

---

**LA DIGITALIZACIÓN Y LA CONSERVACIÓN DEL  
PATRIMONIO: EL PROYECTO PATRIMONIO DE  
CANTABRIA**

---

---

**MÁSTER DE PATRIMONIO HISTÓRICO Y TERRITORIAL  
CURSO 2012-2013**

**TRABAJO FIN DE MÁSTER REALIZADO POR: MARÍA CINTIA  
TORRELLAS GONZÁLEZ**

**DIRIGIDO POR: PROF. AURELIO Á. BARRÓN GARCÍA**

---

## ÍNDICE

|  |    |
|--|----|
| <b>PRESENTACIÓN</b> .....  | 3  |
| <b>1. EL PROYECTO DE NUESTRA ACTIVIDAD</b> .....                   | 4  |
| 1.1. INSTITUCIONES PARTICIPANTES .....                             | 4  |
| 1.1.1. Universidad de Cantabria .....                              | 4  |
| 1.1.2. Fundación Botín .....                                       | 6  |
| 1.1.3. Patrimonio y Lengua.....                                    | 7  |
| 1.2. Proyecto Ontología de Cantabria.....                          | 8  |
| 1.3. Proyecto Patrimonio de Cantabria: Web semántica 3.0 .....     | 14 |
| <b>2. ACTIVIDADES REALIZADAS</b> .....                             | 21 |
| 2.1. Introducción a la digitalización.....                         | 21 |
| 2.2. El local y los medios de que dispone.....                     | 23 |
| 2.3. Tareas realizadas: el Archivo de la Autoridad Portuaria ..... | 30 |
| 2.3.1. Digitalización.....   | 31 |
| 2.3.1.1. Escaneado.....  | 31 |
| 2.3.1.2. Tratamiento de las imágenes .....                         | 40 |
| 2.3.1.2.1. Reconocimiento óptico previo.....                       | 40 |
| 2.3.1.2.2. Retoque informático .....                               | 41 |
| 2.3.2. Generación de ficheros de servicio.....                     | 47 |
| <b>3. PROBLEMAS Y POTENCIALIDADES</b> .....                        | 50 |
| <b>4. CONCLUSIONES</b> .....                                       | 51 |
| <b>5. BIBLIOGRAFÍA Y RECURSOS WEB</b> .....                        | 53 |

## **PRESENTACIÓN**

La presente memoria versa sobre las prácticas realizadas desde el 2 de Abril al 5 de Junio de 2013 en la Unidad Técnica de Cantabria Campus Internacional marco del proyecto *Patrimonio de Cantabria: Web semántica 3.0.* donde he tenido la posibilidad de realizar las 250 horas referentes a las prácticas externas del Máster de Patrimonio histórico y territorial de la Universidad de Cantabria.

En dicho escrito se ha realizado una breve introducción sobre las instituciones participantes en el proyecto, así como los estudios previos y un desarrollo del mismo.

Posteriormente se desarrollan las actividades realizadas durante el periodo mencionado, con una breve introducción al concepto de digitalización, la descripción del local y los medios que posee para, más tarde, centrarnos en las actividades realizadas.

En la conclusión de las prácticas se ha realizado una valoración de los problemas y potencialidades del proyecto, y una justificación sobre la importancia del Patrimonio así como de su difusión, divulgación y las iniciativas llevadas a cabo para tal fin.

Agradezco a María Celestina Losada, Lorena Sainz y Sara Melendi por permitirme ser partícipe de un proyecto de tal envergadura, por enseñarme las técnicas de la digitalización y hacerme copartícipe del equipo.

Agradezco a Manuel Suarez-Cortina y a Aurelio Barrón, por los acuerdos llevados a cabo para que fuera factible la realización de las prácticas en dicho proyecto.

## **1. EL PROYECTO DE NUESTRA ACTIVIDAD**

El proyecto es resultado de la elección de la Universidad de Cantabria como Campus de Excelencia en 2009 y resultado del acuerdo, en 2010, de Cantabria Campus Internacional con la Fundación Botín. Estas entidades son participes, dentro de la línea de investigación Patrimonio y Lengua, de un proyecto de investigación a escala regional.

### **1.1. INSTITUCIONES PARTICIPANTES**

#### **1.1.1. Universidad de Cantabria**

Una de las instituciones participantes en el proyecto es la Universidad de Cantabria, entidad pública configurada por dos campus, uno situado en Santander llamado de las Llamas y otro en Torrelavega ubicado en Tanos.

Creada como institución en 1927 con el nombre de Universidad de Santander, pasó en 1940 a denominarse como la conocemos en la actualidad, Universidad de Cantabria.

Las primeras escuelas en surgir fueron las Escuelas profesionales de Comercio y de Náutica, las cuales aparecieron gracias a la prosperidad de la ciudad como foco portuario. Más tarde, se establecieron las Escuelas de Derecho hasta llegar a 1978 con la creación de la última facultad, Filosofía y Letras. Finalmente, años más tarde las Escuelas de la Marina civil y la de Enfermería se incorporarían como nuevas titulaciones universitarias.

Fue seleccionada, en 2009, por el Gobierno de España como Campus de Excelencia Internacional, dando paso al proyecto Cantabria Campus Internacional.

### Cantabria Campus Internacional

Cantabria Campus Internacional es un proyecto impulsado por el Gobierno de España dentro del programa Campus de Excelencia Internacional.

El proyecto está basado en la cooperación institucional y empresarial con la finalidad de alcanzar un reconocimiento en el ámbito del conocimiento, con recursos académicos, científicos-tecnológicos y de transferencia de la región autónoma.

El proyecto se estructura en cuatro áreas estratégicas: Agua y Energía; Biomedicina y Biotecnología; Banca, Finanzas y Actividad Empresarial; y Patrimonio y Lengua.

Para poder llevar a cabo esto, se debían cumplir una serie de requisitos:

- Había que destacar de forma científica y docente con un nivel excelente.
- Debía de ser una universidad internacional.
- Había que desarrollar de forma completa la cadena de transferencia del conocimiento.
- Debía poseer un soporte de infraestructuras alto.
- Tenía que haber un reconocimiento intrínseco de la región.
- Había que tener una cooperación con diversas instituciones.
- Para el futuro, se perseguía la creación de empresas y entidades que generaran recursos.

La Universidad de Cantabria, cumplió nueve de los diez requisitos, por lo que en 2009 fue una de las nueve primeras universidades seleccionada por el Gobierno de España.<sup>1</sup>

---

1 Cantabria Campus Internacional:  
<http://www.cantabriacampusinternacional.com/ceiOfrece/Paginas/Que-es-CCI.aspx> (Consultado el 22/08/2013).

### 1.1.2. Fundación Botín

La otra institución participe, esta vez de carácter privado, es la Fundación Botín.

Creada en 1964 por Marcelino Botín Sanz de Sautuola y su esposa Carmen Yllera, con el fin de "paliar las necesidades y promover el desarrollo social" de Cantabria. Hoy en día, la Fundación Botín se ha convertido en la primera fundación privada de España en volumen de inversión social.

Tal y como la Fundación Botín explica en su portal web: "la misión de la Fundación Botín es contribuir al desarrollo integral de la sociedad, detectando y apostando por el talento creativo que hay en ella y explorando nuevas formas de generar riqueza cultural, social y económica."

Estos fines son desarrollados a través de varias líneas estratégicas: Desarrollo artístico y cultural; Desarrollo socioeconómico (educación, ciencia, desarrollo rural); Acción social y el Observatorio de Tendencias. Este Observatorio, cuenta con programas para la detección y desarrollo de talentos, tanto en los sectores públicos como privados.<sup>2</sup>

En cuanto a las sedes, la Fundación posee seis infraestructuras. Cuatro de ellas en Santander, la sede principal está situada en la calle Pedrueca, antigua casa del famoso paleontólogo Marcelino Sanz de Sautuola, donde se llevan a cabo numerosas actividades entre las que destacan cursos y conferencias, además de estar dotada de una biblioteca con una importante colección bibliográfica. Otra sede es la sala de exposiciones situada en la calle Marcelino Sanz de Sautuola, donde se realizan exposiciones llevadas a cabo por el programa de artes plásticas de la propia Fundación.

Las otras dos sedes en la capital cántabra se sitúan en el Paseo Pérez Galdós, el Promontorio, casa familiar del empresario Emilio Botín y Villa Iris, dedicada a la investigación y actividades de formación.

---

2 Fundación Botín: <http://www.fundacionbotin.org/la-fundacion-mision.htm> (Consultado el 22/08/2013).

El resto de infraestructuras las encontramos en el municipio de Polaciones (Cantabria), la Casa Rectoral de Puente Pumar, centro de operaciones de la Fundación Botín en el Valle del Nansa desde donde se gestiona y mantiene como parte del programa Patrimonio y Territorio. Y por último en Madrid, donde se aloja el observatorio de Tendencias y el equipo gestor del programa Ciencia y Transferencia Tecnológica.

Además, existe un proyecto en construcción, el Centro Botín, cuya apertura se estima que ocurrirá en 2014.

### 1.1.2. Patrimonio y Lengua

La Fundación Botín, como se ha mencionado anteriormente, sigue diversas estrategias de intervención en materias dispares. Una de estas líneas se dirige hacia la investigación y, en el campo de las humanidades, destaca la cooperación con Cantabria Campus Internacional, especialmente en el área de Patrimonio y Lengua.

Colabora con dicho campus en torno a cinco áreas de conocimiento: Patrimonio y Desarrollo Rural; Patrimonio Cultural y Natural; Prehistoria; Ontología del Patrimonio y Lengua Española.

En la actualidad, las tres últimas áreas correspondientes a Prehistoria, Ontología del Patrimonio y Lengua Española, son las que desarrollan los proyectos más innovadores.

En el área de Prehistoria, la riqueza arqueológica de Cantabria hizo que se creará el Instituto Internacional de Investigaciones Prehistóricas (IIIP) gracias a la colaboración entre la Universidad de Cantabria, el Gobierno Regional y el Banco Santander. Este Instituto está asociado al CSIC y sus investigaciones están relacionadas con el Museo Altamira.

En la actualidad cuenta con numerosas investigaciones tanto de ámbito nacional como internacional, así como publicaciones, libros, etc.

En lo referente a la Lengua Española, la Universidad Internacional Menéndez Pelayo (UIMP), institución de enseñanza de la lengua española, y la Fundación Comillas, centro de conocimiento del español y su cultura, sirven como contexto del Centro Internacional de Estudios Superiores del Español, impulsando las investigaciones para dar una gran capacidad formativa y divulgativa.

Actualmente, se intenta impulsar el conocimiento de la lengua española con múltiples cursos así como con un máster en Enseñanza del Español como Lengua extranjera.

En cuanto a la Ontología del Patrimonio, hay varios proyectos sobre preservación y difusión del Patrimonio de Cantabria, como son la web semántica del Patrimonio de Cantabria o el proyecto Patrimonio y Territorio. Con la presencia de la Fundación Albériz en Cantabria y su amplio estudio sobre la música se completan los estudios sobre Patrimonio en este área.<sup>3</sup>

### 1.2. Proyecto Ontología de Cantabria

El proyecto *Ontología de Cantabria* surgió en el año 2006, promovido y financiado por la Fundación Marcelino Botín en una de sus líneas estratégicas más importantes como es la protección y difusión del Patrimonio de Cantabria.

---

3 Cantabria Campus Internacional:  
<http://www.cantabriacampusinternacional.com/areasExcelencia/cipl/Paginas/Presentacion-Patrimonio-y-la-Lengua.aspx> (Consultado el 25/06/13).

Contó con la colaboración del Departamento de Historia Moderna y Contemporánea de la Universidad de Cantabria (UC), donde se formó un equipo de investigación dirigido por Manuel Suárez-Cortina, catedrático de historia contemporánea de la UC, y coordinado por Gonzalo Capellán, vicerrector en ese momento. Además se integró a la empresa española iSOCO que se encargó de la construcción e implantación de los recursos informáticos necesarios para la puesta en marcha del proyecto.<sup>4</sup>

El objetivo fue reunir de forma ordenada el conocimiento bibliográfico existente sobre Cantabria en sus diferentes variantes, desde el Patrimonio histórico hasta el etnográfico, pasando por cada una de sus diversas manifestaciones. Los tipos de Patrimonio contemplados en esta Ontología son aquellos establecidos por la Ley vigente de Cantabria, además de otro tipo de Patrimonio, más temporal como por ejemplo los periodos cronológicos o periodos culturales.

