-db2fcb71-7eaf73f4-8e12470c-a4d6d7cf

|                  |  |
|------------------|--|
| ID               | 6c65289b-db2fcb71-7eaf73f4-8e12470c-a4d6d7cf   |
| LastUpdate       | 20220112T115150  |
| StudyDescription | Neck*1HEAD_NECK_PETCT  |
| StudyDate        | 20040304   |
| AccessionNumber  | 1657271  |
| PatientName      | COMUNIX  |
| PatientID        | fYET5.0  |
| Series           | 7124dba7-09803f33-98b73826-33f14632-ea842d29-318603c5-03e8cfc-a82b6ee1-3ccd3c1e-18d7e3bb |

Figura 4.14.- Ventana emergente 'Detalles de estudio'

Además, entre las funciones que contiene la pantalla 'Estudios' destaca la posibilidad de anonimizar o seudonimizar los estudios. Este proceso comienza mediante el botón denominado 'Anonimizar':



Fecha desde: dd/mm/aaaa

Fecha hasta: dd/mm/aaaa

| StudyDate | AccessionNumber | PatientName | PatientId | PatientSex | PatientBirthDate | Series   |
|-----------|-----------------|-------------|-----------|------------|------------------|--|
| 0304      | 1657271         | COMUNIX     | fYET5.0   |            | 19410901         | 7124dba7-09803f33-98b73826-33f14632-ea842d29-318603c5-03e8cfc-a82b6ee1-3ccd3c1e-18d7e3bb |
| 0718      |                 | D1-0064     | D1-0064   | F          |                  | d2119318-1824667e-af1c376a-0454c9fa-ef93fcc  |

Figura 4.15.- Botón para anonimizar/seudonimizar estudios



Mediante este botón, se activa los ‘checkbox’ que permiten seleccionar cuales son los estudios que se quieren anonimizar. Por otro lado, también se presentan 3 nuevos botones, uno para poder volver atrás, otro para seleccionar o deseleccionar los estudios y otro para continuar con el proceso. En nuestro caso, le daremos al botón ‘Continuar’ para poder seguir con el proceso:

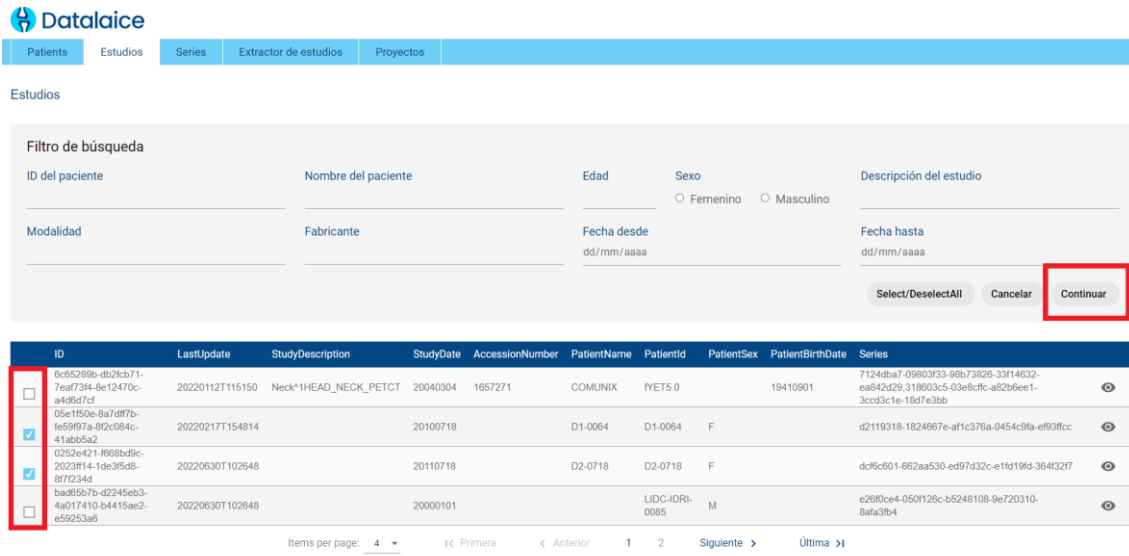


Figura 4.16.- Pantalla para seleccionar los estudios

A continuación, se abre una pantalla emergente que presenta dos opciones: crear un nuevo proyecto o asignar esos estudios que se quieren anonimizar a un proyecto ya existente:

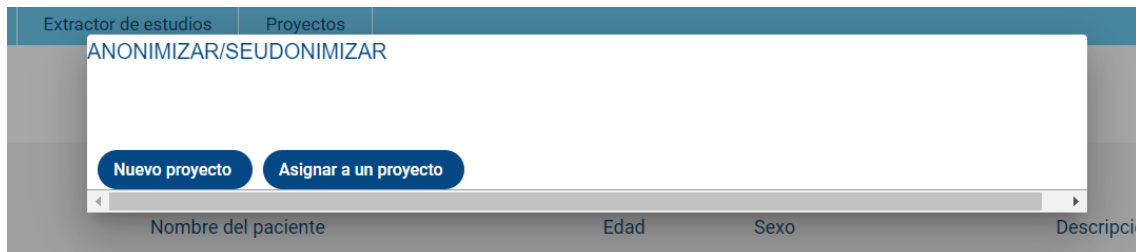


Figura 4.17.- Ventana emergente para anonimizar/seudonimizar

En el caso de que se quiera crear un nuevo proyecto, se abre la siguiente ventana en la que es necesario introducir un nombre y una descripción para el proyecto:

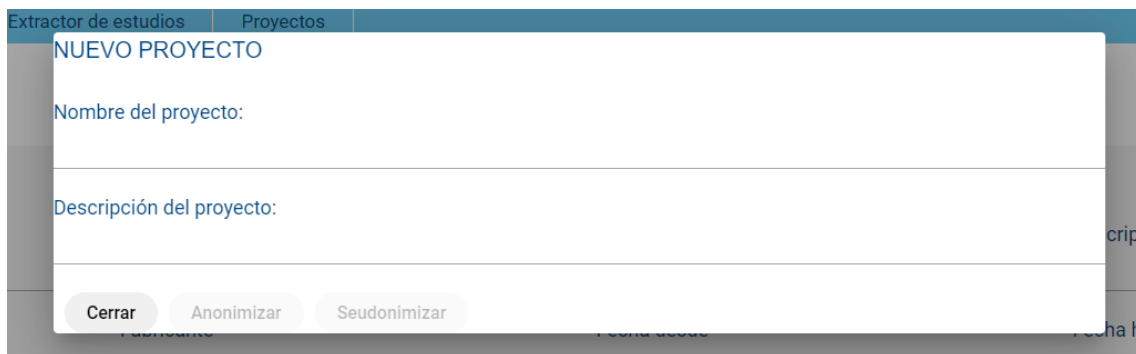


Figura 4.18.- Ventana emergente para crear un proyecto

En caso contrario, si se elige la opción de asignar los estudios a un proyecto ya existente se abre la siguiente ventana, con una lista desplegable que contiene los nombres de los proyectos:

