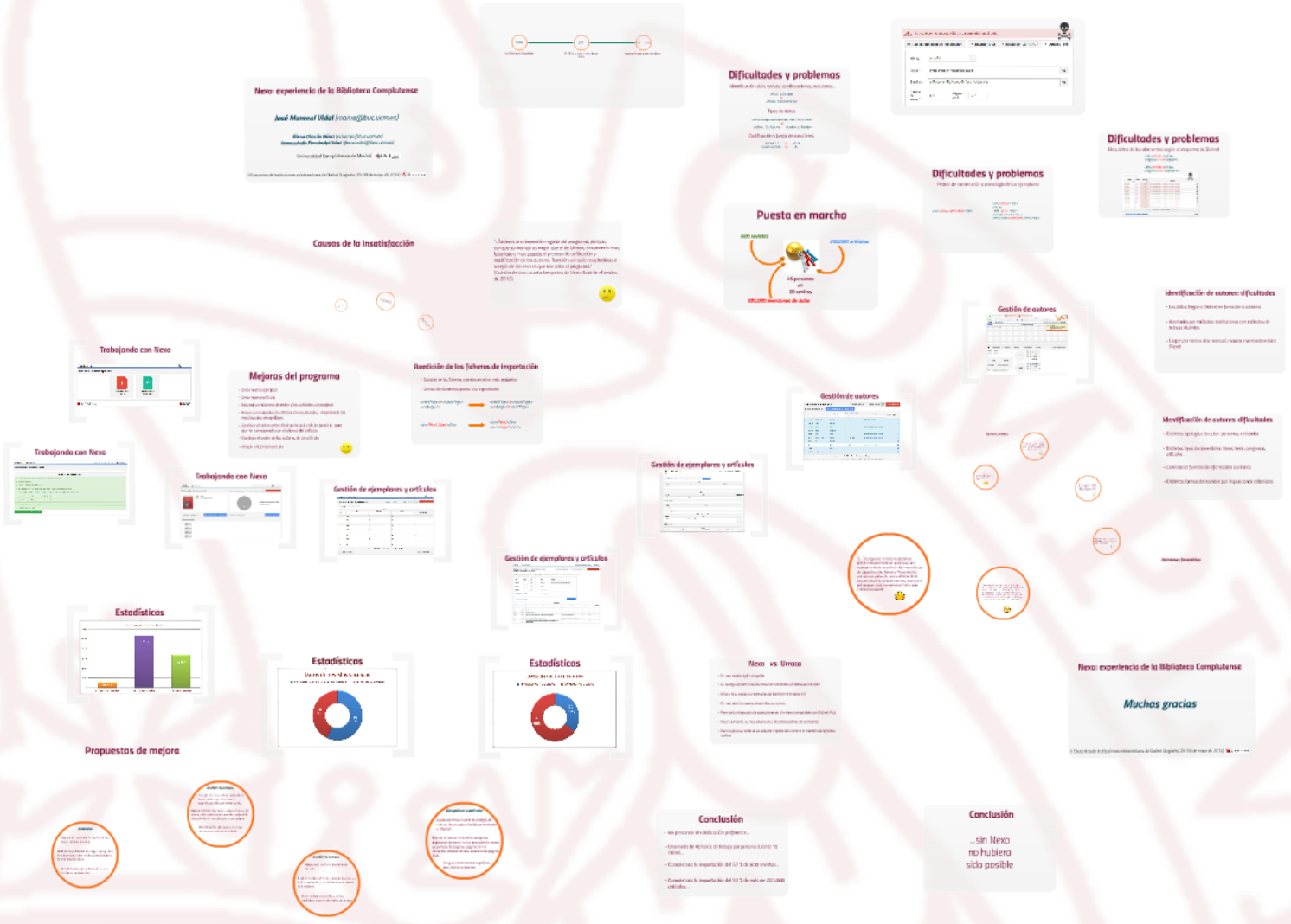
Nexo: experiencia de la Biblioteca Complutense



Nexo: experiencia de la Biblioteca Complutense

José Monreal Vidal *(monre@buc.ucm.es)*

Elena Chacón Pérez *(echacon@buc.ucm.es)*
Inmaculada Fernández Sáez *(ifernandez@buc.ucm.es)*