El problema que surgió, desde el principio, fue cómo abordar la gran variedad de temas y cómo integrar la información existente, pues ésta se encontraba de forma fragmentada en fuentes dispares. Tras estudiar los diversos casos se pretendió unificar de forma estructurada toda esta información.

La Ontología de Cantabria busca servir al público como un sistema de información contextualizada e interrelacionada. Este tipo de sistema informativo, recogido de diferentes fuentes, necesita un diseño de estructuras de información y navegación que muestren las relaciones para poder realizar las búsquedas, una especie de base de datos que permita mediante una misma codificación relacionar toda la información.

---

4 HERNÁNDEZ CARRASCAL, F., “Ontología del Patrimonio de Cantabria”, en: *EL profesional de la información*, 2008, v.17, nº 1, pp. 92-98. (Consultado el 26/06/2013). Disponible en: <http://www.elprofesionalde lainformacion.com/contenidos/2008/enero/11.pdf>, p. 92.

Todo ello tendrá como fruto el desarrollo de un portal especializado en el Patrimonio de Cantabria que mostrará a los usuarios diferentes campos unidos entre sí por relaciones semánticas desde una unidad de información básica a sus atributos y propiedades accediendo por enlaces directos.

El problema de esto es que parte de la información existente sobre el Patrimonio de Cantabria recogida es escasa y primaria, por lo que la información contextual es difícil de registrar y, por lo tanto, hace complicada la tarea de su integración en las relaciones semánticas.

Estas lagunas se pretenden solventar con la interrelación de campos que aumentarían las relaciones entre los datos y que crearían más puntos de información en común. Sucede que pueden existir uniones que a veces son inapreciables y con este sistema se intentan vincular.

Para poder crear estas relaciones semánticas se utilizó el modelo de referencia conceptual del Comité Internacional del ICOM para la Documentación (CIDOC), que proporciona un campo semántico común y extensible para que cualquier información sobre Patrimonio cultural pueda ser mapeada permitiendo la unión entre diferentes campos de información.<sup>5</sup>

La recopilación de la información para poder llevar a cabo el proyecto se hizo con las técnicas establecidas por la bibliografía para la categorización y organización de los diferentes documentos, es decir se fue desde un primer nivel con bibliografía de bibliografías hasta un último nivel donde se incorporan los documentos primarios.

---

5 CIDOC: <http://www.cidoc-crm.org/> (Consultado el 12/08/2013).

Tal como se ha distribuido, la información se encuentra en cuatro niveles. Un primer nivel donde estará todo lo publicado y estudiado sobre Cantabria, donde un pilar fundamental ha sido *Historia de Cantabria: un siglo de historiografía y bibliografía (1900-1994)*. Era la bibliografía más completa conocida hasta el momento sobre Cantabria, aunque necesitaba ser actualizada incluyendo el período comprendido desde 1994 hasta 2006. Esto fue llevado a cabo y realizado por un equipo de la Universidad de Cantabria, integrado por Manuel Suárez-Cortina y Gonzalo Capellán.

En un segundo nivel, estarán las descripciones de las unidades que forman parte del Patrimonio de Cantabria y las instituciones que lo configuran, gestionan y difunden. Para esta sección fueron de importancia los catálogos e inventarios realizados, entre otros recursos útiles.

En un tercer nivel se encontrará la gestión que existe sobre el patrimonio, desde la legislación y difusión hasta las entidades que se encargan de ello.

Y por último, se hallará la información básica sobre ese patrimonio, como son los nombres, los lugares, las fechas, etc.<sup>6</sup>

Además de toda la información existente en diferentes soportes, se incluyeron las copias digitales de algunos documentos. Decir que, muchas de las fuentes de información que se utilizaron no se encontraban en formato digital y por lo tanto tuvieron que hacerse procesos de informatización y hacer cambios en las bases de datos para conseguir normalizar los documentos y fuentes no normalizados.

---

6 HERNÁNDEZ CARRASCAL, F., “Ontología del Patrimonio de Cantabria”, en: *EL profesional de la información*, 2008, v.17, nº 1, pp. 92-98. (Consultado el 26/06/2013). Disponible en: <http://www.elprofesionalde lainformacion.com/contenidos/2008/enero/11.pdf>. Pp. 93-94.

Para este cambio se siguió un sistema de codificación basado en la Normativa Nacional e Internacional para Bibliotecas, Archivos y Museos. Esto se realizó de este modo, debido a q si se debía integrar nueva información en el proyecto con un sistema de codificación diverso al utilizado, esto fuera posible y, además, la web de este modo servirá como complemento a las páginas de las instituciones originarias creando enlaces a las mismas. De este modo será más fácil acceder y proteger la integridad de la documentación.

A la hora de realizar el proyecto se dieron cuenta de que existía una carencia considerable en cuanto a la codificación, ya que no está normalizada y hay problemas a la hora del intercambio de información. Esto no es sólo un problema de Cantabria sino de toda España, lo que crea un inconveniente a la hora de crear relaciones de datos entre las diferentes organizaciones de las regiones autonómicas.<sup>7</sup> Una vez terminado el portal, aunque no está abierto al público, se ha distribuido la información en torno a las siguientes áreas: Personas, Instituciones, Periodos cronológicos y Eventos, Obras y Monumentos, Lugares y Tipos, y Conceptos.

Para facilitar el acceso a la información se insertó un buscador semántico que permite realizar búsquedas muy concretas y diferentes, es decir, si se quiere buscar el patrimonio inmueble realizado entre 1939 a 1949, nos dará por medio de interrelaciones los resultados concretos.

Mediante un mapa interactivo se podrán geoposicionar todos los datos del patrimonio de Cantabria, así como establecer rutas entre diferentes puntos y a partir de toda la información recopilada se podrán crear recursos didácticos digitales.

---

7 HERNÁNDEZ CARRASCAL, F., “Intercambio de información sobre patrimonio cultural: la Ontología del Patrimonio de Cantabria”, en *XI Jornadas de Gestión de la Información: Servicios polivalentes, confluencia entre profesionales de archivo, biblioteca y documentación*, Madrid, 2009 (Consultado el 26/06/2013). Disponible en: <http://eprints.rclis.org/13938/1/XIJGI-Hernandez.pdf>. Pp.137-152.

Asimismo, existen otras herramientas útiles como un instrumento que analiza las acciones realizadas por cada usuario donde se pueden practicar con propuestas concretas de actuación. Hay otra utilidad que permite la realización de rutas turísticas junto con coordenadas GPS para el coche e información sobre la zona a visitar a interés de cada usuario.

Se integró un tipo de Wikipedia semántica que permite la aportación de datos, comentarios, documentos etc.



1. Página de inicio al portal de Patrimonio de Cantabria

De cara al futuro en que la web esté completada, Francisca Hernández, directora del proyecto y asesora técnica de la Fundación Botín, propone crear una ontología integradora o crear una conexión entre las diferentes ontologías parciales que existen para que a la hora de hacer una consulta la información se siga un mismo modelo y la integración de más información sea posible.

El proyecto inicial finalizó en 2009, momento en el que comenzó un nuevo proyecto llamado *Patrimonio de Cantabria: web semántica 3.0*.

### 1.3. Proyecto Patrimonio de Cantabria: Web semántica 3.0

El proyecto *Patrimonio de Cantabria: Web semántica 3.0* se origina del proyecto ontológico de Cantabria y del proyecto de actuación integral sobre el Valle del Nansa, trabajos anteriores realizados también en colaboración por la Universidad de Cantabria y la Fundación Botín. Este convenio ha sido realizado en el marco del Campus de Excelencia Internacional, más exactamente en el área de Patrimonio y Lengua, ya explicados anteriormente, con el objetivo de digitalizar la documentación generada por los diversos entes tanto públicos como privados.

Su inicio fue en verano (Agosto) de 2010, con la adaptación de las instalaciones habilitadas para el proyecto en la calle Rualasal donde se instaló un equipamiento informático y de digitalización, más un equipo humano cualificado para las distintas actividades del proyecto.

Este equipo humano se divide en dos sectores. Por un lado el equipo científico cuyo objetivo es determinar la naturaleza de los materiales, objeto de digitalización e incorporación al portal, además de establecer unas tipologías, la prioridad del trabajo y el calendario de realización. En este grupo hay un coordinador general, donde en una primera instancia correspondió a Gonzalo Capellán y, posteriormente hasta la actualidad, a Manuel Suárez Cortina, cuyo cometido en un primer momento fue la dirección científica.

Este equipo, dentro de las actividades mencionadas anteriormente, se encarga de las actividades de selección de contenido, llevadas a cabo por un grupo de especialistas en diferentes sectores. Estos son:

- Historia y Derecho: Manuel Estrada, Profesor titular de Historia del Derecho.
- Geografía y Ordenación del Territorio: Leonor de la Puente, Profesora titular de Geografía Humana.

- Lengua y Literatura: Raquel Gutiérrez Sebastián, Profesora de Lengua y Literatura Española.
- Arte y Patrimonio: Miguel Ángel Aramburu Zabala, Catedrático de Historia del Arte.
- Ciencias y Técnicas Historiográficas: Virginia Cuñat, Profesora Titular de Ciencias y Técnicas Historiográficas.

Por otro lado se encuentra el equipo técnico y de gestión, encargados del tratamiento digital e informático de los materiales previamente seleccionados. Este grupo está formado por cuatro personas: dos digitalizadoras, una informática y una coordinadora:

- Coordinadora en gestión y responsable de la Unidad Técnica: María Celestina Losada Varea, Doctora en Historia del Arte.
- Informática: Sara Melendi, Ingeniera en Telecomunicaciones.
- Digitalizadoras: Lorena Sainz y Eloisa González, Licenciadas en Historia.

A este último equipo, he pertenecido durante mis prácticas con la función de digitalizadora, en ausencia de Eloisa González.

Además de todo el personal, la Fundación Botín cuenta con una asesora técnica, Francisca Hernández, documentalista y especializada en el tratamiento y digitalización de la documentación.<sup>8</sup>

El objetivo principal de esta Unidad Técnica es la adquisición y gestión de contenidos digitales sobre Patrimonio para la integración de los mismos en el Portal del Patrimonio de Cantabria. Se persigue crear una relación entre todos los campos pudiendo pasar de unos elementos a otros sin una relación aparente.

---

<sup>8</sup> CAPELLÁN, G., *Web semántica del Patrimonio de Cantabria. Informe de Septiembre de 2010* (Consultado el 25/06/2013). Informe cedido por la Coordinadora en gestión y responsable de la Unidad Técnica, María Celestina Losada.

Además, se pretende acumular todo el conocimiento existente sobre el Patrimonio de Cantabria para posteriormente difundirlo de forma estructurada proporcionando un medio para la consulta de la información que atienda a las necesidades de los diferentes usuarios como continuación del Proyecto Ontología de Cantabria, desarrollado con anterioridad.

Todo ello contribuye a enriquecer, conservar y hacer visible, a través de la difusión, nuestro patrimonio.<sup>9</sup>

Por ejemplo, en la biografía de Menéndez Pelayo, entre sus muchas relaciones, habrá una que será la Biblioteca Menéndez Pelayo. A través de la unión semántica se puede acceder a la información sobre la Biblioteca y dentro de ésta se incluiría en localización la palabra Santander, a través de un enlace directo se podrá acceder a información sobre Santander. Y así sucesivamente.

Pero, ¿Qué es una web semántica? Extraigo dos definiciones del libro *Tendencias en documentación digital*<sup>10</sup> y cito literalmente:

“1. La Web Semántica es un conjunto de iniciativas destinadas a promover una futura Web cuyas páginas están organizadas, estructuradas y codificadas de tal manera que los ordenadores sean capaces de efectuar inferencias y razonar a partir de sus contenidos.