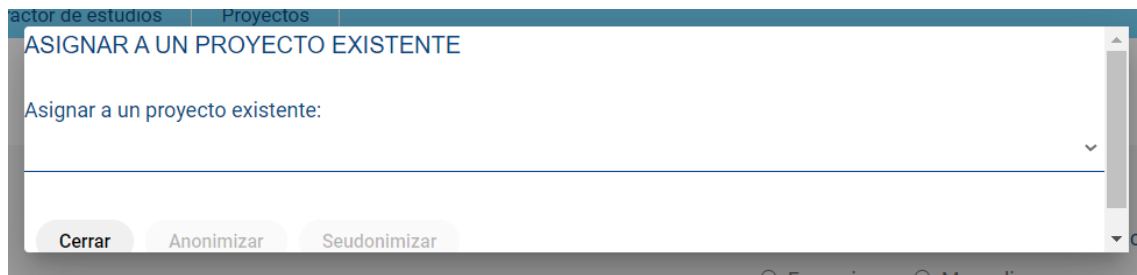


Figura 4.19.- Ventana emergente con los proyectos existentes

Se puede contemplar que, en ambos casos, se permite la opción de elegir entre anonimizar o seudonimizar los estudios. Ambos botones permanecen deshabilitados hasta que se asigne un nombre al proyecto nuevo o se elija el nombre del proyecto ya existente.

Una vez que se haya elegido cuál de los dos tipos de procesos se quiere llevar a cabo, el fichero se anonimiza o seudonimiza, tal y como ha sido explicado en el apartado anterior, y se muestra en la tabla de estudios. Para mostrarlo, haremos un ejemplo creando el proyecto 'Proyecto1' y anonimizando un estudio seleccionado:

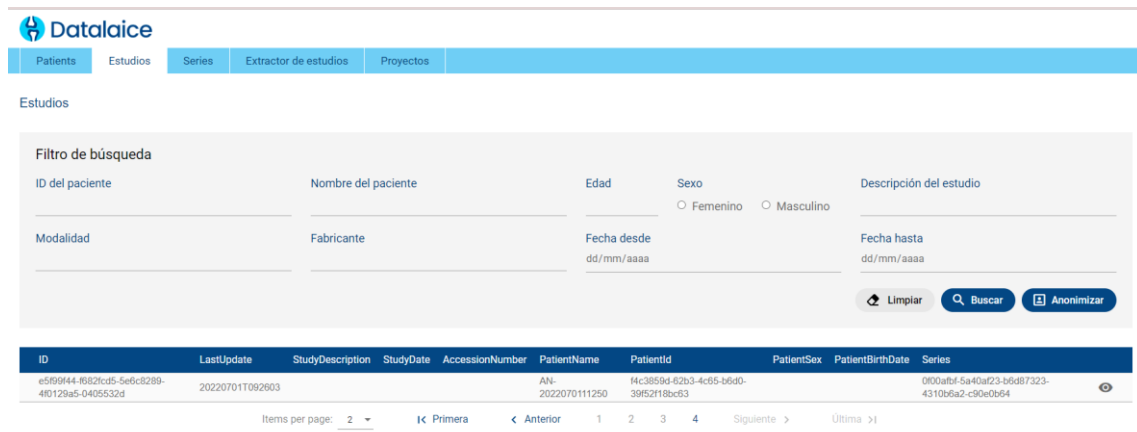


Figura 4.20.- Pantalla de 'Estudios' con el nuevo estudio anonimizado

Se comprueba a su vez, que se ha creado el proyecto en la pantalla 'Proyectos' con sus respectivos datos, aunque posteriormente se haga un análisis mas detallado de su funcionamiento:

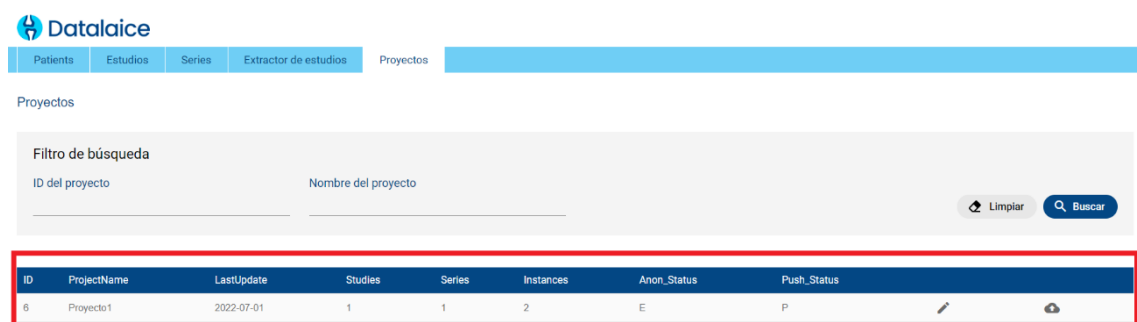
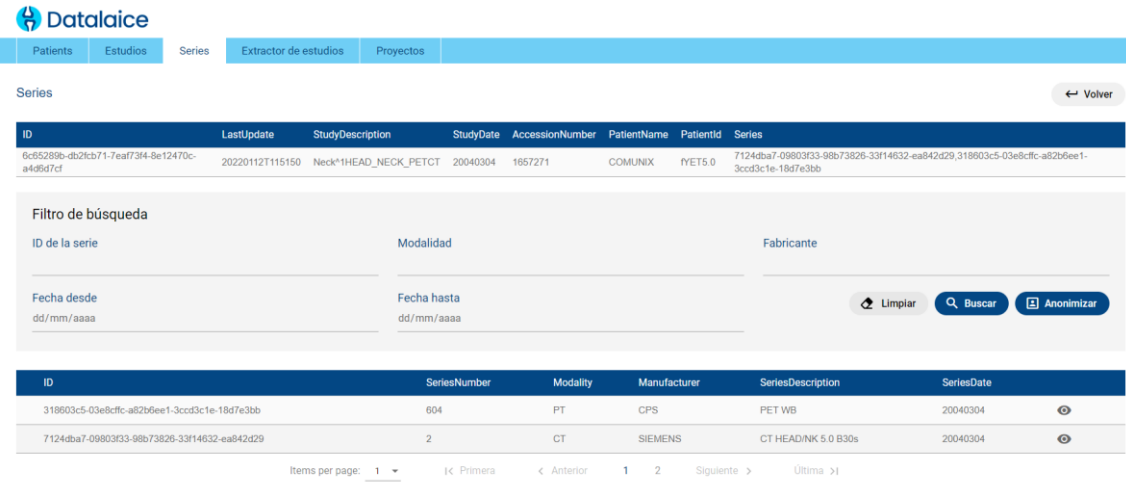