Universidad Complutense de Madrid

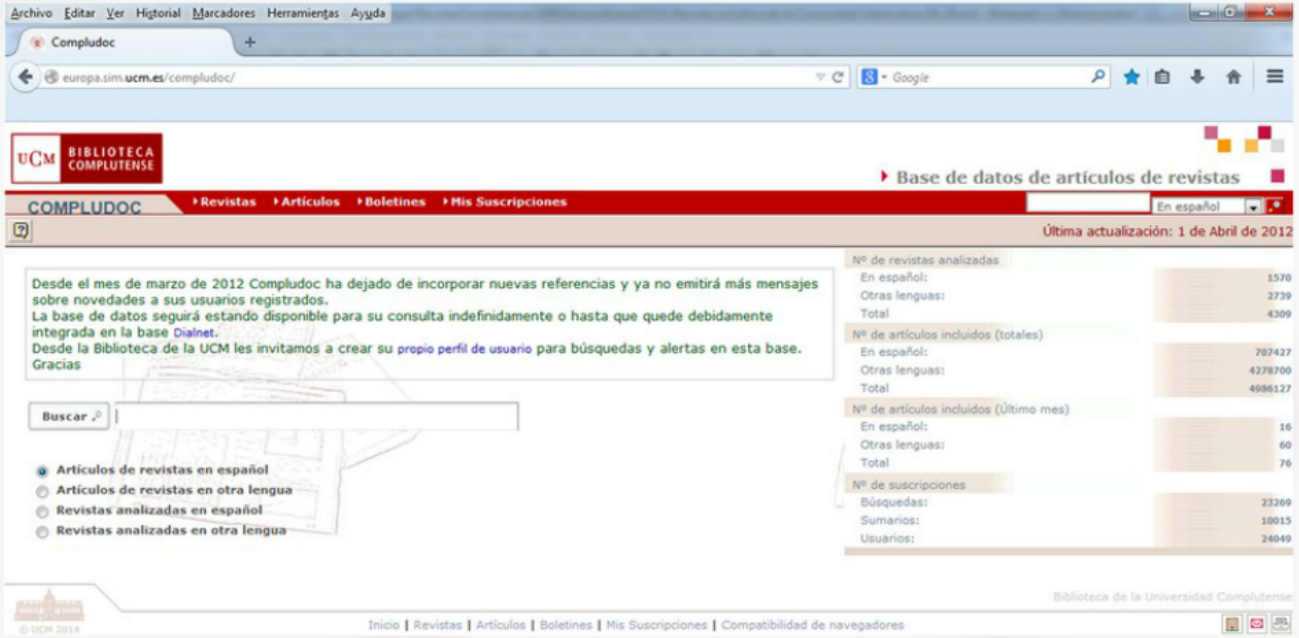
II Encuentro de Instituciones colaboradoras de Dialnet (Logroño, 29-30 de mayo de 2014)

Prezi

1998

En 1998 nace Compludoc

Prezi



- Trece años de vida:
  - 4.310 revistas
  - 4.975.602 artículos
- Vaciado de revistas en español
- Swets proporciona sumarios de revistas en otros idiomas