2. La Web Semántica es un conjunto de iniciativas destinadas a convertir el World Wide Web en una gran base de datos capaz de soportar un procesamiento sistemático y consistente de la información.”

---

9 Fundación Botín: [http://www.fundacionbotin.org/creada-la-unidad-de-trabajo-para-ampliar-y-mantener-los-contenidos-de-la-web-de-cantabria-de-la-fundacion-botin\\_publicacion-fundacion-botin-12037422.htm](http://www.fundacionbotin.org/creada-la-unidad-de-trabajo-para-ampliar-y-mantener-los-contenidos-de-la-web-de-cantabria-de-la-fundacion-botin_publicacion-fundacion-botin-12037422.htm) (Consultado el 25/06/2013).

10 TRAMULLAS J. (Coord), *Tendencias en documentación digital*, Gijón, Ediciones Trea S.L., 2006. P.10.

En resumen, la web semántica es un tipo de base de datos trasladada a la red. Lo que se intenta es que en vez de realizar búsquedas con caracteres se puedan hacer por conceptos, es decir que el usuario a través de una búsqueda con un objetivo, aunque no sea conciso, sea ayudado por el software para elaborar una estrategia interpretativa de la información y pueda adaptarla al contexto. Esta información está compuesta por lenguajes documentales, metadatos y ontologías para responder de forma acertada al usuario.

Para la realización de este proyecto se hicieron diferentes bloques, extraídos de las áreas realizadas para la Ontología de Cantabria:

- Personas.
- Instituciones.
- Períodos cronológicos y Eventos.
- Obras y Monumentos (Bienes muebles e inmuebles).
- Lugares y Tipos (Parques naturales, elementos geográficos destacados etc.).
- Conceptos (Planes de protección).

Además se volcó la Gran Enciclopedia de Cantabria, cuya primera edición se llevó a cabo en 1985 dirigida por Jesús Pindado Uslé. Esta primera publicación estaba formada por 8 tomos. Más tarde, en 2002, se hizo una segunda edición con una ampliación de tres tomos más que completaba el periodo cronológico comprendido entre 1985 hasta 2002. Ésta poseía las fichas con la información recogida hasta el momento y sirvió de base.

Aparte, para cada una de las instituciones y personas relacionadas con el Patrimonio de Cantabria se elaboraron fichas informativas con el programa Excel. Se realizaron fichas modelo, donde se rellenaron los campos con los datos correspondientes y dentro de estos campos, se debía plasmar de dónde se había extraído la información así como el enlace con acceso directo a la fuente original. Esta base de datos se actualiza periódicamente.

En el caso del Patrimonio inmueble y mueble, se crearon fichas con fotografías, además de digitalizar los libros que contienen información útil referente a dicho Patrimonio. Un ejemplo de ello es el libro *Altamira* de Hugo Obermayer, publicado en 1911.

En lo que respecta a la documentación histórica, la lista es más amplia. Se han digitalizado imágenes, postales antiguas de la región, los primeros diez años del Diario Montañés, los veinte primeros años de las publicaciones de la Cámara de Comercio, parte de los libros publicados por la Fundación Botín y por la Universidad de Cantabria.

De la Biblioteca Municipal se han digitalizado cinco obras de Joaquín Telesforo Trueba y Cossío y trece de Baldomero Madrazo. Del Colegio de Arquitectos se han digitalizado los planos de Leonardo Rucabado, sobre todo los correspondientes a sus obras de Bilbao, aunque también hay algunos de sus proyectos en Santander. Igualmente se han incorporado los datos conservados en el Archivo de la Confederación Nacional del Trabajo (CNT) y en la Fundación Botín.

Del Centro de Documentación e Investigación de Música de Cantabria (CDIMC) se ha realizado la digitalización de las partituras del fondo Antonio Gorostiaga, Arturo Duo Vital y José Lucio Mediavilla. En el caso de estos fondos se puso en la portada de cada partitura una tabla de colores para la corroboración del color de la misma, ya que el centro pidió que tuviera que ser de determinado tono. Además, se incorporó una ficha de catálogo proporcionada por ellos.

Otro fondo interesante que se ha incorporado al proyecto es el del arquitecto Javier González de Riancho del que se extrajeron planos y dos álbumes, uno de edificios característicos de Cantabria y otro de asuntos variados que había realizado él, en este último alternaba dibujos con fotografías.

Además de la gran cantidad de archivos mencionados anteriormente, otras entidades privadas cedieron parte de sus documentos para la digitalización de los mismos.

Una vez digitalizados estos documentos, se prosiguió a guardarlos en formato digital. Para ello se utilizó un patrón de registros.

La mayoría de la documentación es guardada en formato eXtensible Markup Language (XML). XML es un sistema de metadatos que permite definir tipos de documentos y los conjuntos de etiquetas necesarias para etiquetar dichos documentos.

Cuando las referencias XML de un documento ya existen en la red, como ocurre en el caso de los libros de la Universidad de Cantabria o las Bibliotecas en general, se cogen de las web las referencias XML de la misma. En cambio, si esta información no existe, se realiza una ficha modelo. En el proyecto, esta ficha modelo ha sido elaborada en Madrid.

Posteriormente, se realizan para cada uno de los documentos fichas METS, las cuales sirven para cuando un Portable Document Format (PDF) excede de un tamaño razonable. En este caso, el documento se divide en diferentes partes con menor tamaño sin tener que realizar varios PDF.

Dentro de las actividades realizadas en el proyecto se encuentran:

- El establecimiento de relaciones y contactos con vista a la digitalización de fondos
- Digitalización de la documentación cedida por los contactos previamente acordados.
- Se realiza una descripción de los objetos digitales anteriormente ejecutados.
- Se efectúa un mantenimiento en red del Portal del Patrimonio de Cantabria.

### Establecimiento de relaciones y contactos con vista a la digitalización de fondos

En un primer momento, las entidades, instituciones y particulares entran en contacto con el director de la Unidad científica, en este caso Manuel Suárez. Éste se encarga de que estos vean las instalaciones, se les explica qué se va a hacer y cómo.

Este proceso suele durar varios meses desde que se inician los contactos hasta que se firma el convenio.

La firma de acuerdos tiene un efecto positivo dentro del proyecto, ya que influye en el próximo convenio dado el éxito del anterior, esto ocurre sobre todo si son acuerdos que producen resultados.

El encargado de firmar los convenios necesarios es la Fundación Botín, que aporta una asesoría jurídica para que los convenios tengan un valor legal y para que el uso de la documentación sea el adecuado atendiendo a los derechos de autor.

La documentación recibida es evaluada por el Comité científico, el cual se encarga de dar una fecha aproximada para la ejecución del encargo, y es entonces cuando interviene la Unidad Técnica, responsable de ordenar la documentación, seleccionarla y manipularla en el periodo de tiempo establecido por dicho Comité.

Una vez hecho esto, la documentación pasa a formar parte del proyecto, donde se digitalizará y guardará en formato PDF. Una de las copias va para la entidad que la ha cedido y la otra va destinada para ser volcada en la red de Patrimonio de Cantabria 3.0, siempre que pueda ser publicada.

Los siguientes pasos realizados en el proyecto serán desarrollados a continuación, ya que han sido parte de las tareas efectuadas durante el periodo de prácticas.

## **2. ACTIVIDADES REALIZADAS**

### **2.1. Introducción a la digitalización**

Para poder realizar una introducción al concepto de digitalización es imprescindible definir qué es digitalizar. El término digitalizar es definido como: “Convertir cualquier señal de entrada continua en una serie de valores numéricos haciendo así que un material físico pase a archivo digital mediante equipos especializados para ello”.<sup>11</sup>

Esta idea hace referencia al proceso de crear imágenes digitales y por consiguiente almacenarlas en soportes electrónicos, ya sean discos duros, Internet, etc.

Como los procesos de digitalización, son duraderos y costosos, para poner en marcha el desarrollo de un proyecto de digitalización hay que tener en cuenta diversos factores. Entre estos se encuentran los requisitos técnicos que se poseen, los aspectos legales pertenecientes a los materiales, el presupuesto que se debe considerar a la hora de realizar un proyecto de tal envergadura, los recursos humanos que son necesarios para poder llevar a cabo la tarea, el desarrollo y mantenimiento de la infraestructura web donde se quiera realizar, así como en un futuro, la preservación de ese contenido digital.

---

<sup>11</sup> Diccionario de informática: <http://www.alegsa.com.ar/Dic/digitalizar.php> (Consultado el 27/06/2013).

De este modo, todos los factores mencionados anteriormente hacen que un proyecto de digitalización no sea algo sencillo y de fácil factura, sino que sea una tarea que debe llevar una planificación previa y un programa estructurado. Y debido a esto, se debe valorar que, dentro de un proyecto de digitalización existan tanto ventajas como inconvenientes.

Dentro de las ventajas existentes destaca el incremento de acceso a diversos usuarios, permitiendo una rápida aproximación a la información sin ser necesaria la consulta física del material, mejorando así, los servicios y proporcionando cierta comodidad al usuario. Esto, por tanto, desencadenaría en múltiples utilidades. Al digitalizar un documento y hacer posible al usuario la visión de una copia digital, se reduce la manipulación del mismo y, por consiguiente, el uso de los materiales originales frágiles, de este modo se asegura la futura preservación y conservación del documento.

En cuanto a las instituciones competentes en la materia, la digitalización da la oportunidad para el desarrollo de sus infraestructuras técnicas así como de sus empleados, proporcionando un conocimiento avanzado en esta clase de tecnología. Además, se establecen recursos cooperativos con otras instituciones para la creación de colecciones virtuales en conjunto y poder ofrecer un recurso electrónico más rico en información. Por lo que, esta cooperación resulta ventajosa debido a que se comparten los gastos económicos y, por lo tanto, rentabiliza el proyecto.

Como se ha mencionado anteriormente, en un proyecto de digitalización cabe la posibilidad de encontrarnos con que dentro del mismo haya también inconvenientes.

La digitalización resulta, por tanto, un proceso costoso, tanto por el uso de tecnologías avanzadas como por el mantenimiento de las mismas. Además de resultar oneroso, los procesos de digitalización son de larga duración y dificultosos, pues requieren la realización de diferentes pasos de actuación. Dentro estas partes se encuentran la manipulación que se debe hacer al documento, su catalogación, el escaneado, el tratamiento digital posterior, la conservación electrónica del mismo así como la creación de una infraestructura digital para la puesta en uso por otros usuarios.

La digitalización tiene como fin satisfacer al mayor número de usuarios posibles y dependiendo del origen de la entidad se complace a unos usuarios u otros. Si es una entidad privada tenderá a un proyecto concreto y por lo tanto a unos consumidores específicos, en cambio si es pública deberá haber una mayor diversificación del contenido, haciendo así que haya más disparidad entre los beneficiarios.

Esto es posible gracias a que la digitalización da acceso a materiales a los que no es posible acceder ya sea del tipo que sea, desde manuscritos hasta imágenes originales. Estos materiales en muchos casos solo se podrían consultar en el lugar donde se encuentran, con carnet de investigador en otras, o incluso hay algunos que no es posible su consulta por el estado en el que se encuentran.<sup>12</sup>

### 2.2. El local y los medios de que dispone.

El local está situado en la calle Rualasal número 6, sitio que ocupaba la antigua librería universitaria, la cual ha sido reformada y diseñada conforme al uso al que estaba destinado el proyecto.

---

<sup>12</sup> VVAA, *Directrices para proyectos de digitalización de colecciones y fondos de dominio público, en particular para aquellos custodiados en bibliotecas y archivos*, España, Ministerio de Cultura, 2002. Pp. 9-13.