Figura 4.21.- Pantalla de 'Proyectos' con un nuevo proyecto

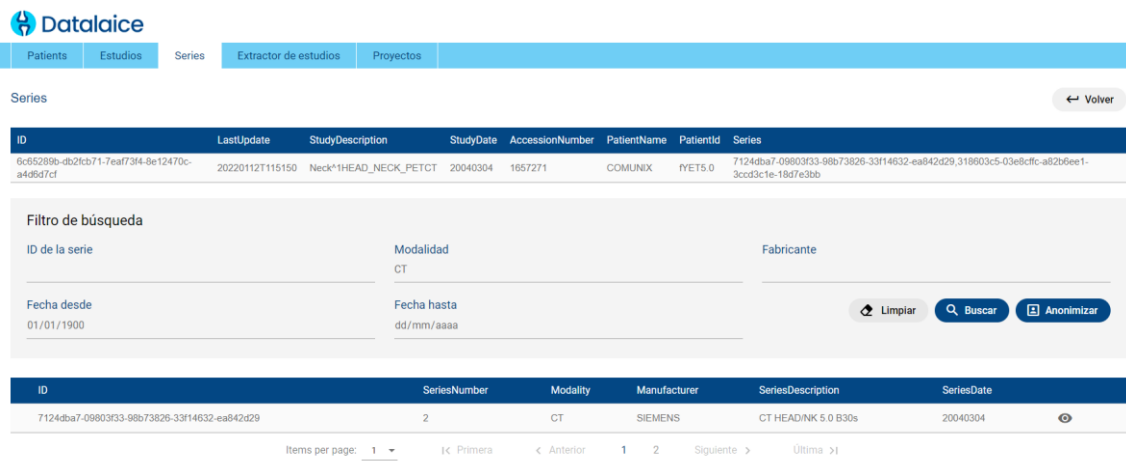
Por último, otra de las funciones que ofrece esta pantalla trata sobre poder consultar las series que pertenecen a cada estudio, simplemente haciendo clic en la fila del estudio. El resultado quedaría de la siguiente manera:



The screenshot shows the 'Series' page in the Datalaice application. The page has a navigation bar with 'Patients', 'Estudios', and 'Series' tabs. The 'Series' tab is active, and there are sub-tabs for 'Extractor de estudios' and 'Proyectos'. Below the navigation bar, there is a table of series data. The table has columns for ID, LastUpdate, StudyDescription, StudyDate, AccessionNumber, PatientName, PatientId, and Series. The first row shows a series ID of '6c55289b-db2fcb71-7eaf73f4-8e12470c-a4d9d7cf' and a study description of 'Neck\*1HEAD\_NECK\_PETCT'. Below the table, there is a search filter section with fields for 'ID de la serie', 'Modalidad', 'Fecha desde', and 'Fecha hasta'. The 'Modalidad' field is set to 'CT'. There are buttons for 'Limpiar', 'Buscar', and 'Anonimizar'. At the bottom, there is a table with columns for ID, SeriesNumber, Modality, Manufacturer, SeriesDescription, and SeriesDate. The first row shows a series ID of '318903c5-03e8dfc-a82b6ee1-3ccd3c1e-18d7e3bb' and a modality of 'PT'. The second row shows a series ID of '7124dba7-09803f33-98b73826-33f14632-ea842d29' and a modality of 'CT'.

Figura 4.22.- Pantalla de 'Series' correspondiente a un estudio concreto

Mediante esta figura, se contempla que la pantalla resultante al consultar las series de un estudio concreto también ofrece la posibilidad de búsqueda mediante diferentes filtros y la de consultar los detalles de cada serie mediante el símbolo del ojo:



The screenshot shows the 'Series' page in the Datalaice application, similar to the previous one, but with search filters applied. The 'Modalidad' field is set to 'CT' and the 'Fecha desde' field is set to '01/01/1900'. The table below the search filter section shows a single row with a series ID of '7124dba7-09803f33-98b73826-33f14632-ea842d29' and a modality of 'CT'. The 'Buscar' button is highlighted, indicating that the search has been executed.

Figura 4.23.- Filtros de búsqueda 'Modalidad' y 'Fecha desde' en series

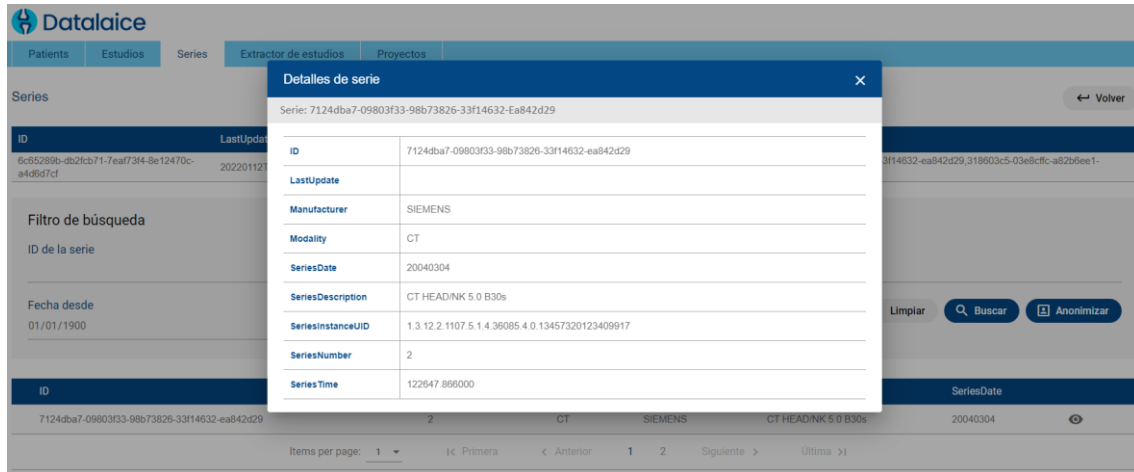


Figura 4.24.- Ventana emergente 'Detalles de serie' de un estudio

Por otra parte, se puede observar cómo esta pantalla también dispone de la posibilidad de anonimizar o seudonimizar una serie:

