



## **EXPERIENCIA DE CREACIÓN DE PORTAL TEMÁTICO SOBRE PRENSA DE LA SEGUNDA GUERRA MUNDIAL.**

### **EXPERIENCE CREATING PRESS THEMATIC PORTAL ON WWII.**

#### **Autor:**

Juan Manuel Dévora Lorenzo. Universidad Rey Juan Carlos. Servicio de Aplicaciones Corporativas. [juanmanuel.devora@urjc.es](mailto:juanmanuel.devora@urjc.es)

#### **Resumen:**

El Archivo Fernández-Xesta está compuesto por una vasta colección de prensa editada durante la Segunda Guerra Mundial. Mediante convenio de colaboración con la Universidad Rey Juan Carlos, se está realizando la digitalización completa del archivo y la creación de un portal donde podrá consultarse toda la información, que actualmente asciende a 1.100.000 páginas.

#### **Abstract:**

Fernández-Xesta Archive consists of a vast collection of newspapers and journals published during the Second World War. By an agreement of collaboration with the Universidad Rey Juan Carlos, is being made complete file scanning and creating a portal where all information will be available, which currently amounts to 1,100,000 pages.

#### **Palabras Clave:**

Prensa, Segunda Guerra Mundial, digitalización de fondos.

#### **Keywords:**

Press, WWII, digitized collections, image scanning.

El Archivo Fernández-Xesta está centrado en la prensa editada durante la Segunda Guerra Mundial (1939-1945) y comprende tanto periódicos y revistas de todos los contendientes y de los principales países neutrales, como pasquines que por su diseño, su título común y estar numerados y fechados, pueden ser considerados como prensa, aunque sea de propaganda. Se trata, muy posiblemente, de la mayor colección del mundo sobre el periodo en manos particulares. Este material comenzó



a ser reunido a mediados de los años cincuenta y en la actualidad comprende alrededor de 36.000 ejemplares (30.000 revistas, 5.000 periódicos y unos 1.000 pasquines). El total estimado de páginas sobrepasa largamente el millón. Aparte de los ejemplares de prensa, el Archivo cuenta con un apartado de miscelánea que incluye todo tipo de objetos y publicaciones de la época, como carteles anunciadores de las revistas, fajas de reparto de periódicos, hojas de suscripción, distintivos de los correspondientes de guerra, proyectiles para lanzar los pasquines sobre las trincheras enemigas, etc. Entre las piezas más singulares del archivo se pueden citar las siguientes:

- Ejemplares miniaturizados de la revista *Signal* lanzados con bombas volantes V1 y V2 sobre Londres.
- El periódico confeccionado en ciclostil por los defensores de Tobruk durante el cerco de ocho meses establecido por el Afrika Korps de Rommel en torno a esa ciudad.
- Las revistas *Time* y *Newsweek* reducidas y reproducidas sobre papel fotográfico, tras ser lanzados los negativos a las guarniciones norteamericanas aisladas en remotas islas del Pacífico (Navy V Mail).
- El ejemplar de *Volkischer Beobachter* falsificado por el MI5 británico para hacer creer a Rudolf Hess, que su familia estaba siendo perseguida en Alemania, tras haberse dirigido este a Gran Bretaña.
- Cuatro de los siete únicos números del periódico *Panzerbär*, creado por los defensores de Berlín en pleno cerco de la ciudad, que culminó con la muerte de Hitler y la caída de la capital del Reich.
- El ejemplar del leaflet del SHAEF, el Cuartel General Supremo Aliado, que, en cuatro idiomas, da cuenta oficial de la rendición de Alemania.
- Ejemplares del periódico *Pravda* realizados y puestos en circulación por los alemanes en las zonas ocupadas de la Unión Soviética, conocido entre los expertos como “el Pravda de Hitler”.