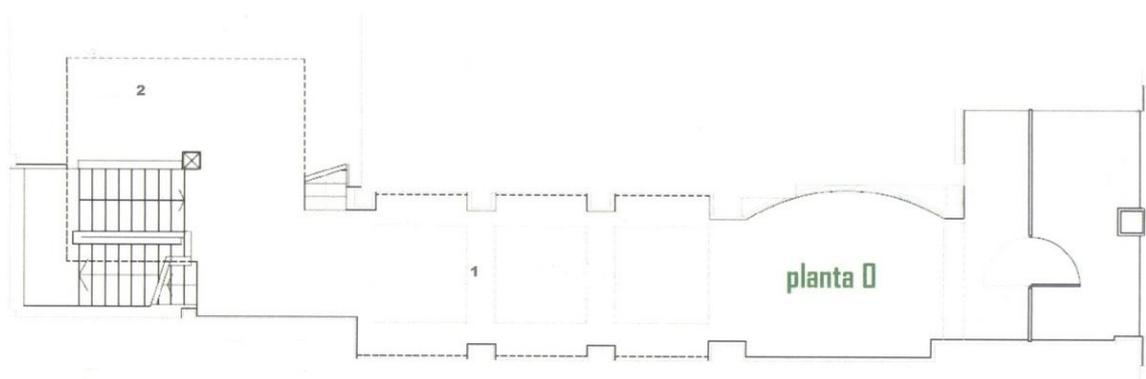
## Memoria de prácticas

El espacio se distribuye en dos plantas, una primera de 103 metros cuadrados y un sótano con 142 metros cuadrados, con un total de 245 metros cuadrados.

Dentro de la primera planta, nos encontramos con vestíbulo donde hay un pequeño mostrador que da acceso a dos despachos, marcados en el plano con el número 1. Estas mesas de trabajo corresponden a María Celestina Losada, encargada de la Unidad, y a Sara Melendi, ingeniera en Telecomunicaciones. Ambas están dotadas de equipamientos informáticos, al igual que con diversas estanterías destinadas a diferentes usos.

Prosiguiendo con la primera planta, nos encontramos con el espacio numerado con un 2, que corresponde a la situación de tres mesas de trabajo dotadas con su equipo informático y los escáneres utilizados para el proyecto.

El sótano está distribuido de la siguiente manera: en la zona inmediata a las escaleras nos encontramos con un almacén provisional donde están aglomeradas diversas cajas pertenecientes a los equipos informáticos arriba situados. En las zonas marcadas como 1 y 2 se hallan otras zonas de almacenaje. En una de ellas están situados los discos externos. Por último, en la zona 3 está el espacio destinado al aseo.



## Memoria de prácticas



2. Planos correspondientes a la primera planta y el sótano



3. Imagen de la entrada



4. Imagen correspondiente a los despachos principales



5. Imagen de mi mesa de trabajo durante las prácticas

El establecimiento está dotado de una tecnología avanzada con varios equipos de digitalización, donde cada uno de ellos está especializado en una tarea distinta y por lo tanto se ha utilizado para diferente documentación.

Los tipos de escáner son:

- i2s Digibook Scanner Copybook:

Este escáner cuenta con unas medidas de 810mm x 990mm, y pesa alrededor de 70 kg. Gracias a las dimensiones que posee permite que se puedan digitalizar documentos encuadernados de hasta un formato DIN A2.

Entre las ventajas que detenta este escáner se encuentra su funcionamiento con luz ambiente, por lo que no necesita ninguna luz adicional. Esto, permite una mejor conservación de la documentación antigua gracias a la ausencia de rayos ultravioleta e infrarrojos.

## Memoria de prácticas

El escáner tiene un uso asequible, pues permite un uso automatizado mediante la elección de los ajustes automáticos que se amoldan perfectamente a cada tipo de documentación. Esto es posible gracias a los avanzados modos de captura que utiliza así como por los algoritmos integrados que posee el escáner.

Además, tiene un diseño tecnológico avanzado que se ajusta a las necesidades del trabajador. Ofrece una gran comodidad en el entorno de trabajo, acompañado de una pantalla de cristal líquido (LCD), de un ratón táctil, etc. Debido a estos avances tecnológicos, la calidad de la imagen es de una alta resolución, que oscila entre los 300 pixeles por pulgada (PPP) hasta los 600. A lo largo del proyecto, hemos trabajado en la unidad con una medida estándar de 400 PPP aunque para imágenes de alta calidad se ha llegado a usar hasta 600.

Asimismo, son de una alta productividad debido a su rapidez a la hora de trabajar. Esto hace que el tiempo de digitalización varíe dependiendo del formato, de la tonalidad y otros factores. Dentro del proyecto hemos digitalizado todo tipo de documentos y, por ejemplo, un documento con un formato DIN A2 en tonalidad de grises tiene un tiempo de captura de 2,5 segundos aproximadamente. En las mismas condiciones la toma a color se produce en 6 segundos.

También, dentro de la amplia gama de avances tecnológicos mencionados, posee una balanza portalibros motorizada lo que permite una manipulación del documento menor y más sencilla a la hora de ser digitalizado.



## Memoria de prácticas

- Canon DR-7090C:

El siguiente equipo de digitalización es el escáner Canon DR-7090C. Este escáner gracias a las dimensiones que posee es capaz de realizar digitalizaciones en formato DIN A3 en color.

Su uso es sencillo debido a que la composición del panel de control es asequible y de fácil manejo para cualquier usuario. Además de poder usarse a través del panel de control, tiene la posibilidad de manejarse desde el ordenador con un software especial para el escáner. Este programa se llama Capture Perfect 3.0.

Una vez digitalizada la documentación, ofrece en el soporte informático una buena calidad de imagen que varía desde los 100 PPP hasta los 600 PPP. Asimismo ofrece también una serie de procesamientos avanzados en la imagen.

Este escáner es muy útil en el proyecto debido a sus dos entradas para escanear. Se divide en una unidad de escaneado con pantalla fija en la parte inferior con dimensiones de hasta un DIN A3 y en la parte superior a ésta, hay una bandeja diseñada para escanear automáticamente por lotes con una capacidad de hasta 100 hojas y un tamaño variable según las dimensiones necesitadas.



- Hasselblad flextight x5:

Es un escáner de última generación con un sistema óptico vertical que permite colocar el sensor de reproducción hacia abajo para crear mayor contacto entre la lente y el original, lo que proporciona un mejor enfoque al material que se va a escanear, y con este posicionamiento permite la colocación de los diferentes soportes de una manera más sencilla.

Para la puesta en marcha del escáner, se debe abrir este sistema óptico vertical desde donde se enciende. Después, se escoge la imagen que se desee escanear y se adapta a uno de los soportes, mostrados en la parte inferior. Ese adaptador se encaja en la bandeja de luz desde donde se digitaliza la imagen y pasa al ordenador.

Al contrario que los otros dos escáneres, este último escáner no lo he utilizado durante las prácticas sino que solo he podido ver el procedimiento de actuación.



### 2.3. Tareas realizadas: el Archivo de la Autoridad Portuaria

Respecto a las tareas realizadas en la Unidad Técnica del proyecto *Patrimonio de Cantabria 3.0*, la función que desempeñé dentro del mismo, fue la de digitalizadora.

A lo largo de los tres meses en prácticas mi cometido ha sido digitalizar el Archivo Histórico de la Autoridad Portuaria junto con Lorena Sainz, digitalizadora perteneciente al proyecto.

Esta tarea había sido iniciada con anterioridad, por lo que ya se había extraído la documentación de las cajas y se había colocado en estanterías.

Los documentos vinieron en cajas numeradas, un total de 8, dentro de las cuales había diversas carpetas con su correspondiente signatura que pertenecían a diferentes proyectos.



9. Estanterías con la documentación en carpetas del Archivo de la Autoridad Portuaria

Además, la Autoridad Portuaria adjuntó un listado a la documentación con la información básica necesaria de las carpetas, donde aparecía la signatura, el nombre, las páginas, el número de planos y el número de pliegos de cada una. Esta lista también sirvió como orientación a la hora de estimar una fecha de recogida así como para seguir el patrón establecido.

Se colocaron en una estantería las diferentes cajas, enumeradas del 1 al 8, desde donde se han ido digitalizando por ese orden y han sido compartimentadas por carpetas en el ordenador. El modelo seguido para nombrar los archivos ha sido el siguiente, al inicio de todas las carpetas aparecen las iniciales AP, abreviación de Autoridad Portuaria, a continuación se coloca la numeración correspondiente con la caja y la nomenclatura. Por ejemplo, la caja 2 con el número 2571 sería AP\_CAJ\_02\_2571.

Dentro del Archivo de la Autoridad Portuaria, nos hemos encontrado con una documentación muy heterogénea, donde ha primado la documentación histórica de expedientes, proyectos de obra, entre otros asuntos, junto con planos históricos adjuntos a esa documentación.

Una vez ordenada y colocada la documentación, y tras su revisión y manipulación. Se procedió a su digitalización. Tarea que paso a comentar, pues participe en ella.

### 2.3.1. Digitalización

#### 2.3.1.1. Escaneado

Los escáneres que he utilizado durante el periodo de las prácticas han sido dos, el i2s Digibook Scanner Copybook y el Canon DR-7090C.

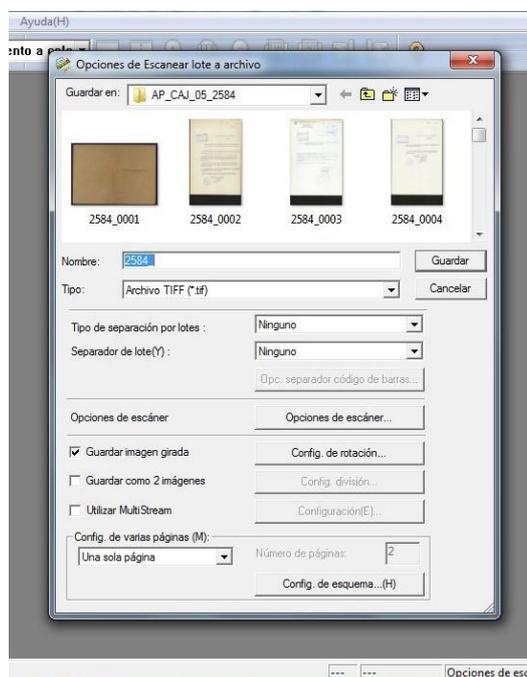
Primero, un breve apunte sobre el uso del escáner Canon DR-7090C. El escáner Canon fue utilizado durante la digitalización del Archivo Portuario con un fin específico. Ha servido para escanear lotes de documentos extensos con un formato DIN A4 o menor.

Mediante el software *Capture perfect 3.0* programé el escáner para poder ser digitalizado automáticamente. Una vez abierto el programa, seleccionamos en la pestaña de escáner la opción de escanear por lote a archivo, esta opción nos permite poder escanear múltiples hojas a la vez de forma automática.

## Memoria de prácticas

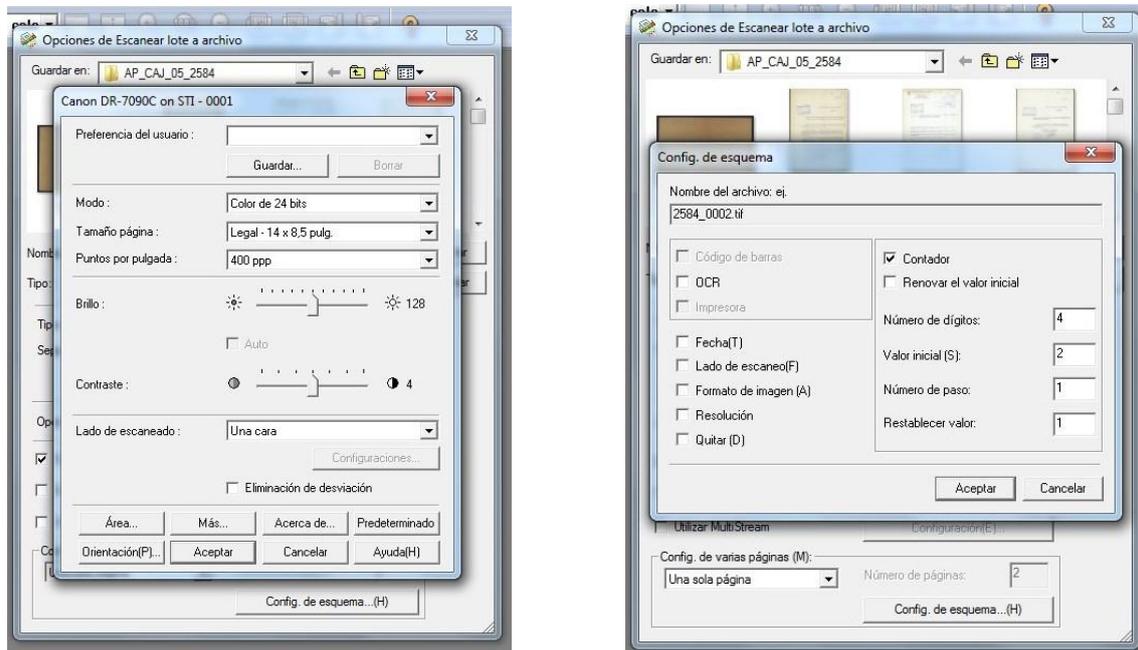
Una vez seleccionado esto, accedemos a una nueva ventana, donde escogeremos la carpeta en la cual irán guardadas las copias después le daremos un nombre que normalmente es la nomenclatura de la carpeta seguido de la numeración correspondiente a cada una de las digitalizaciones.