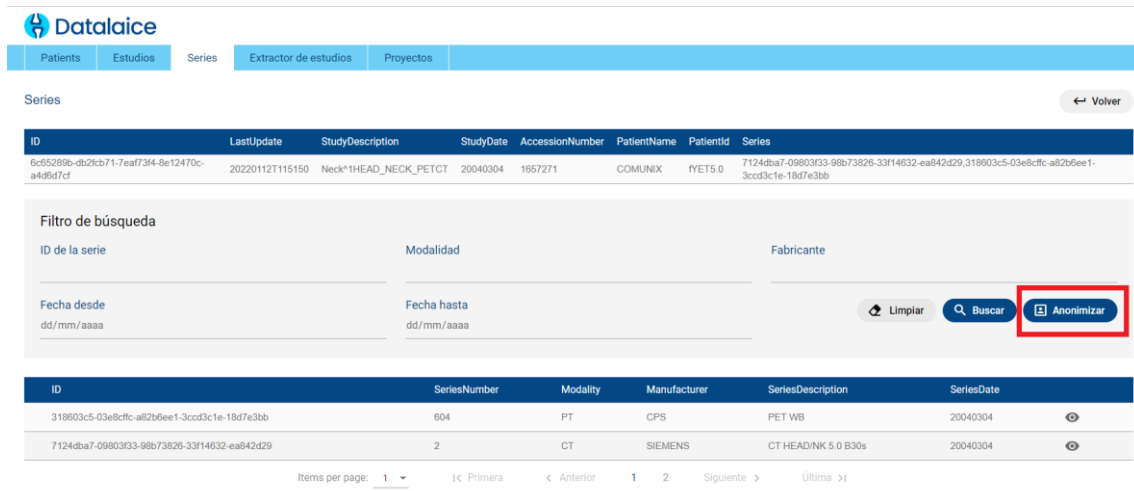


Figura 4.25.- Botón para anonimizar/seudonimizar series de un estudio

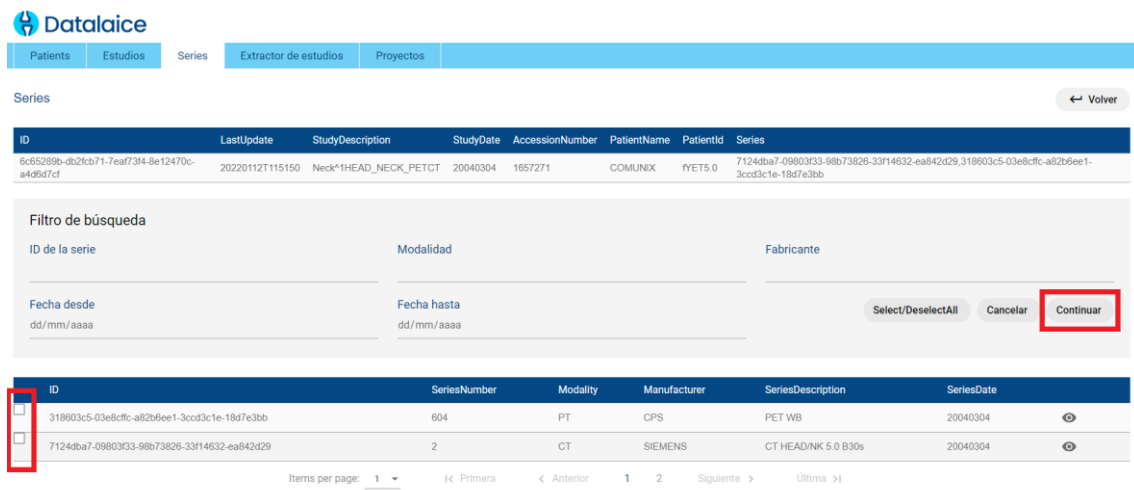
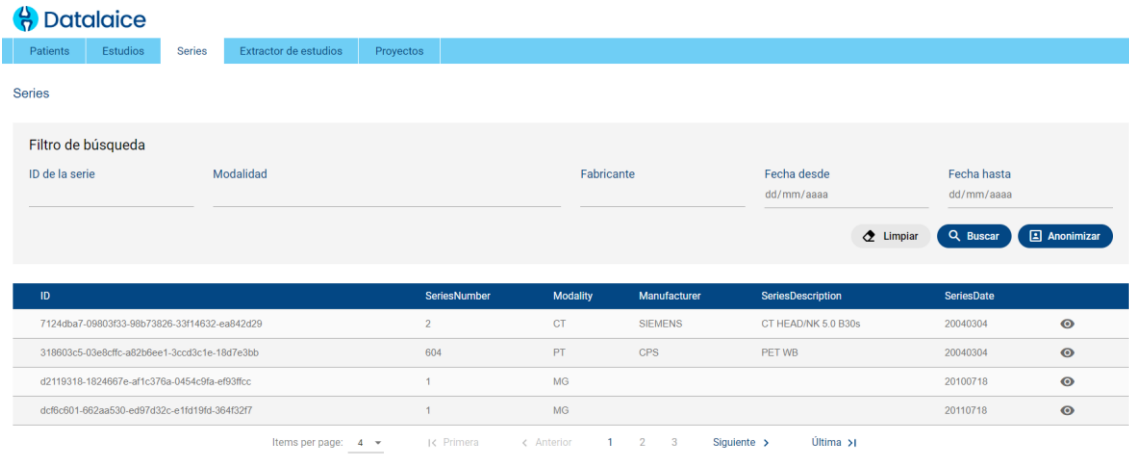


Figura 4.26.- Pantalla para seleccionar las series de un estudio

El resto de las pantallas emergentes son las mismas que se han especificado en las figuras 4.17, 4.18, y 4.19.

Por lo que se refiere a las series, la interfaz web esta compuesta por una pantalla que muestra todas las series con sus respectivos datos:



**Filtro de búsqueda**

ID de la serie:

Modalidad:

Fabricante:

Fecha desde: dd/mm/aaaa

Fecha hasta: dd/mm/aaaa

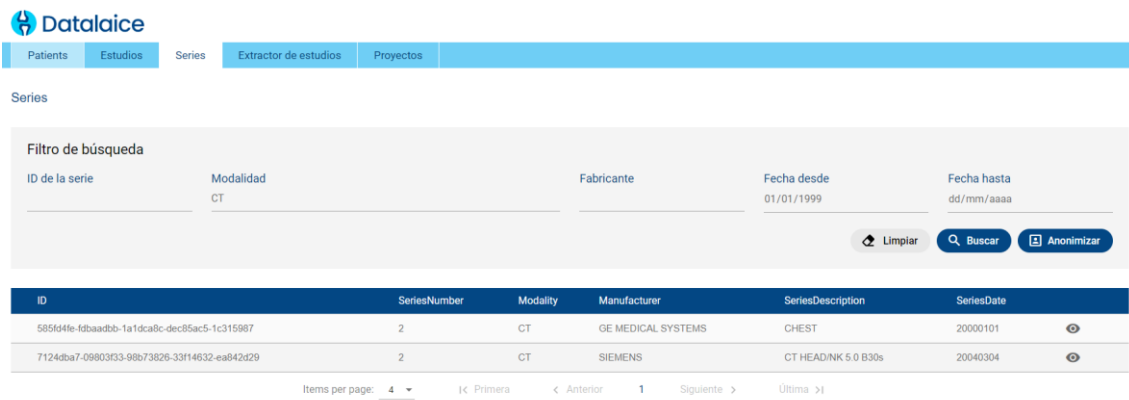
[Limpiar](#) [Buscar](#) [Anonimizar](#)

| ID  | SeriesNumber | Modality | Manufacturer | SeriesDescription    | SeriesDate |
|---|--------------|----------|--------------|----------------------|------------|
| 7124ba7-09803f33-98b73826-33f14632-aa842d29 | 2            | CT       | SIEMENS      | CT HEAD/INK 5.0 B30s | 20040304   |
| 318603c5-03e8cfc-a82b6ee1-3ccd3c1e-18d7e3bb | 604          | PT       | CPS          | PET WB               | 20040304   |
| d2119318-1824667e-af1c378a-0454c9fa-e93ffcc | 1            | MG       |              |                      | 20100718   |
| dcf6c601-662aa530-ed97d32c-e1fd19d-364f3277 | 1            | MG       |              |                      | 20110718   |

Items per page: 4 [Primera](#) [Anterior](#) 1 2 3 [Siguinte](#) [Última](#)

Figura 4.27.- Pantalla de 'Series'

Al igual que en la pantalla de 'Estudios', esta pantalla ofrece varias funcionalidades. En primer lugar, contamos con un filtro de búsqueda que cuenta con varios parámetros:



**Filtro de búsqueda**

ID de la serie:

Modalidad: CT

Fabricante:

Fecha desde: 01/01/1999

Fecha hasta: dd/mm/aaaa

[Limpiar](#) [Buscar](#) [Anonimizar](#)

| ID   | SeriesNumber | Modality | Manufacturer       | SeriesDescription    | SeriesDate |
|--|--------------|----------|--------------------|----------------------|------------|
| 585fd4fe-fd8aadbb-1a1dca8c-dec85ac5-1c315987 | 2            | CT       | GE MEDICAL SYSTEMS | CHEST                | 20000101   |
| 7124ba7-09803f33-98b73826-33f14632-aa842d29  | 2            | CT       | SIEMENS            | CT HEAD/INK 5.0 B30s | 20040304   |

Items per page: 4 [Primera](#) [Anterior](#) 1 [Siguinte](#) [Última](#)

Figura 4.28.- Filtros de búsqueda 'Modalidad' y 'Fecha desde' en series

Por otro lado, también podemos ver los detalles de cada serie, mediante el símbolo del ojo:

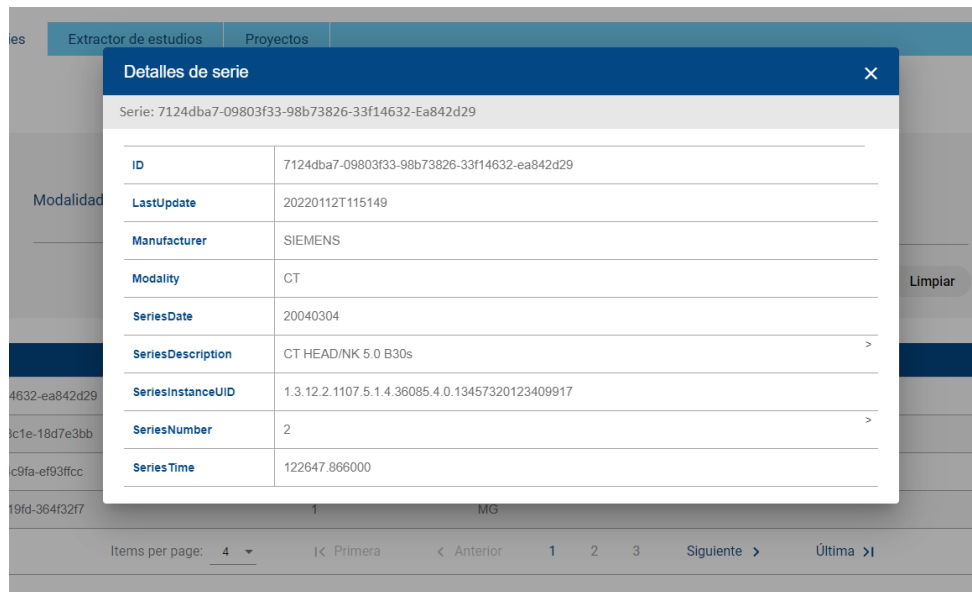


Figura 4.29.- Ventana emergente 'Detalles de serie'

Asimismo, también tenemos la posibilidad de anonimizar o seudonimizar una serie concreta desde esta misma pantalla. En este caso, vamos a seudonimizar una serie en un proyecto nuevo. La manera de hacerlo es igual que en las anteriores pantallas:

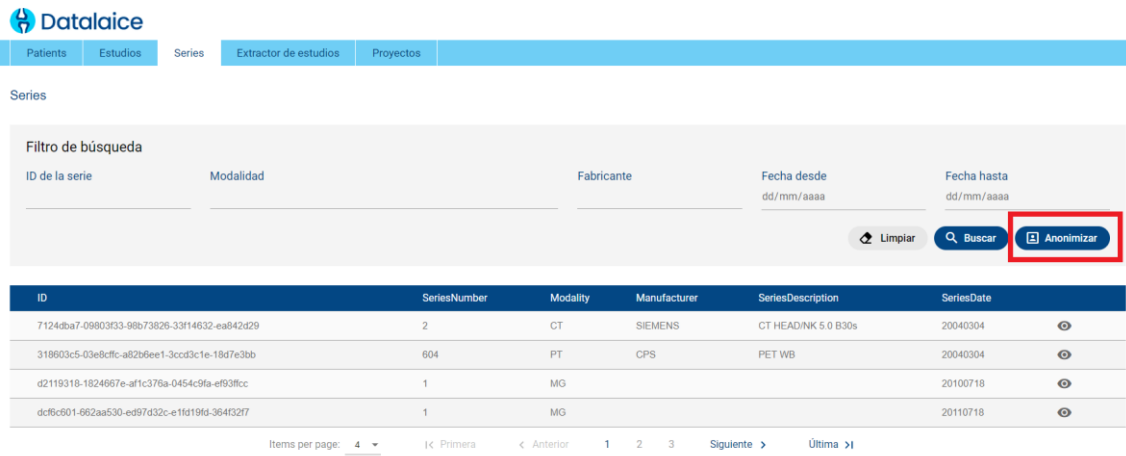


Figura 4.30.- Botón para anonimizar/seudonimizar series

Series

Filtro de búsqueda

ID de la serie      Modalidad      Fabricante      Fecha desde dd/mm/aaaa      Fecha hasta dd/mm/aaaa

Select/DeselectAll    Cancelar    **Continuar**

| ID  | SeriesNumber | Modality | Manufacturer | SeriesDescription    | SeriesDate |
|---|--------------|----------|--------------|----------------------|------------|
| <input type="checkbox"/> 7124ba7-09803f33-98b73826-33f14632-aa842d29  | 2            | CT       | SIEMENS      | CT HEAD/INK 5.0 B30s | 20040304   |
| <input type="checkbox"/> 318603c5-03e8cfc-a82b6ee1-3ccd3c1e-18d7e3bb  | 604          | PT       | CPS          | PET WB               | 20040304   |
| <input type="checkbox"/> d2119318-1824667e-af1c378a-0454c9fa-e93ffcc  | 1            | MG       |              |                      | 20100718   |
| <input type="checkbox"/> dc9fc601-682aa530-ed97d32c-e1fd19fd-364f32f7 | 1            | MG       |              |                      | 20110718   |

Items per page: 4    < Primera    < Anterior    1 2 3    Siguiete >    Última >

Figura 4.31.- Pantalla para seleccionar las series

Extractor de estudios    **Proyectos**

**NUEVO PROYECTO**

Nombre del proyecto:  
Proyecto2

Descripción del proyecto:  
Proyecto seudonimizadq

Cerrar    Anonimizar    Seudonimizar

Figura 4.32.- Ventana emergente para crear un nuevo proyecto

Se comprueba que la serie ha sido seudonimizada desde la pantalla de estudios, ya que se muestra cual es el estudio nuevo anonimizado mediante el valor del nombre del paciente, valor que se muestra desde la pantalla de estudios. Es decir, se ha creado un nuevo estudio con la serie que se ha mandado a seudonimizar:

**Datalaice**

Patients    Estudios    **Series**    Extractor de estudios    Proyectos

Estudios

Filtro de búsqueda

ID del paciente      Nombre del paciente      Edad      Sexo  Femenino  Masculino      Descripción del estudio

Modalidad      Fabricante      Fecha desde dd/mm/aaaa      Fecha hasta dd/mm/aaaa