- Más de medio centenar de cabeceras distintas de periódicos del frente confeccionados por la Wehrmacht: Desde el publicado en el lejano norte (*Polar Kurier*) a los desiertos de Africa (*Der Oase*) o en el punto más extremo alcanzado por el ejército alemán en toda la guerra (*Kämpfer im Kaukasus*).
- Podemos citar además numerosas publicaciones restringidas y de circulación prohibida editadas por ambos bandos; números especiales para los combatientes, en tamaño miniatura, de periódicos (como *The Times* de Londres o el *Boston Globe*) o revistas (como la americana *Time* o *The New Yorker*). Y el caso de la única prensa británica que apoyó al Tercer Reich: *The Evening Post* de Jersey o *The Star* y el *Evening Press* de Guernsey, en las islas anglonormandas (las islas del Canal).
- Todo ello junto con otras decenas de miles de copias que incluyen hasta casi una treintena de ediciones distintas de la revista militar americana *Yank*, confeccionada y editada simultáneamente en numerosas ciudades de los cinco continentes. O el periódico *Star and Stripes*, también editado en varias decenas de puntos a lo largo de todo el planeta. Y las ediciones en una quincena de lenguas de la revista británica *Neptune* o la veintena que en diferentes idiomas editó la alemana *Signal*, de la que el archivo dispone de casi millar y medio de ejemplares distintos.

Muchos de tales ejemplares han debido de vivir auténticas odiseas antes de integrarse en el Archivo, ya que editados bajo el régimen nazi, han pasado medio siglo en los escondites más insospechados, en los países del Este, durante la época en la que Europa estaba dividida por el Telón de Acero y la posesión de una de estas publicaciones podía suponer largas condenas e incluso en algunos momentos, posiblemente la muerte. Porque, al margen de la cantidad o de la calidad de algunas piezas, cada uno de los ejemplares que se encuentra en el Archivo Fernández-Xesta, tiene una larga y a veces trágica historia. Es un superviviente no sólo a las terribles vicisitudes bélicas vividas por el mundo en aquellos años o políticas de la etapa posterior, sino, que incluso debió desafiar el carácter de efímero que tiene una

revista o un periódico recién leído y que únicamente recupera o acrecienta su auténtico valor con el paso de los años, cuando adquiere una categoría histórica.. Y otras muchas agresiones, desde las grapas que se oxidan, a la tinta que se desvanece o el papel, cuarteado por la sequedad o deshecho por un ambiente excesivamente húmedo. Por no citar las condiciones en las que ha estado guardado hasta el momento de su recuperación. Por ello, cada ejemplar es único, con una historia singular.

Por medio de un convenio de colaboración firmado entre la Fundación de la Universidad Rey Juan Carlos y el Archivo Fernández-Sexta, se materializa un proyecto de creación de un portal temático que incluye la digitalización, clasificación, caracterización de la información y publicación de dicha obra. El resultado del proyecto, es un portal que incluye el más del millón de páginas digitalizadas, consultables a texto completo e indexado por todo su contenido, con un volumen total de 3TB de información.

### **El portal WWIIPress.**

Para la consulta del fondo se está desarrollando un portal web, que se espera publicar a lo largo de 2013, con el nombre de WWIIPress, multi-idioma y de acceso libre, donde se puede navegar por la colección (en función de la temática, del países de edición de la revista...), buscar en el fondo a texto completo con multitud de filtros (idioma de la revista, temática, fechas de publicación...), consultar las biografías de los principales personajes de la Guerra, etc.

Además de la consulta completa de la colección, hay otras secciones destacadas:

- **Imágenes:** conjunto de imágenes relacionadas.
- **Videos y documentales:** enlaces a vídeos y documentales relevantes.
- **Día a día:** sección con un breve resumen de los hechos acaecidos en dicho día y un enlace al periódico de la colección del día siguiente donde se relatan los hechos.

El portal asimismo trata de fomentar la investigación en los contenidos y recogerá investigaciones y publicaciones relacionadas con el fondo.