También le dimos el formato, que siempre es formato TIFF. Por último, configuré el esquema del documento a escanear.



10. Captura de pantalla del programa

A continuación, en opciones de escáner escogí la calidad de la imagen, el tamaño del lote de documentos a escanear, etc. Y después, antes de digitalizar se ha de elegir la configuración de esquema deseada para aplicar una numeración correlativa a los documentos dentro de la carpeta seleccionada.



11. Capturas de pantalla del programa

Una vez realizado esto, se coloca la documentación en la bandeja del escáner y sin haber cerrado el programa, se pulsa el comando guardar para que automáticamente se comience a escanear por lotes, es decir por grupos de documentos. Si en un momento determinado se quisiera parar el proceso de escaneado y cambiar la configuración, no habría problema ya que al introducir la documentación por lotes, el escáner capta cuando no hay material en la bandeja y da esa opción.

El uso del escáner i2s Digibook Scanner Copybook es más laborioso.

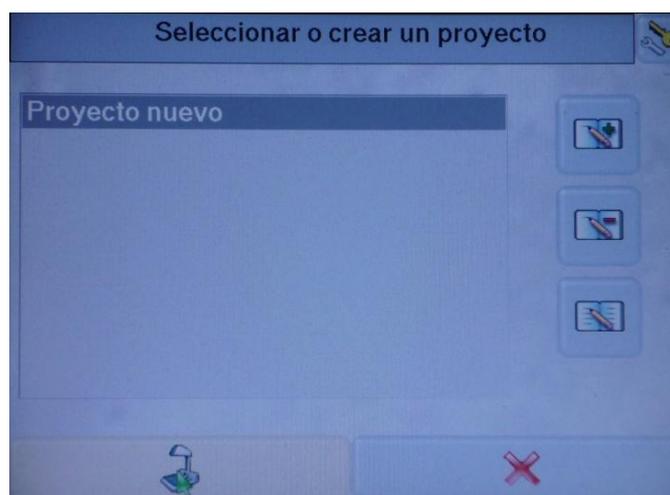
Antes de comenzar a trabajar con el escáner, hay que encender las lámparas para dejarlas calentar entre 20 y 25 minutos, con el fin de que tomen la temperatura de color adecuada. También debemos cerciorarnos de que las luces del entorno estén apagadas ya que el reflejo de los fluorescentes puede salir en la copia digital, dejando brillo en la digitalización. Aparte de que la temperatura de color de la fluorescencia del local difiere de la del escáner.

## Memoria de prácticas

Además, se debe limpiar el cristal con un trapo suave y un producto sin amoníaco especializado para preservar el material de una forma óptima. Este proceso solo se realiza en caso de que haya suciedad y sea necesario.

Una vez encendido el escáner, escogemos el idioma y tenemos dos opciones o escogemos un proyecto ya realizado donde ya estarían ajustadas todas las características o se crea un proyecto nuevo que se ajuste a las características de la documentación próxima a digitalizar.

En las prácticas, lo primero que se hizo fue escoger un proyecto ya realizado que se ajustará a las características de la documentación del Archivo Portuario, esto, dada la diversidad del material creó problemas por lo que se inició un proyecto nuevo.



12. Panel principal del escáner

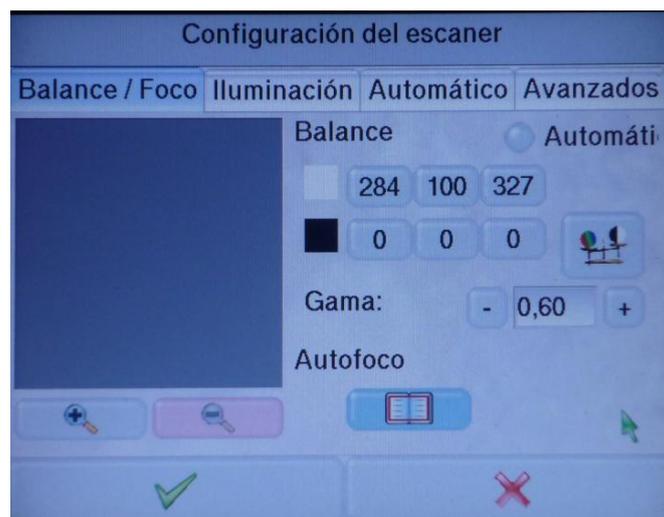
Al crear un proyecto nuevo, lo primero que se hace es darle una denominación y posteriormente se ajusta la iluminación. Para ajustar la iluminación del proyecto se pulsa en configuración de escáner y en la opción de Balance/foco encontramos este icono , que sirve para calibrar el color.

## Memoria de prácticas

Cogemos una cartulina blanca de trama rugosa proporcionada por el fabricante, la situamos sobre la superficie del escáner y procedemos a hacer la calibración. Después se coloca una cartulina negra, de la misma textura que la blanca, encima de la cartulina blanca haciendo reconocible los dos tonos y se realiza el balance de blancos y negros.

Este paso es muy importante porque puede condicionarnos todo el proceso. En las prácticas hubo un problema con la luz en los planos, unos aparecían muy claros y otros muy oscuros.

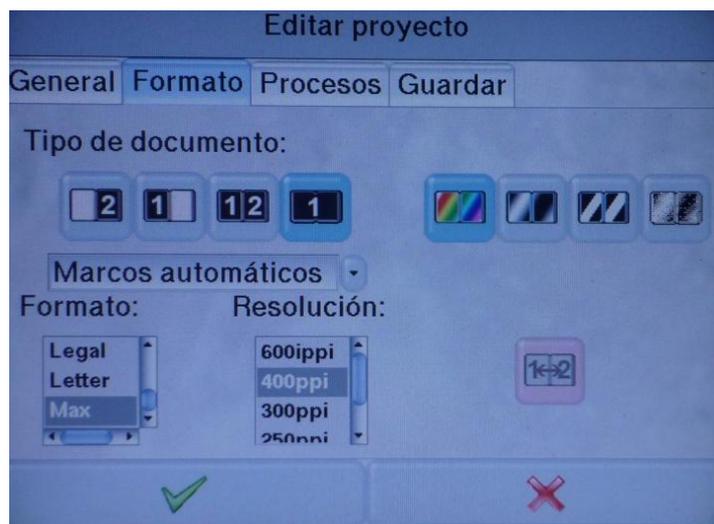
El problema era que se había desenfocado la iluminación. Se tuvieron que borrar todas las configuraciones creadas anteriormente y crear una nueva específicamente para el proyecto del Archivo Portuario.



13. Panel de configuración del escáner

Tras realizar el balance de blancos y negros se configura la iluminación. Utilicé la frecuencia dada por el fabricante, y la etiqueta se denomina igual que al proyecto creado. Después colocamos la cartulina blanca que hemos utilizado anteriormente, calibramos la iluminación y una vez realizado en corrección de la iluminación seleccionamos la opción que acabamos de establecer.

Una vez terminada la configuración, se comienza a escanear. Ya en la pantalla principal nos encontramos con una imagen a tiempo real de la plataforma de digitalización y en la parte inferior vemos dos botones, uno con la imagen de un libro que sirve como vista preliminar de la última digitalización realizada. El segundo botón representa un pequeño escáner en miniatura donde podremos editar el proyecto.



14. Panel para editar el proyecto

Dentro de este panel nos encontramos con varios apartados, el primero es el panel general desde donde podemos dotar de un nuevo nombre al proyecto.

En el menú formato accedemos a la estructura que se va a dar al documento a la hora de digitalizarse. Podemos escoger el tipo de documento, el primero solo guarda las páginas pares, el segundo las impares, el tercero los dos encuadres, y el cuarto lo guarda como una única página. En las prácticas los más utilizados han sido el tercero para los libros y para documentación suelta que se encontraba en una situación precaria, y el cuarto, sobre todo para los planos, aunque también para alguna hoja suelta con un tamaño superior al DIN A4 y algunas con los marcos de forma irregular.

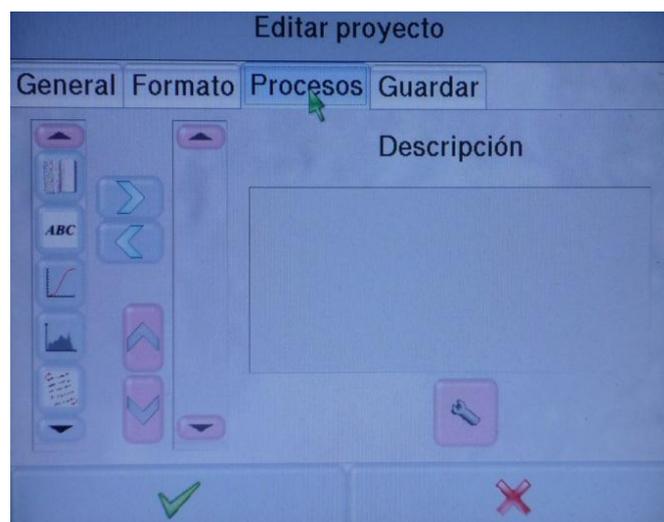
## Memoria de prácticas

Debajo, hay una pestaña donde se puede seleccionar el tipo de marcos. Los utilizados durante las prácticas han sido los marcos automáticos, ya que el programa del escáner reconoce el borde de los documentos. Otra opción es el procesado manual, utilizado para los planos y para la documentación que el marco automático no reconoce.

A la derecha, hay cuatro cuadrados donde se selecciona la gama cromática que se desea aportar al documento. Existen cuatro posibilidades: color, escala de grises, blanco y negro y por puntos. En las prácticas solo se ha usado a la hora de digitalizar el formato de color ya que al tratarse de documentos históricos se obtiene una imagen más fidedigna.

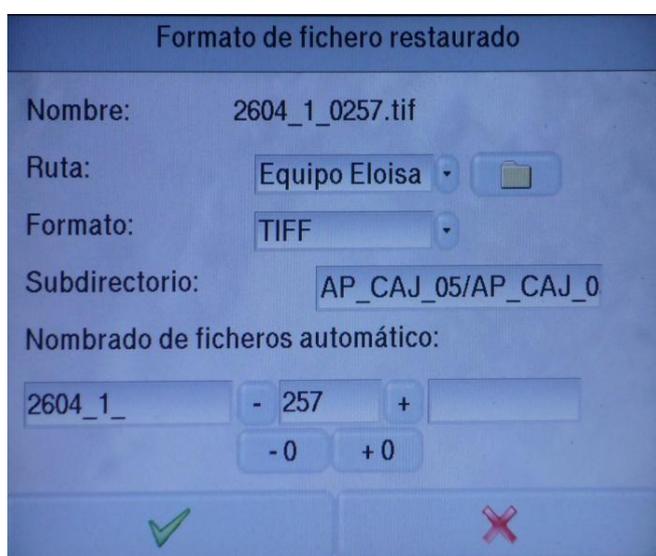
En cuanto al formato, el escáner ofrece variedad de tamaños aunque para el proyecto del Archivo Portuario siempre se ha utilizado el máximo dada la gran diversidad de documentos registrados. Lo mismo ocurre con la resolución, siempre se ha utilizado como configuración estándar 400 PPP, aunque el escáner ofrezca una amplia gama que va desde los 150 PPP hasta los 600 PPP.