Limpiar    Buscar    Anonimizar

| ID  | LastUpdate      | StudyDescription | StudyDate | AccessionNumber | PatientName      | PatientId                            | PatientSex | PatientBirthDate | Series                                      |
|---|-----------------|------------------|-----------|-----------------|------------------|--------------------------------------|------------|------------------|---|
| e5f9944-4827cd5-5e6c8289-483420e5-40e8c3a4  | 20220701T092603 |                  |           |                 | AN-2022070111590 | 4c3859d-62b3-4c95-b6d0-20f5182b1e83  |            |                  | 0f00afbf-5a40a23-b9d87323-4340e6a3-90e016c7 |
| 4bcf3a48-7f34c28-74a17e2a-e87e938b-5a71d2c7 | 20220701T095920 |                  |           |                 | AN-2022070111590 | ecba8a1a-3f84-4b83-a8e1-cb678d9d9759 |            |                  | 7319b49b-4bed9f30-b248fb3-1334b481-34216f78 |

Items per page: 2    < Primera    < Anterior    1 2 3 4 5    Siguiete >    Última >

Figura 4.33.- Pantalla de 'Estudios' con el nuevo estudio que contiene la serie seudonimizada

Por último, se comprueba también que el proyecto ha sido correctamente creado y almacenado con sus respectivos cambios.



Figura 4.34.- Pantalla de 'Proyectos' con un proyecto nuevo

En esta aplicación web también se hace un control para que en un mismo proyecto no se pueda anonimizar y seudonimizar simultáneamente. Por lo que, por ejemplo, si intentamos anonimizar en el último proyecto 'Proyecto2' que han sido seudonimizadas las instancias, nos sale la siguiente pantalla:

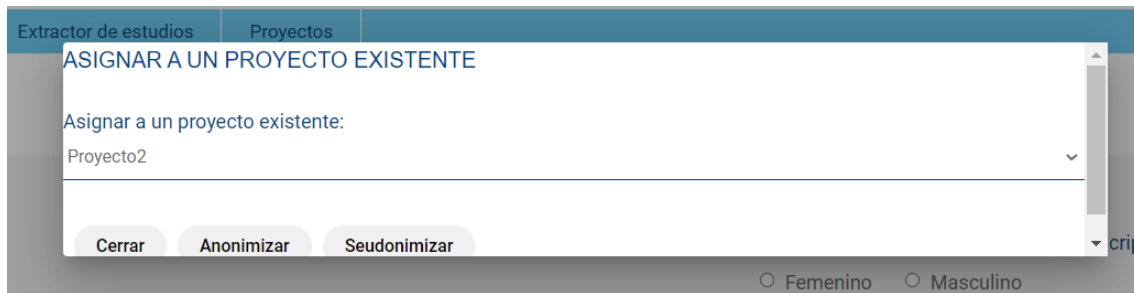


Figura 4.35.- Ventana emergente para asignar un estudio a un proyecto existente

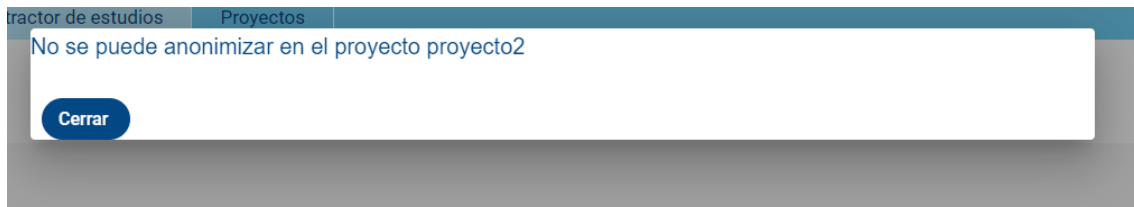


Figura 4.36.- Ventana emergente de error

Finalmente, la interfaz web cuenta con la pantalla 'Proyectos', que ya ha sido expuesta anteriormente. Esta pantalla cuenta con un filtro de búsqueda basada en dos parámetros:

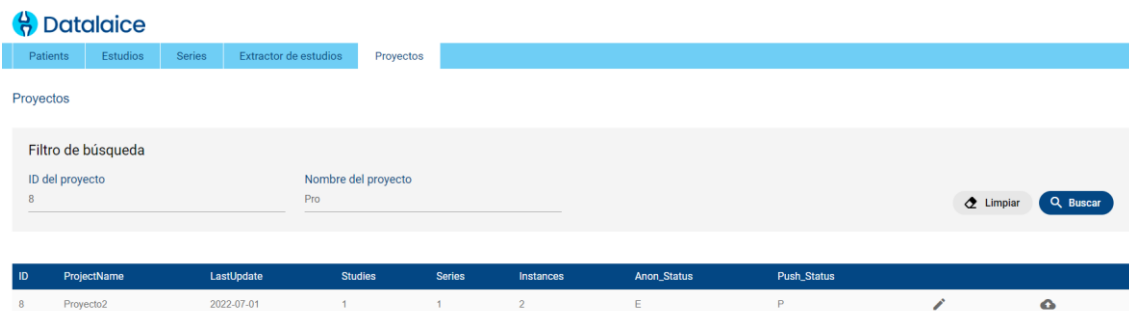


Figura 4.37.- Filtros de búsqueda 'ID del proyecto' y 'Nombre del proyecto' en proyectos



Por otra parte, también ofrece la posibilidad de cambiar la descripción del proyecto de la siguiente manera:

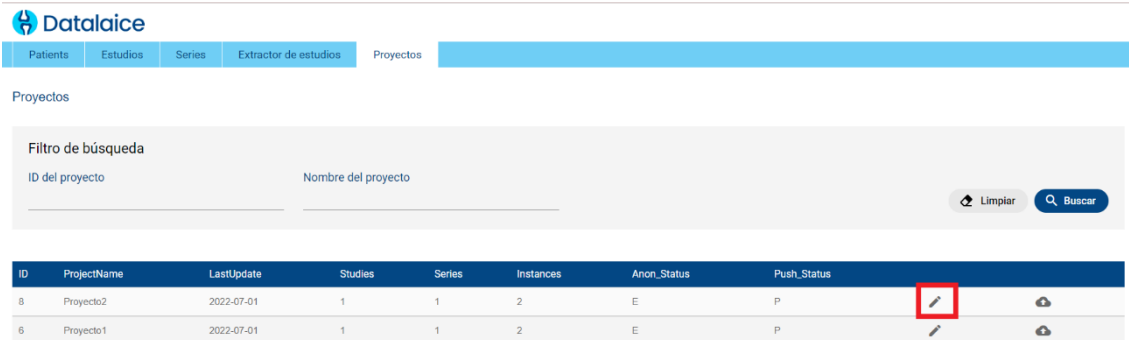


Figura 4.38.- Referencia al símbolo para editar la descripción en la pantalla 'Proyectos'

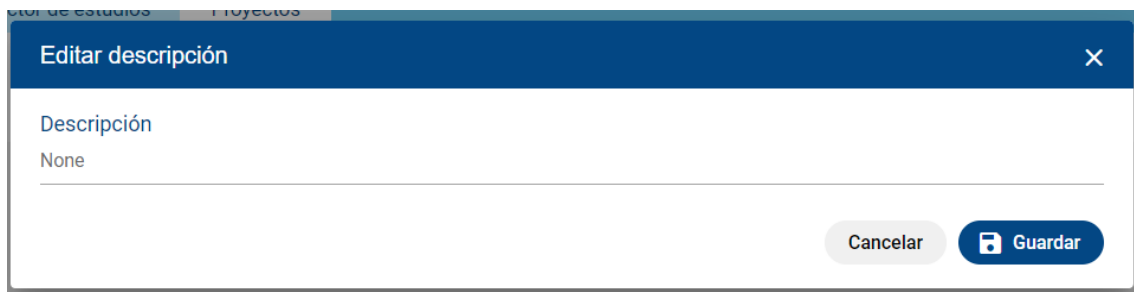


Figura 4.39.- Ventana emergente 'Editar descripción'

Esta función cambiara simultáneamente el valor de la columna *LastUpdate*, en caso de que haya sido realizada en una fecha diferente.