Prezi

2011

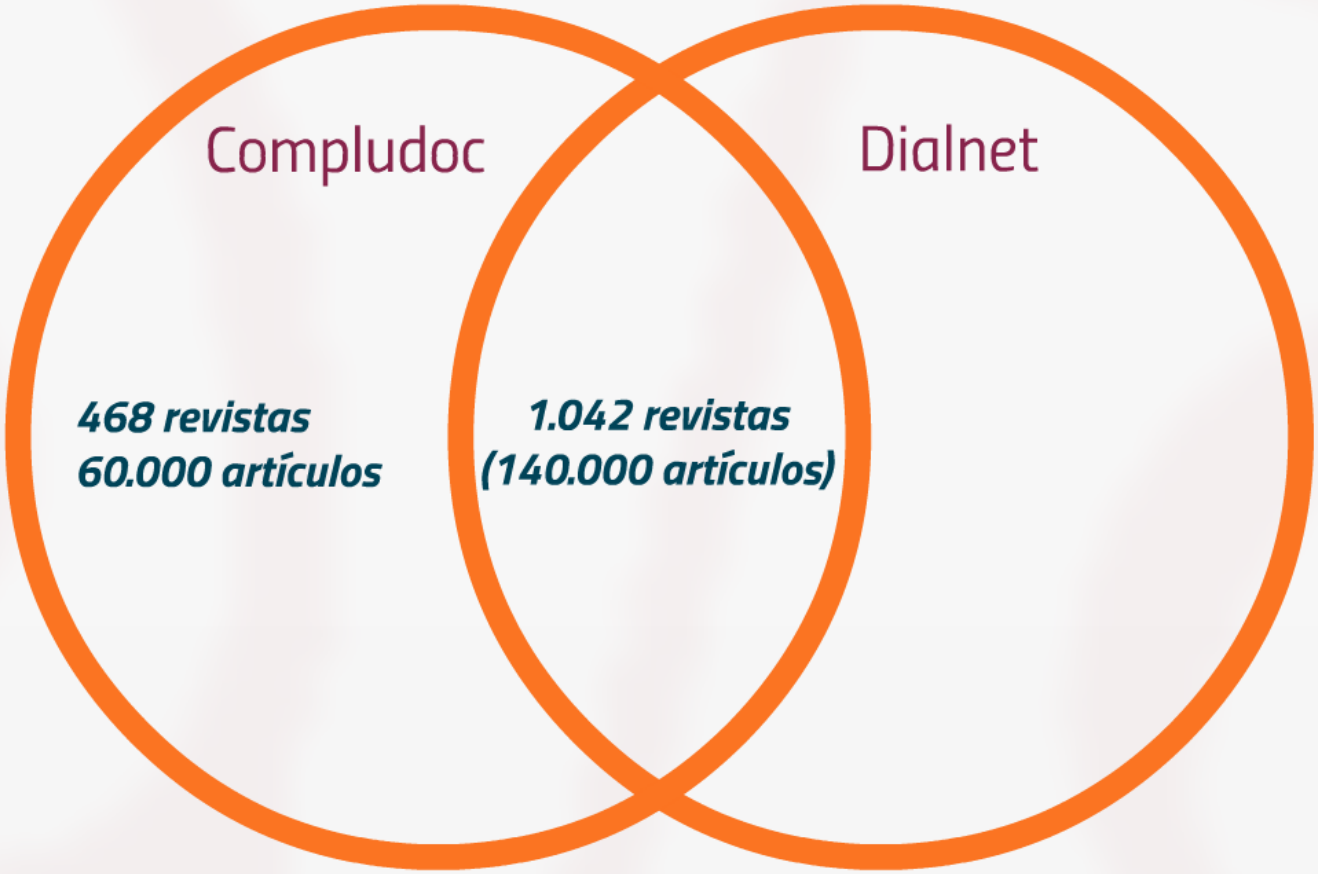
En 2011 se firma el acuerdo con Dialnet

Prezi

Cruzado solo de los títulos vaciados por la BUC

- 1.510 revistas
- 699.676 artículos

Resultado del cruzado Compludoc-Dialnet



Prezi

2011-2013

Proceso de conversión de datos

Prezi

Conversión de los sumarios al formato XML 'DIALNET_JOURNAL'

```
<title>Odiseo, 1998 NOV 9; Vol. 10 (1406)</title>
...
```

```
<td class="smt">Esbozo de gramática sentimental /
<a title="...">Rubio Montaner, P.</a></td>
<td class="smp">5-7</td>
```

```
<dialnetId>6666</dialnetId>
<journal>
<title>Odiseo</title>
<issue>
<year>1998</year>
<volume>10</volume>
<number>1406</number>
<description>NOV 9</description>

<article lang="ES">
<title><![CDATA[Arco Agoniza: crónica]]></title>
<startPage>5</startPage>
<endPage>7</endPage>
<author>
<origSignature>Rubio Montaner, P.</origSignature>
<name>P.</name>
<sn>Rubio Montaner</sn>
</author>
</article>
```

Prezi

Dificultades y problemas

Identificación de la revista: continuaciones, escisiones...

Odiseo, ISSN papel

vs.

Odiseus, ISSN electrónico

Tipos de datos

<title>Trabajos de estadística, 1961; XII (II)</title>

vs.