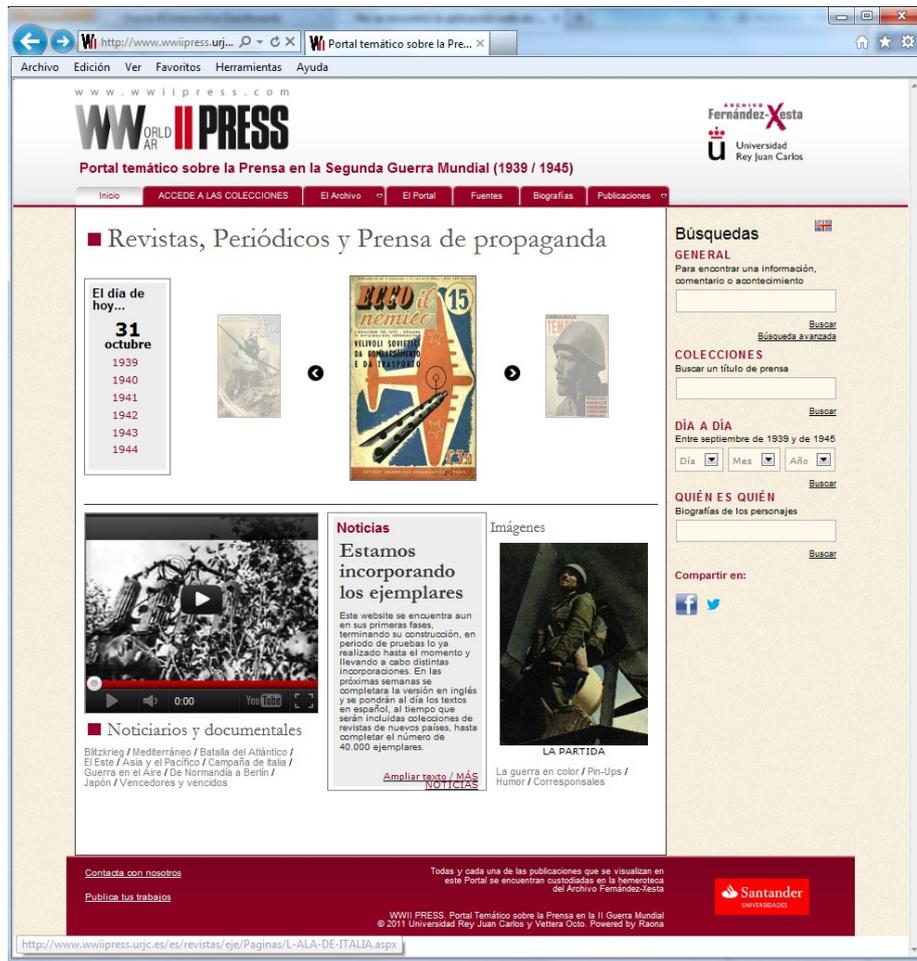


Ilustración 1. Portada del portal WWIIPress.

## Digitalización del Archivo.

Dada la antigüedad y delicadeza de los elementos del Archivo, la digitalización se ha realizado físicamente en la sede del mismo. Se ha utilizado un escáner con formato máximo de digitalización DIN A2, a una resolución de 300 puntos por pulgada y 24 bits de profundidad de color, y un cabezal de cámara 3x35MPixels.

El proceso de digitalización incluía dos fases, uno en la sede del Archivo y otra de postprocesado:

- En la sede del Archivo:
  1. Manipulación de revistas: Las revistas son extraídas de sus correspondientes fundas por la documentalista, que a la vez es la

encargada de introducir los siguientes metadatos asociados en una hoja de cálculo para su posterior envío junto con la digitalización:

- Medición/Recuento de páginas.
- País al que pertenece y condición (eje, aliados o neutrales), idioma, años publicación, periodicidad (semanal, quincenal, etc.), edición, editor, distribución, color (blanco y negro, color, etc.), dimensiones, tamaño de la revista y número de páginas, fecha, época, volumen y número.

## 2. Proceso de digitalización:

- Digitalización de revistas formato JPEG a doble página.
- Revisión de las imágenes digitalizadas para detectar errores.

### - Postprocesado posterior:

- Proceso de corte de imágenes, para suprimir bordes negros página a página.
- Revisión del proceso anterior (corte) para constatar la fiabilidad del proceso de corte de imágenes.
- Se preparan las portadas de las revistas para introducirlas en carpetas independientes, a una resolución menor de ancho 500 píxeles.
- Se generan los PDF optimizados, a la máxima calidad.
- Generación de proceso de OCR a los ficheros PDF.
- Generación de proceso de inserción de marca de agua a los ficheros PDF.
- Indexación de los PDF.