El siguiente panel es la pestaña del menú de procesos que permite mejorar la calidad del documento.



15. Panel para editar el proyecto

En la última opción, la pestaña guardar, nos aparecen las diferentes formas en las que podemos archivar las imágenes que vamos a digitalizar. La opción que se utilizó durante el proyecto del Archivo Portuario fue la imagen restaurada. Dentro de ésta se escoge el equipo informático donde se enviarán las imágenes digitalizadas así como la carpeta correspondiente a cada grupo de documentos. Las imágenes se guardan en formato TIFF, para no perder calidad visual, y se enumeran. Delante y detrás de la enumeración, hay dos cuadros donde se puede añadir una nomenclatura.



16. Panel de enumeración y destino

Una vez editado el proyecto en el escáner, en el ordenador creamos un directorio para la posterior manipulación de la documentación.

En el disco C: / creamos una carpeta llamada escáner dentro de la cual se generan dos subcarpetas llamadas brutas y retocadas. Las brutas son aquellas digitalizaciones originales, que sin ningún tipo de manipulación, llegan desde el escáner y las retocadas son aquellas que siguen un procedimiento de cambio a través de un programa informático, que en este caso ha sido el programa Adobe Photoshop C5. Los archivos retocados son guardados con la misma denominación pero con la prolongación de copia.

## Memoria de prácticas

A la hora de digitalizar la documentación del Archivo Portuario, como ya se ha mencionado anteriormente, nos hemos encontrado con gran variedad de documentación de todo tipo: carpetillas, hojas sueltas, fragmentos de periódicos, fotografías, planos, etc.

En la práctica, la digitalización se ha realizado de tres maneras diferentes dependiendo del tipo de documento. La primera, cuando son documentos con marco regular -como son las carpetas, fotografías, etc.- se ha utilizado el marco automático. La segunda manera, para documentos con bordes irregulares, donde el marco automático no capta el perímetro del documento y debemos utilizar el manual, se ha procedido a manipularlos para amoldar los resultados a lo deseado que era que guardara correspondencia con el original.

Por último, para los planos, dependiendo del tamaño se optó por un escaneado completo o en fragmentos. Si hay que hacerlo en fragmentos, pondremos el marco manual para seleccionar el máximo espacio posible. Se comienza por el anverso creando puntos coincidentes para que después a la hora de montar el plano sea más fácil y si el reverso contiene información lo digitalizaremos igualmente.

Dado el peso de almacenamiento de tantas imágenes digitales, ha habido ciertas ocasiones en las que el escáner ha dado problemas. A veces en el proceso de pasar la imagen digital al ordenador se ha colapsado. Lo que origina una especie de línea negra en el documento. En estos casos, inevitablemente se ha de repetir la copia.

A parte del Archivo Portuario, durante mi estancia se trajeron a la Unidad Técnica 15 láminas de François Maréchal pertenecientes al fondo de la Biblioteca Municipal, porque debido a problemas técnicos de esta institución solicitaron que se llevara a cabo el trabajo en esta Unidad.



17. Proceso de digitalización

### 2.3.1.2. Tratamiento de las imágenes

Una vez escaneada la documentación se procede a tratar las imágenes a través del software de retoque fotográfico, Adobe Photoshop C5.

#### 2.3.1.2.1. Reconocimiento óptico previo

Antes de proceder a retocar el documento debemos hacer un reconocimiento óptico de los caracteres. Con el programa Adobe Photoshop C5, abrimos la imagen y con la herramienta de la lupa lo acercamos lo máximo que nos permitan los píxeles de la imagen, comprobamos que los caracteres sean legibles así como las imágenes estén definidas. En el caso de que no lo estén, debemos repetir la imagen.

Esto es un paso previo al reconocimiento óptico de caracteres u OCR (abreviación en inglés- Optical Character Recognition- utilizado más comúnmente). Se trata de un software que permite el reconocimiento óptico de los caracteres contenidos en una imagen de tal manera que estos se convierten en un texto reconocible para el ordenador que, por lo tanto, puede ser editable y puede trasladarse a programas de edición de textos como el Microsoft Word.<sup>13</sup>

### 2.3.1.2.2. Retoque informático

Una vez realizada la comprobación de la imagen se inician una serie de retoques con el fin de mejorarla.

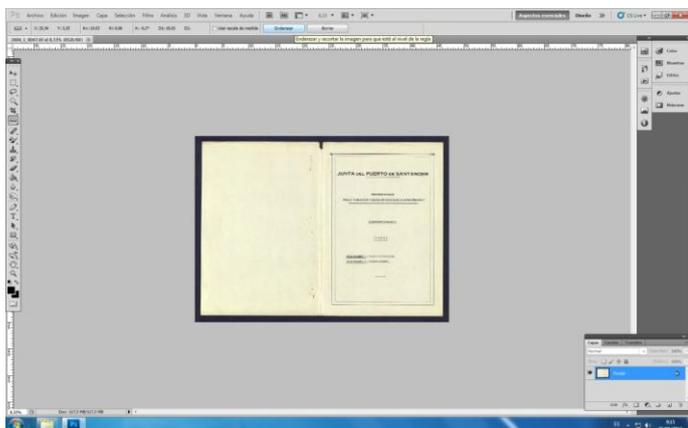
A la hora de proceder al tratamiento de la documentación, las acciones utilizadas con más frecuencia han sido enderezar, es decir, alinear y recortar para mejorar la visibilidad, contraste y definición de las imágenes. Estas tareas son realizadas de forma automática con el comando automatizar recortar y enderezar, aunque a veces el procedimiento no se hace de forma correcta por lo que se procede a una intervención manual.

Una manera rápida para realizar a la vez estas acciones en toda la documentación es hacerlo por lotes en los que unas determinadas manipulaciones se aplican a un conjunto de archivos o carpetas y subcarpetas. Es decir, se escoge la carpeta completa y se aplican de forma automática para, con posterioridad, repasar las imágenes por si se hubieran producido fallos o errores y, de este modo, poderlos enmendar.

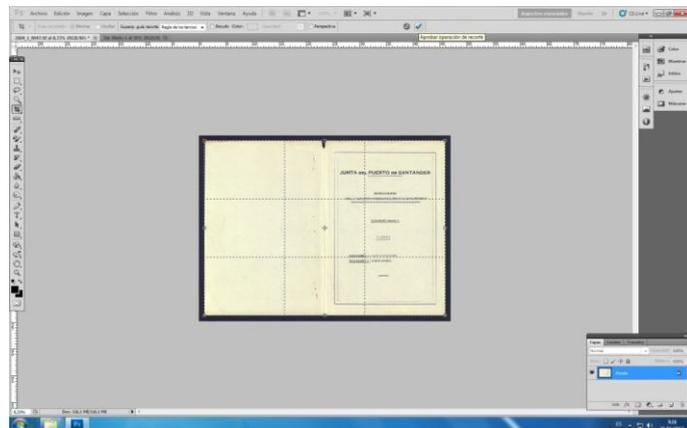
---

13 Kurturklik, espacio interactivo de la cultura vasca: <http://www.kulturklik.euskadi.net/wp-content/uploads/2011/04/pildora-OCR-2.pdf> (Consultado el 26/08/2013).

## Memoria de prácticas



18. Proceso de enderezar



19. Proceso de recortar

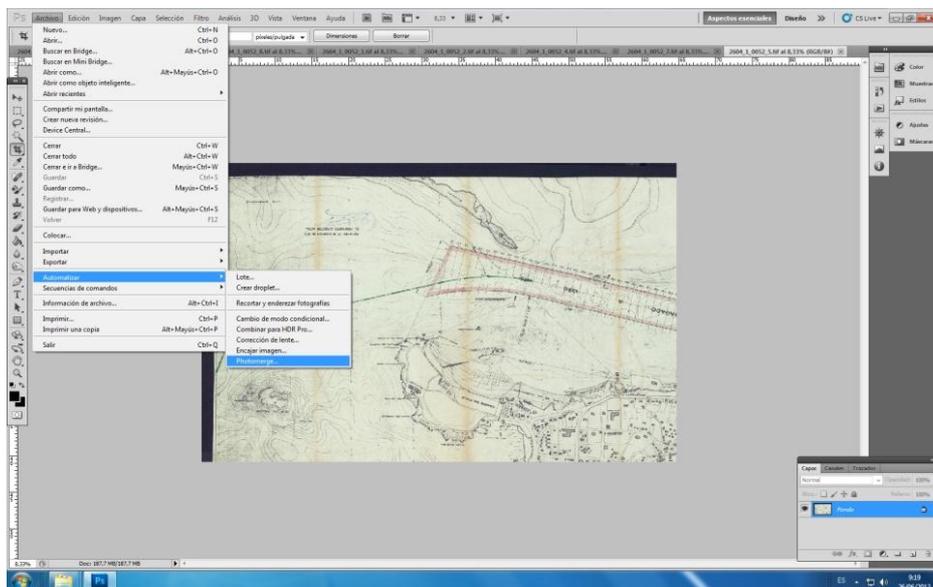
Como las operaciones que se aplican a las imágenes son frecuentes y previsibles, es muy conveniente crear acciones de Adobe Photoshop para poderlas aplicar automáticamente, después, a las imágenes escaneadas. Las acciones se generan pulsando el botón , situado a la derecha, con el cual se crean nuevas acciones. Se escoge la función que se quiera realizar, se le da un nombre y se selecciona la tecla que se quiere utilizar para que haga esa función. Por ejemplo, la tecla F12 va a servir para recortar y enderezar y cada vez que se pulse se realiza la función automáticamente.

Para que esto quede registrado, se debe dar a la tecla de grabar, se realiza la acción y una vez terminada esta, se pulsa “stop” quedando así registrada.

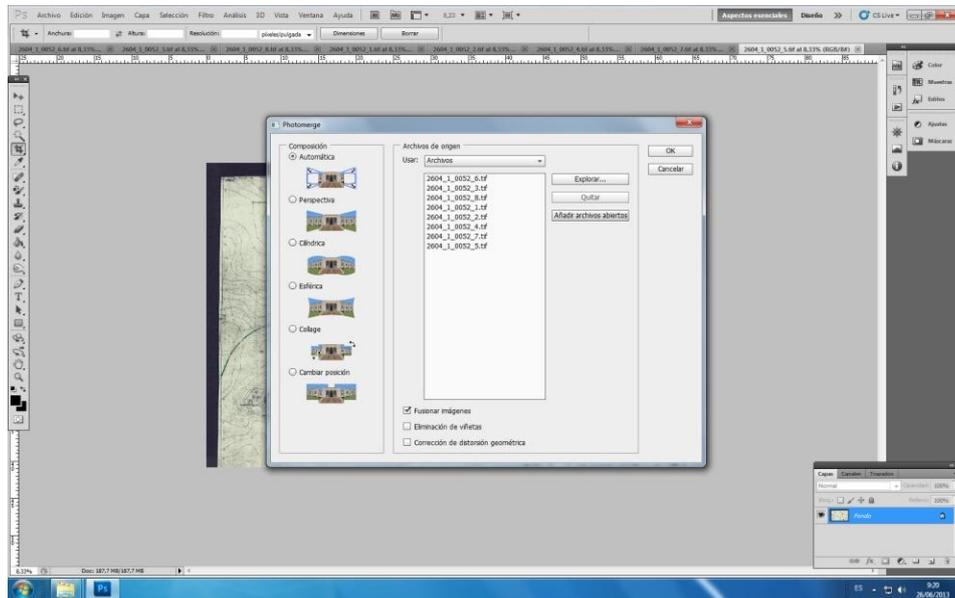
Para los planos, el tratamiento es distinto. Si son planos de dimensiones DIN A4 simplemente se enderezan y recortan de forma manual. En cambio si nos encontramos con un plano de un formato mayor debemos reconstruir el plano con las diferentes partes. Para ello se abren los diferentes fragmentos en los que se compone, y se procede a realizar el montaje, el cual puede ser hecho de dos maneras diferentes. Una primera, con el comando Photomerge, el cual automáticamente une los diversos puntos que coinciden en las distintas partes creando un panorama de salida. Una vez abierta la ventana del Photomerge hay que añadir los archivos que se quieren automatizar, de ahí la importancia de tener los fragmentos abiertos.