Por último, la funcionalidad más importante de esta pantalla es la gestión del *push*. Como ya ha sido mencionado anteriormente, las instancias anonimizadas o seudonimizadas de cada proyecto se lanzan a un nuevo Orthanc CLOUD para su posterior gestión y manipulación. Esta acción se realiza mediante el siguiente botón:

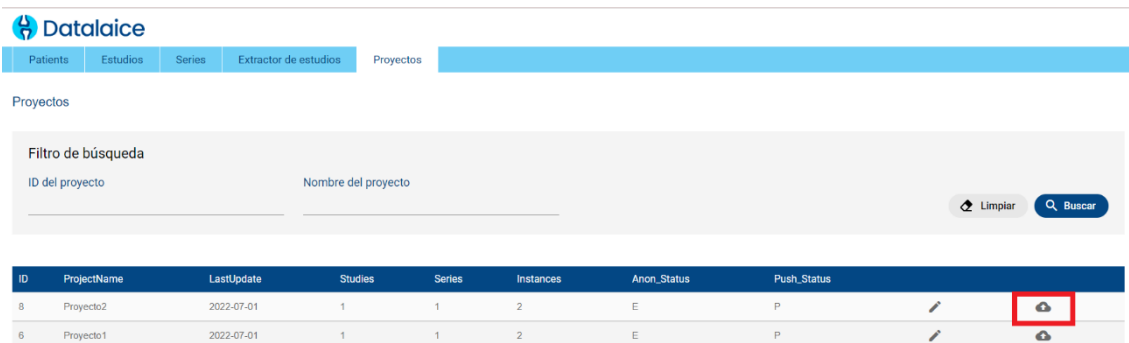
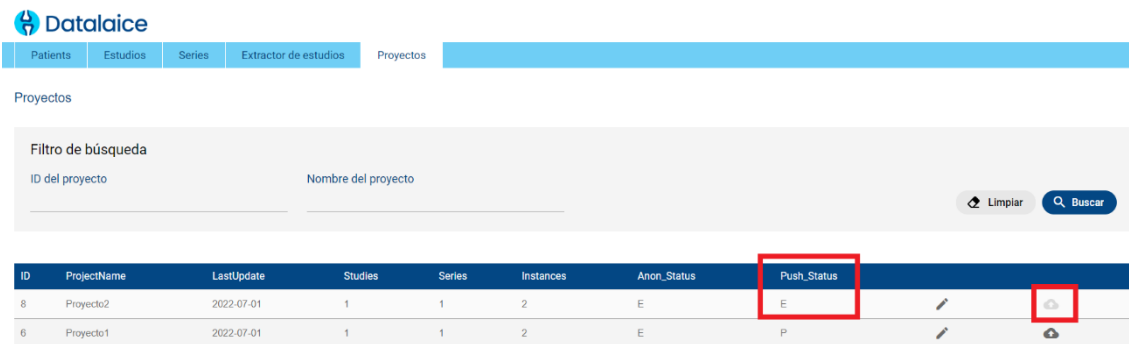


Figura 4.40.- Referencia al botón *push*

Como se puede observar, el botón para realizar el *push* está habilitado y se debe a que el valor de la columna *push\_status* es igual a 'P' que significa que todavía hay instancias anonimizadas o seudonimizadas en ese proyecto que no han sido lanzadas. Una vez que se haya presionado el botón y las instancias de ese proyecto sean lanzadas al nuevo Orthanc, el botón se deshabilitara.

Recordemos, que el estudio que esta almacenado en el proyecto denominado 'Proyecto2' tiene el nombre de paciente 'AN-2022070111590', por lo que, ese será el nombre del paciente que deberá mostrarse en el Orthanc CLOUD una vez que se haya hecho el *push*. Para demostrarlo, se necesita acceder a un servidor propio de la empresa, por lo que será mostrado detalladamente el día de la defensa.

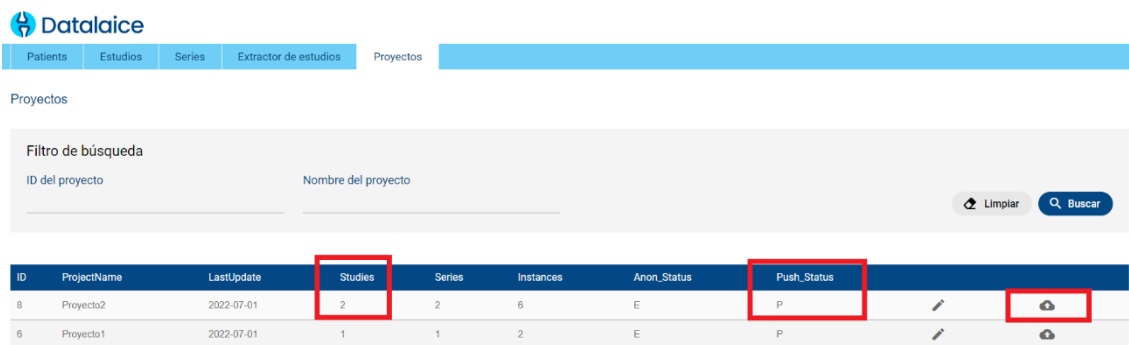
Lo que se puede ver a través la interfaz es como el valor de la columna *status\_push* a cambiado a 'E', que significa que las instancias de ese proyecto han sido lanzadas, y como el botón *push* esta deshabilitado.



| ID | ProjectName | LastUpdate | Studies | Series | Instances | Anon_Status | Push_Status |  |  |
|----|-------------|------------|---------|--------|-----------|-------------|-------------|--|--|
| 8  | Proyecto2   | 2022-07-01 | 1       | 1      | 2         | E           | E           |  |  |
| 6  | Proyecto1   | 2022-07-01 | 1       | 1      | 2         | E           | P           |  |  |

Figura 4.41.- Pantalla de proyectos modificada

Para una comprobación más exhaustiva, hacemos la prueba de añadir nuevas instancias a ese mismo proyecto y así comprobar que el botón vuelve a estar habilitado y el valor de la columna *push\_status* vuelve a ser 'P':



| ID | ProjectName | LastUpdate | Studies | Series | Instances | Anon_Status | Push_Status |  |  |
|----|-------------|------------|---------|--------|-----------|-------------|-------------|--|--|
| 8  | Proyecto2   | 2022-07-01 | 2       | 2      | 6         | E           | P           |  |  |
| 6  | Proyecto1   | 2022-07-01 | 1       | 1      | 2         | E           | P           |  |  |

Figura 4.42.- Pantalla de proyectos con modificaciones en un proyecto

Podemos comprobar, que ese proyecto consta de dos estudios, que el valor de *push\_status* es igual a 'P' y que el botón *push* está habilitado.