Todo este proceso, dado el volumen del archivo, ha sido una tarea ardua que se ha ido realizando a lo largo de 6 años.

## Clasificación y estructura de la información.

El Archivo está dividido en tres colecciones principales:

- Revistas.
- Periódicos.
- Prensa de propaganda (Air-leaflets).

www.wiipress.com

# WW II PRESS

Portal temático sobre la Prensa en la Segunda Guerra Mundial (1939 / 1945)

Inicio | ACCEDA A LAS COLECCIONES | El Archivo | El Portal | Fuentes | Biografías | Publicaciones

Inicio > Colecciones

## ■ Colecciones

[Revistas](#)  
[Periódicos](#)  
[Prensa de Propaganda \(Air-leaflets\)](#)

El Archivo Fernández-Xesta consta de secciones diferenciadas de Revistas, de Periódicos y de Prensa en pasquines (air-leaflets) que, a su vez, se organizan por países y dentro de cada país por temas concretos. Elija que sección del Archivo desea consultar y una vez en ella, seleccione el país y el tema correspondiente.

**Búsquedas**

**GENERAL**  
Para encontrar una información, comentario o acontecimiento

[Buscar](#)  
[Búsqueda avanzada](#)

**COLECCIONES**  
Buscar un título de prensa

[Buscar](#)

**DÍA A DÍA**  
Entre septiembre de 1939 y de 1945

Día  Mes  Año

[Buscar](#)

**QUIÉN ES QUIÉN**  
Biografías de los personajes

[Buscar](#)

Compartir en:

[f](#) [t](#)

[Contacta con nosotros](#) | [Publica tus trabajos](#)

Todas y cada una de las publicaciones que se visualizan en este Portal se encuentran custodiadas en la hemeroteca del Archivo Fernández-Xesta

WWII PRESS, Portal Temático sobre la Prensa en la II Guerra Mundial © 2011 Universidad Rey Juan Carlos y Vettera Octo. Powered by Raona

### Ilustración 2.Colecciones.

A medida que se recibe la digitalizan de los elementos, se realiza un proceso de revisión e introducción de metadatos y contenidos en el portal, consistente en:

- Revisión del pdf para comprobar la calidad del mismo y detectar errores en la digitalización (páginas invertidas, páginas que pudieran faltar, etc.).

- Por cada título de la colección:
  - Se introduce una ficha que incluye los metadatos principales del elemento (Título, Editor, Origen, Idioma, País de distribución, Periodicidad, Formato, Paginación, Color de la portada, Color interior, Publicidad).
  - Se genera el logotipo de la revista y se sube al portal.



Ilustración 3. Ficha de revista.

- Se incluyen en el portal los ejemplares de cada título. Para cada uno de ellos se introducen la imagen de la portada de la revista y los siguientes metadatos:
  - Fecha de publicación.
  - Volumen.
  - Número.
  - Observaciones.



Ilustración 4. Portada de una revista.

- Se compila el pdf de la revista para publicarla en una formato específico para su lectura (ver siguiente sección), de tal forma que cuando el usuario haga clic sobre la portada del ejemplar se abrirá la revista (periódico o pasquín) para su lectura.

www.wiipress.com

# WWORLD II PRESS

Portal temático sobre la Prensa en la Segunda Guerra Mundial (1939 / 1945)

Inicio | ACCEDA A LAS COLECCIONES | El Archivo | El Portal | Fuentes | Biografías | Publicaciones

Inicio > Revistas > Reino Unido General > FLIGHT

## FLIGHT

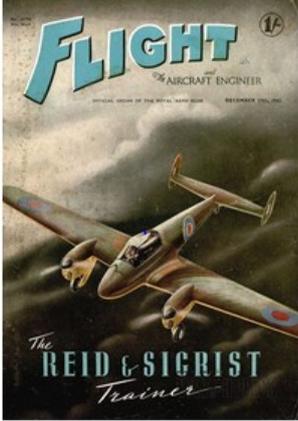
13 ejemplares

**Ficha**

Cada número de *Flight* abría sus páginas redaccionales con un largo comentario sin firma bajo el título de *The Outlook* (Las Perspectivas) que pretendía hacer un análisis de los principales acontecimientos de la semana y una suerte de pronóstico sobre lo días venideros. Otra sección habitual era *Here and There* (Aquí y allá) con pequeñas noticias de muy diversa índole, siempre relacionadas con el mundo aéreo. Y con el estallido de la guerra, ésta empezó a ocupar un lugar preeminente entre los temas a tratar por *Flight*, bajo la cabecera *War in the Air*