Si la panorámica de salida es correcta se acopla la imagen y se guarda, pero esto muchas veces no funciona por diversas razones, y es donde entra la segunda opción, la forma manual. Aquí se enderezarían y guardarían las diferentes partes manualmente, después se abre un nuevo documento en blanco con las dimensiones que nosotros queramos, si la capa base se nos queda pequeña no pasaría nada ya que Photoshop tiene la opción de ampliar lienzo.

Por último, unimos las partes encima de la capa fondo, dando a una de las partes una opacidad del 50 % aproximadamente para que la unión sea más sencilla de realizar y más exacta. Una vez coincidan, se seleccionan las diferentes capas exceptuando la capa fondo y se fusionan. Tras haber finalizado la fusión, corroboramos que ésta haya finalizado con éxito y acoplamos la imagen a la capa de fondo para crear una imagen homogénea.



20. Automatización de la unión de los fragmentos de un mapa



21. Panel del Photomerge

En este proceso, tras unir el plano, debemos recortar los bordes del mismo dejando un marco simétrico con el cual se pueda apreciar el perímetro original del plano.

La imagen, como ya se ha dicho anteriormente, está en formato TIFF. ¿Por qué el uso de este formato? He extraído de *Las recomendaciones para la digitalización de documentos en los archivos* de la Junta de Castilla y León<sup>14</sup> un fragmento que nos explica qué es un formato TIFF y que ventajas contiene.

“TIFF (Tagged Image File Format): Es un formato de fichero digital que acepta la *compresión* sin pérdidas, por lo que se conserva la información original completa. Se trata de un formato estándar y de uso libre, y es el más utilizado con fines de conservación. Proporciona gran calidad para imágenes en color (32 bits), escala de grises y blanco y negro. Su principal inconveniente es el alto volumen de almacenamiento requerido motivado por el tamaño del archivo que genera”.

14 Archivos de Castilla y León: [http://www.aefp.org.es/NS/Documentos/Guias-Manuales/JCYLRecomendaciones\\_Digitalizacion\\_Archivos2011.pdf](http://www.aefp.org.es/NS/Documentos/Guias-Manuales/JCYLRecomendaciones_Digitalizacion_Archivos2011.pdf) (Consultado el 28/08/13).

Es decir, que aunque el archivo resultante es voluminoso, se trata de un formato no propietario, de uso libre, que se ha generalizado y que, además, puede contener la información completa obtenida tras el escaneado, prácticamente un negativo digital.

Gracias al uso de este formato, obtenemos unas imágenes de alta calidad aunque de grandes dimensiones medidas en píxeles. Esto ha causado problemas en ocasiones, ya que el escáner se ha trabado y ha originado errores en Photoshop puesto que demanda una gran cantidad de la memoria RAM disponible en el sistema. Este problema de almacenamiento en la RAM lo hemos solucionado cambiando algunos parámetros del programa ya que consumía más memoria de la que nos permitía el ordenador. Hemos entrado en edición, preferencias, generales y en el rendimiento hemos cambiado el tamaño de azulejos de caché y hemos tenido que bajar el rendimiento de otras propiedades.

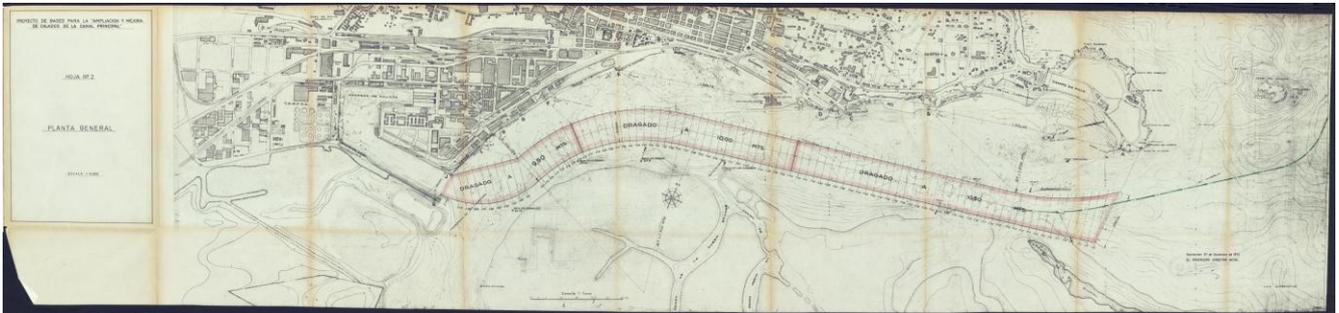
La imagen obtenida en formato TIFF se pasa a formato JPG ya que, aunque se pierda calidad en la imagen, el tamaño es menor, algo fundamental a la hora de montar los PDF finales. Las imágenes normales se automatizan a una resolución de 150 PPP y a un tamaño de 1.080 del lado más largo, puesto que el lado menor se calcula automáticamente en el programa Photoshop para mantener la relación de escala. El otro lado es calibrado automáticamente por Photoshop.

Para los planos se optó por la misma resolución de 150 PPP. Las medidas dependen del tamaño del plano escaneado. Si ocupan un tamaño de 3000 a 5000 píxeles, se reducen a 3000 píxeles en el lado mayor. Si son imágenes que contienen de 5001 a 7000 píxeles, se reducen a 3500 píxeles en el lado máximo. Y, por último, las imágenes mayores de 7000 píxeles –hasta las de 10.000 o mayores- se limitan los JPG finales a 4000 píxeles de lado.

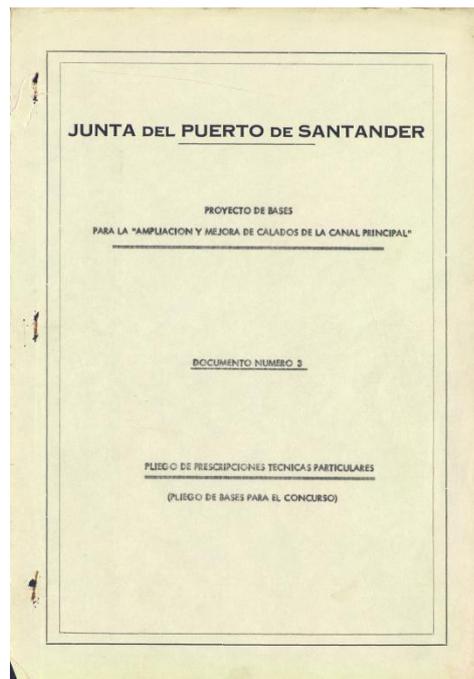
Una vez terminada manipulada de la documentación se almacena en un disco duro externo para liberar espacio en el disco duro del ordenador, así como para poder ser compartido por todos los equipos de la Unidad. De este proceso se responsabiliza Sara Melendi, encargada de los procesos posteriores.

## Memoria de prácticas

A continuación ofrezco un ejemplo final del escaneado de un plano grande y de la portada de una cuartilla de la documentación del Archivo Portuario.



22. Ejemplo de plano



23. Ejemplo de portada de una cuartilla de la documentación del Archivo Portuario

Aparte del programa Photoshop, la Unidad cuenta con otros programas para el tratamiento de la documentación como son *Capture Perfect 3.0*, mencionado anteriormente para la utilización del escáner Canon.

Otro software utilizado es *Total Commander*, un gestor de archivos con mayor número de funciones que el que ofrece el administrador de archivos que contiene el sistema Microsoft Windows por defecto. Con este programa, una vez manejada la documentación, se procede a cambiar la numeración y el nombre de los archivos en lotes automatizados. Con la herramienta multinombrado se ha podido cambiar las denominaciones sin tener que ir una a una.

### 2.3.2. Generación de ficheros de servicio

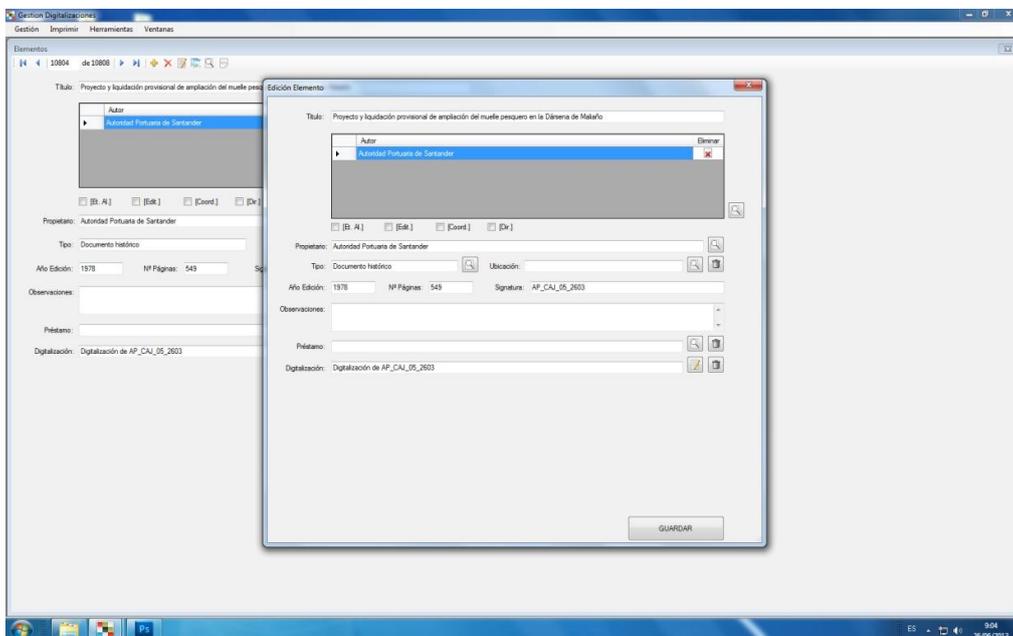
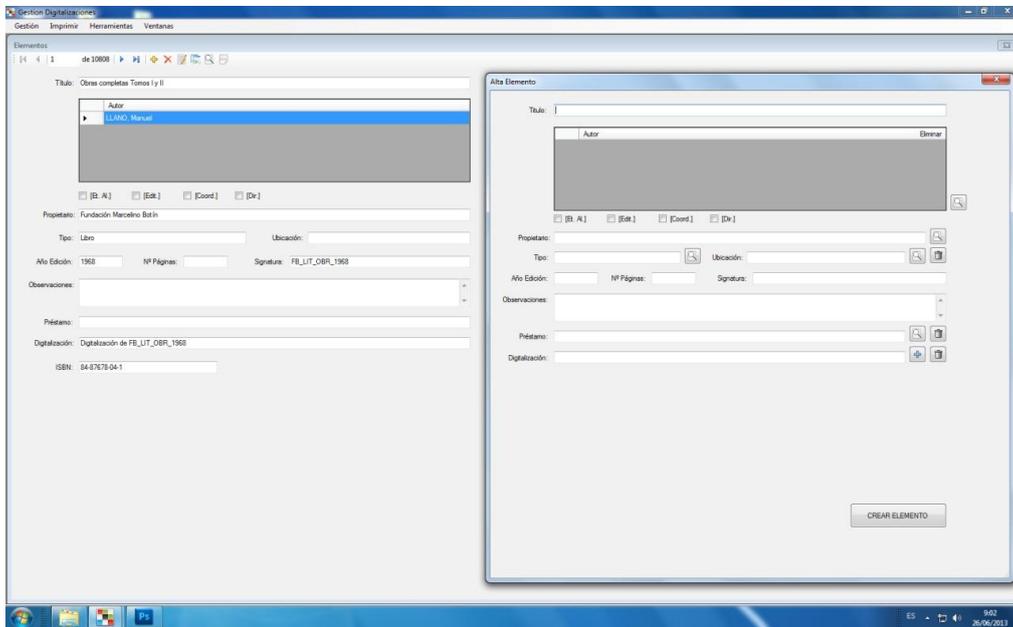
Cada vez que se termina una carpeta se registra en una base de datos compuesta por unas fichas modelos creadas para el proyecto.