# Capítulo 6

## CONCLUSIONES

### Contenido

---

#### 6.1. Conclusiones

---

## 6.1. CONCLUSIONES

---

Los estudios de investigación analizados revelaron el potencial de uso de los metadatos DICOM, pero también los retos asociados a la indexación y gestión de grandes volúmenes de metadatos, a saber, la identificación de los interesados en la imagen médica, las necesidades informáticas y la gestión de la información, la calidad de los metadatos DICOM, los análisis de la exposición individual y de la población y el uso de recursos humanos y materiales. Además, ha sido posible adoptar mecanismos eficientes para adquirir y procesar los metadatos DICOM sin que ello repercutiera en el rendimiento del PACS.

Es por ello, que el rápido desarrollo de las tecnologías de la información y la comunicación (TIC) en las últimas décadas ha traído grandes cambios al ámbito de la salud, convirtiéndolo en uno más dinámico y global. El presente proyecto trata de propulsar una gran variedad de oportunidades de la salud personalizada y la dimensión de la prioridad estratégica de Biociencias/Salud.

Desarrollar una aplicación web basada en estos archivos médicos permite conocer a fondo cual es la estructura de ellos, como están almacenados y como se puede hacer uso de ellos. El motivo que me ha sido impulsado para realizar este proyecto como mi Trabajo de Fin de Carrera ha sido principalmente ese. Se trata de un ámbito completamente nuevo para mí y el cual se me ha sido posible formarme. A través de este proyecto he podido aplicar todo lo aprendido durante mi docencia junto con el conocimiento de nuevos aspectos.

Considero que ha sido una practica que ha aumentado mi conocimiento y me ha aportado gran valor al ámbito personal también, ya que gracias a él considero que estoy más capacitada para poder involucrarme en aspectos informáticos que sean nuevos para mí.

# Capítulo 7

## BIBLIOGRAFÍA

### Contenido

---

7.1. Bibliografía

---

## 7.1. BIBLIOGRAFÍA

- [1] Nevarez, E. L. (2016, 26 febrero). *Desarrollo de aplicaciones DICOM para la gestión de imágenes biomédicas*
- [2] Barberis, L. S. (2009, diciembre). *APLICACIÓN DE UN SISTEMA AUTOMÁTICO DE PROCESAMIENTO DE IMÁGENES MÉDICAS BASADA EN ESTANDARES*
- [3] I. (2021, 24 agosto). *¿Qué es el formato DICOM? Las claves del estándar en imágenes médicas.*
- [4] EcuRed. *Ficheros DICOM - EcuRed.*  
[https://www.ecured.cu/Ficheros\\_DICOM#Estructura\\_f.C3.ADsica\\_del\\_ficher\\_o](https://www.ecured.cu/Ficheros_DICOM#Estructura_f.C3.ADsica_del_ficher_o)
- [5] Universidad Tecnológica de Pereira. (2010, agosto). *DESCRIPCIÓN DEL ESTANDAR DICOM PARA UN ACCESO CONFIABLE a LA INFORMACIÓN DE LAS IMAGENES MÉDICAS.* Scientia et Technica Año XVI.
- [6] Carestream Health. (2015, julio). *Metadatos: creación de acceso relevante a imágenes e información clínicas para todos los usuarios.* CARESTREAM.  
[https://collaboration.carestream.com/sites/default/files/white\\_paper\\_ccp\\_a4\\_2015-07-20\\_es-es\\_0.pdf](https://collaboration.carestream.com/sites/default/files/white_paper_ccp_a4_2015-07-20_es-es_0.pdf)
- [7] Kahn, Jr, Charles & Langlotz, Curtis & Channin, David & Rubin, Daniel. (2010). *Informatics in Radiology: An Information Model of the DICOM Standard 1. Radiographics : a review publication of the Radiological Society of North America, Inc.* 31. 295-304. 10.1148/rg.311105085.
- [8] <https://1library.co/article/aspectos-b%C3%A1sicos-sobre-pacs-sistemas-pacs.qv1rr8dq>
- [9] García, M. V. (2012). *Diseño e integración de un sistema de conversión de video al estándar de comunicaciones DICOM con sistemas PACS de libre acceso.* Biblioteca digital.
- [10] <https://book.orthanc-server.com/users/rest.html#using-https>
- [11] Juan Carlos Battle, Keith Dreyer, Bibb Allen, Tessa Cook, Christopher J. Roth, Andrea Borondy Kitts, Raym Geis, Carol C. Wu, Matt P. Lungren, Jay Patti, Adam Prater, Daniel Rubin, Safwan Halabi, Mike Tilkin, Tom Hoffman, Laura Coombs, Christoph Wald, *Data Sharing of Imaging in an Evolving Health Care World: Report of the ACR Data Sharing Workgroup, Part 1: Data Ethics of Privacy, Consent, and Anonymization, Journal of the American College of Radiology,* Volume 18, Issue 12, 2021,  
<https://www.sciencedirect.com/science/article/pii/S1546144021005809>
- [12] H. (2020, 18 noviembre). *La Seudonimización de los datos y su importancia en el nuevo RGPD. Ayuda Ley Protección Datos.*
- [13] <https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf>
- [14] <https://dicom.nema.org/medical/dicom/current/output/chtml/>

- [15] <https://book.orthanc-server.com/users/anonymization.html#anonymization-of-a-single-instance>
- [16] Bayona, J. V. (2021, 4 febrero). DICOM. eHealths.
- [17] <https://es.mathworks.com/help/images/ref/dicomanon.html>
- [18] [https://dicom.nema.org/dicom/2013/output/chtml/part05/sect\\_6.2.html](https://dicom.nema.org/dicom/2013/output/chtml/part05/sect_6.2.html)
- [19] Milton Santos, Nelson Pacheco Rocha, Outcomes from Indexing Initiatives of Medical Imaging DICOM Metadata Repositories. A Secondary Analysis, Procedia Computer Science, Volume 138, 2018
- [20] Herranz, F. B. (2003, octubre). Desarrollo de aplicaciones DICOM para la gestión de imágenes biomédicas.
- [21] <https://www.orthanc-server.com/>
- [22] <https://code.visualstudio.com/docs/supporting/FAQ>
- [23] <https://desarrolloweb.com/home/angular>
- [24] <https://www.hostinger.es/tutoriales/que-es-angular>
- [25] <https://saasradar.net/mejores-framework-frontend/#:~:text=Se%20trata%20de%20un%20framework%20de%20c%C3%B3digo%20abierto%20con%20licencia,a%20las%20aplicaciones%20de%20escritorio.>
- [26] <http://www.jtech.ua.es/j2ee/restringido/grails/sesion03-apuntes.html>
- [27] <https://dinahosting.com/ayuda/que-es-mysql/>
- [28] <https://www.ticportal.es/glosario-tic/mysql>
- [29] <https://openwebinars.net/blog/que-es-mysql/>