[Leer más...](#)

**Años / Volúmenes**



Revistas relacionadas

- Aeroplane
- Aeroplane
- Flying

### Búsquedas

**GENERAL**  
Para encontrar una información, comentario o acontecimiento

[Buscar](#)  
[Búsqueda avanzada](#)

**COLECCIONES**  
Buscar un título de prensa

[Buscar](#)

**DÍA A DÍA**  
Entre septiembre de 1939 y de 1945

Día:  Mes:  Año:

[Buscar](#)

**QUIÉN ES QUIÉN**  
Biografías de los personajes

[Buscar](#)

Compartir en:

[f](#) [t](#)



Cover	Issue	Volume	Number
Alperts & Their Equipment	2 de febrero 1939	Volumen XXXV	Número: 1571
Military Aircraft Number	14 de marzo 1940	Volumen XXXVII	Número: 1629
Electric Steels for Aero Engines	5 de septiembre 1940	Volumen XXXVIII	Número: 1654
FRIVRE	12 de diciembre 1940	Volumen XXXVIII	Número: 1668

Ilustración 5. Ejemplares de revista.

### Generación de los visualizadores.

Para la visualización de las revistas en el portal, no se muestra el PDF de la misma sino que se ha generado un visor (o versión *folleto* de la misma) para ofrecer al lector una experiencia más realista.



Ilustración 6. Visor de un ejemplar.

La generación del visor para cada revista ha sido un proceso laborioso dado el volumen de la información. Para la producción de todos y cada uno de los ejemplares ha sido necesario un software específico que produce el la versión *folleto*.

Mediante dicho software, y cinco máquinas virtuales corriendo en un servidor durante 24h a lo largo de cuatro meses, se fueron procesando uno a uno cada pdf correspondiente a un ejemplar de la colección y obteniendo como resultado el formato publicable.

Como dato curioso, destacar que el software no permitía la automatización desatendida de la tarea (hay software que sí lo hacen pero por cuestiones de funcionalidad y coste no se tomaron en cuenta). Para solucionar este problema se simuló un *usuario virtual* mediante un software específico (Autolt), se realizó una simulación de los movimientos de ratón y clics correspondientes en cada opción de la aplicación generadora de los folletos, que permitía la entrada de un pdf y generaba la salida del folleto. Este *usuario virtual* repetía la tarea mientras hubiera ficheros a procesar. De esta manera se consiguió automatizar de manera



desatendida la producción en masa de lo que, de manera contraria, hubiera sido una producción uno a uno.

Debido a que las tareas de introducción de portadas y títulos en el portal y la producción de los visualizadores correspondientes se realizaron en paralelo, hubo que solucionar problemas derivados de la asociación del visualizador con su portada correspondiente del portal, de tal manera que quedara perfectamente enlazado. Como solución, se estandarizó un nombrado de carpetas, de tal manera que posteriormente se pudiera realizar de manera automática la vinculación entre el elemento y su ubicación en el sistema de archivos. No obstante, a pesar de dicha estandarización y dado el volumen de la colección, quedaban numerosos elementos huérfanos; como solución se realizaron procedimientos automáticos para buscar, detectar y corregir problemas de normalización en el nombrado.

Con esta normalización en el nombrado queda vinculado el *folleto* con el elemento correspondiente del portal, de manera que no ha sido necesario en el portal vincular de manera expresa la portada de revista con su versión *folleto*, ya que al nombrarse de manera idéntica se ha configurado de manera general.

### **Colaboraciones.**

El desarrollo del proyecto ha sido posible gracias a una red de colaboradores que, de manera desinteresada, han sintetizado biografías, han traducido y revisado textos, han elaborado un resumen con los hechos más destacados de cada día de la guerra... y un largo etcétera de cuestiones relacionadas.