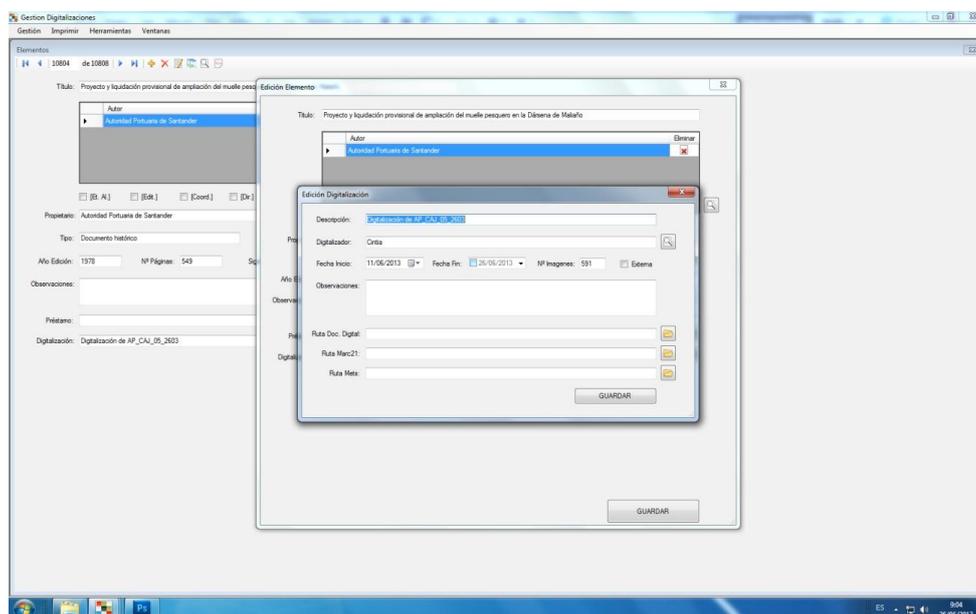
Para esto, se rellenan los datos más elementales de la documentación como son el título, el autor, el propietario, el tipo de documento, la ubicación, el año de edición, el número de páginas, la signatura. Además se añaden observaciones, datos de digitalización y una descripción que suele contener la ruta de almacenamiento, el digitalizador, la fecha de inicio y fin, el número de digitalizaciones, y las observaciones.

Es una ficha sencilla, donde se rellenan los datos básicos para mantener una base de datos actualizada de lo que se va digitalizando durante el proyecto. Nada tiene que ver con los métodos usados en otros lugares como por ejemplo el método utilizado en la Red Digital de Colecciones de Museos de España (CERES), donde se realizan fichas histórico-artísticas en las que se incluyen una serie de datos y metadatos más extensos, como el número de inventario, la descripción, las dimensiones, etc. Son fichas con un contenido más exhaustivo.

# Memoria de prácticas

Ejemplo de fichas realizadas durante el proyecto:





### 24. Capturas de pantalla del gestor de ficheros

Ejemplos de títulos de la documentación digitalizada del Puerto:

- Pliego de bases para la contratación de colaboración para el estudio del Hinterland del Puerto y sus zonas de servicio.
- Estudio del tráfico actual y futuras perspectivas del P.S., Pedro de Aguilar Martínez de la Vega.
- Desdoblamiento de calzada y obra de fábrica especial paso sobre el FFCC de acceso al Puerto de Raos nº635 de Santander a Francia por S.Sebastián y P.K-4/6889 6,706 año 1970.
- Estudio del hinterland y zonas de servicio del puerto de Santander. Autor: Senda 3. Año 1977.

### **3. PROBLEMAS Y POTENCIALIDADES**

Cantabria es una región que demandaba un proyecto de esta índole debido al rico Patrimonio que posee. Este proyecto supone una gran innovación en lo que concierne al campo cultural en la región, ya que nunca se había realizado nada de estas características en el entorno geográfico.

Además de su carácter innovador, el proyecto cuenta con una serie de factores que lo enriquecen, como la avanzada tecnología que posee, los especialistas altamente cualificados en diferentes materias, además de contar con el apoyo de numerosas entidades/instituciones de carácter tanto público como privado que han acogido el proyecto y han colaborado cediendo numerosa documentación.

Es de gran importancia destacar la continuación de este proyecto ya que en el contexto de crisis económica mundial que nos encontramos, es muy difícil que este tipo de proyectos de investigación e inversión continúen y se culminen con éxito. El proyecto también se ha visto afectado por la crisis, aunque se trata del único proyecto que sigue vigente dentro de todas las líneas de investigación llevadas a cabo por la Fundación Botín.

Sin embargo, nos encontramos con el impedimento de que el proyecto cuenta con limitados recursos de personal para la ejecución del proceso de digitalización y manipulación de los materiales, y por consiguiente de la correcta elaboración del mismo.

Tenemos que resaltar que las infraestructuras de la Unidad Técnica donde se desarrolla el proyecto, tanto la situación como las dimensiones, son idóneas. Como ya se ha mencionado anteriormente, el primer piso está destinado al proceso de digitalización así como al recibimiento de los representantes de las entidades colaboradoras. En cambio la parte correspondiente al sótano, con un amplio espacio de 142 metros cuadrados, se encuentra actualmente sin un uso determinado.

Este espacio sería adecuado para la realización de actividades relacionadas con el proyecto, como charlas, conferencias, actividades culturales, etc. o mismamente como sala para la exposición pública de parte de la documentación que se está digitalizando. A modo de ejemplo y con documentación manejada durante mis prácticas sobre el Archivo Portuario, se podrían mostrar algunos de los planos que considero de gran interés histórico ya que son valorados como auténticas joyas del archivo.

### 4. CONCLUSIONES

A modo de reflexión sobre las prácticas realizadas en la Unidad Técnica del proyecto *Patrimonio de Cantabria: web semántica 3.0*, quiero destacar la gran trascendencia que tienen los trabajos de investigación de esta magnitud como forma de preservar el Patrimonio identitario de la Comunidad Autónoma. Un Patrimonio que no es ajeno al paso del tiempo, que sufre cambios, y que debemos recoger, conservar y difundir.

Es de gran interés que aún se emprendan estudios sobre este tema, ya que son una forma de avanzar y de seguir descubriendo el pasado. Un pasado que nos identifica y que nos ha hecho evolucionar hasta lo que somos hoy en día.

El valor que posee la web semántica se debe a la agrupación de los contenidos que realiza y dirige a diversos usuarios, desde el investigador hasta el usuario común que demanda conocimiento sobre el patrimonio.

La web semántica integra la información más variada y facilita el estudio a futuros investigadores de una forma ordenada y clara, ya que reúne dentro de un mismo portal todo el conocimiento existente. Por ello, si se indaga sobre un asunto, con una simple consulta en la web se obtendría la información deseada.

Los modos de consulta han ido evolucionando con el paso del tiempo, desde los papiros de la Biblioteca de Alejandría, pasando por la consulta de libros en las grandes bibliotecas europeas, y por último, la red. Las páginas web son la forma de consulta más utilizada en los últimos años debido a su comodidad, la posibilidad de consultar un documento original sin tener que desplazarte hasta su lugar de origen y por la rapidez que proporciona dicha información. La web semántica ofrece modernidad y supone un avance adecuado a los tiempos presentes y todo esto con formas nuevas y dinámicas.

Es un proyecto novedoso, en el que hay que continuar trabajando para poder seguir dando la importancia que se merece a la cultura porque como dijo el filósofo Epicteto, *sólo el hombre culto es libre*.

## 5. BIBLIOGRAFÍA Y RECURSOS WEB

### BIBLIOGRAFIA

- KEFEER, A. y GALLART, N., *La preservación de recursos digitales. El reto para las bibliotecas del siglo XXI*, Barcelona, Editorial UOC, 2007.
- VIVES, J. (Coord.), *Digitalización del patrimonio: archivos, bibliotecas y museos en la red*, Barcelona, Editorial UOC, 2009.
- VVAA, *Directrices para proyectos de digitalización de colecciones y fondos de dominio público, en particular para aquellos custodiados en bibliotecas y archivos*, España, Ministerio de Cultura, 2002.
- TRAMULLAS J. (Coord), *Tendencias en documentación digital*, Gijón, Ediciones Trea S.L., 2006.

### INFORMES

- *Ontología Patrimonio de Cantabria. Unidad Técnica: Informe de actividades (2010-2012)*. (Consultado el 25/06/2013)
- CAPELLÁN, G., *Web semántica del Patrimonio de Cantabria. Informe de Septiembre de 2010* (Consultado el 25/06/2013)

RECURSOS WEB

- Archivos de Castilla y León:

[http://www.aefp.org.es/NS/Documentos/Guias-Manuales/JCYLRecomendaciones\\_Digitalizacion\\_Archivos2011.pdf](http://www.aefp.org.es/NS/Documentos/Guias-Manuales/JCYLRecomendaciones_Digitalizacion_Archivos2011.pdf) (Consultado el 28/08/2013).

- Cantabria Campus Internacional:  
<http://www.cantabriacampusinternacional.com/ceiOfrece/Paginas/Que-es-CCI.aspx> (Consultado el 13/08/2013).

<http://www.cantabriacampusinternacional.com/areasExcelencia/cipl/Paginas/Presentacion-Patrimonio-y-la-Lengua.aspx> (Consultado el 25/06/2013).

- CIDOC:

<http://www.cidoc-crm.org/> (Consultado el 12/08/2013).

- Diccionario de informática:

<http://www.alegsa.com.ar/Dic/digitalizar.php> (Consultado el 27/06/2013).

- Fundación Botín:

<http://www.fundacionbotin.org/la-fundacion-mision.htm> (Consultado el 22/08/2013).

[http://www.fundacionbotin.org/creada-la-unidad-de-trabajo-para-ampliar-y-mantener-los-contenidos-de-la-web-de-cantabria-de-la-fundacion-botin\\_publicacion-fundacion-botin-12037422.htm](http://www.fundacionbotin.org/creada-la-unidad-de-trabajo-para-ampliar-y-mantener-los-contenidos-de-la-web-de-cantabria-de-la-fundacion-botin_publicacion-fundacion-botin-12037422.htm) (Consultado el 25/06/2013)

- HERNÁNDEZ CARRASCAL, F., “Intercambio de información sobre patrimonio cultural: la Ontología del Patrimonio de Cantabria”, en *XI Jornadas de Gestión de la Información: Servicios polivalentes, confluencia entre profesionales de archivo, biblioteca y documentación*, Madrid, 2009 (Consultado el 26/06/2013). Disponible en: <http://eprints.rclis.org/13938/1/XIJGI-Hernandez.pdf>.
- HERNÁNDEZ CARRASCAL, F., “Ontología del Patrimonio de Cantabria”, en: *EL profesional de la información*, 2008, v.17, nº 1, pp. 92-98. (Consultado el 26/06/2013). Disponible en: <http://www.elprofesionaldelainformacion.com/contenidos/2008/enero/11.pdf>.
- Kurturklik, espacio interactivo de la cultura vasca:  
<http://www.kulturklik.euskadi.net/wp-content/uploads/2011/04/pildora-OCR-2.pdf>  
(Consultado el 26/08/13).
- *Manual de instrucciones del escáner CANON DR-7090C*:  
[https://www.canonresellers.com/content/scanners/user\\_manuals/DR7090CUserManual.pdf](https://www.canonresellers.com/content/scanners/user_manuals/DR7090CUserManual.pdf) (Consultado el 24/05/2013).
- *Manual de instrucciones de Hasselblad Flextight x5*:  
<http://instrucciones-pdf.com/dref/2936633> (Consultado el 26/08/2013).
- *Manual de instrucciones del escáner i2S Copybook*:  
<https://biblioteca.ucm.es/foa/intranet/doc8985.pdf> (Consultado el 24/05/2013).
- Universidad de Cantabria:  
[http://www.unican.es/WebUC/Internet/Informacion\\_General/resena.htm](http://www.unican.es/WebUC/Internet/Informacion_General/resena.htm)  
(Consultado el 12/08/2013).