<volume>12</volume> <number>2</number>

Codificación y juego de caracteres

ISO Latin 1 *vs.* *UTF-8*

A₂ *vs.* A_2

Prezi

GRAVE **Artículo con etiquetas HTML no permitidas en el título**

Página principal de importación › Revista 19406 › Ejemplar - 2011, Nº. 4 › Artículo -658

Idioma * español

Título * Proteomics in Venom Research

Subtítulo a Focus on PLA₂ Molecules

Página de inicio * 629

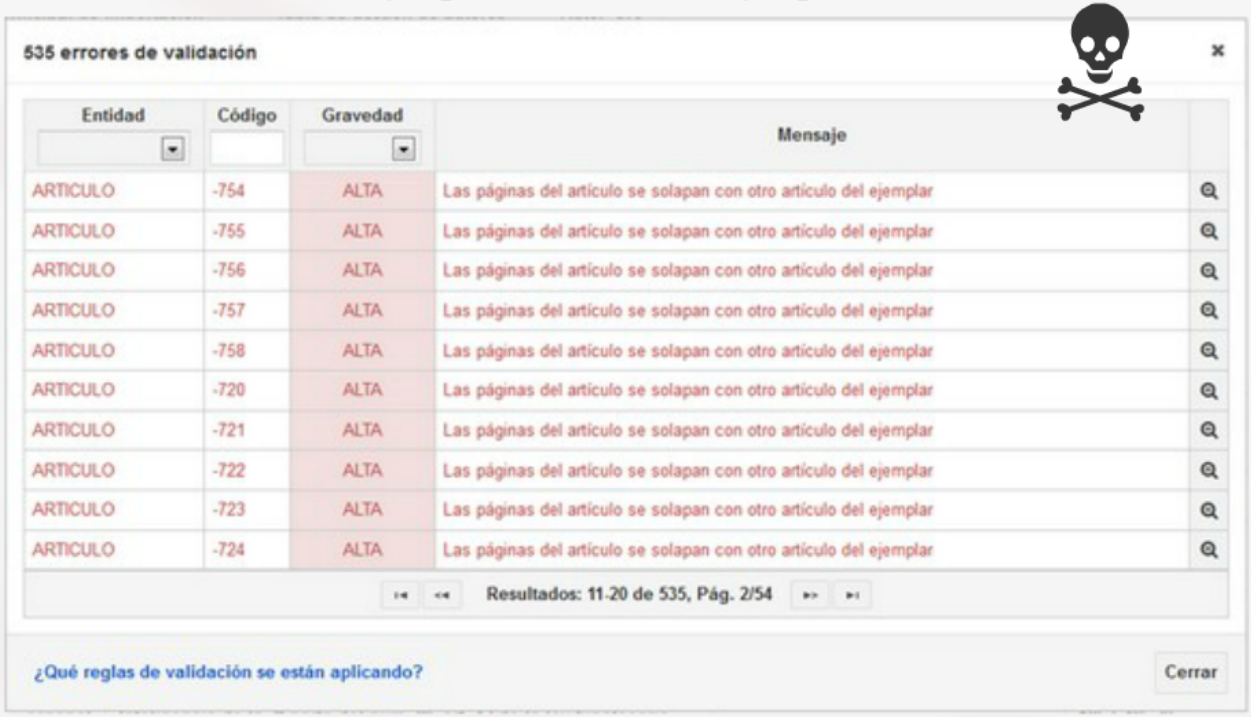
Página de fin 637

Prezi

Dificultades y problemas

Requisitos de los elementos según el esquema de Dialnet

<title>Artículo 1</title>
<paginas>1-10</paginas>
...
<title>Artículo 2</title>
<paginas>10-12</paginas>



Prezi

Dificultades y problemas

Patrón de numeración y cronología de los ejemplares

```
<title>Odiseo, 2011; NOV</title>
```

```
<title>Odiseo</title>
<issue>
<year>2011</year>
<number>11</number>
<description>noviembre</description>
```

Prezi

Puesta en marcha

600 revistas

200.000 artículos

46 personas
en
30 centros

285.000 menciones de autor

Prezi

"...Tuvimos una impresión regular del programa, porque, aunque su manejo es mejor que el de Urraca, nos parecía muy laborioso y muy pesado el proceso de unificación y modificación de los autores. También se hacía muy tedioso el arreglo de los errores que marcaba el programa."
(Opinión de una usuaria temprana de Nexo durante el verano de 2013)

Prezi

Causas de la insatisfacción

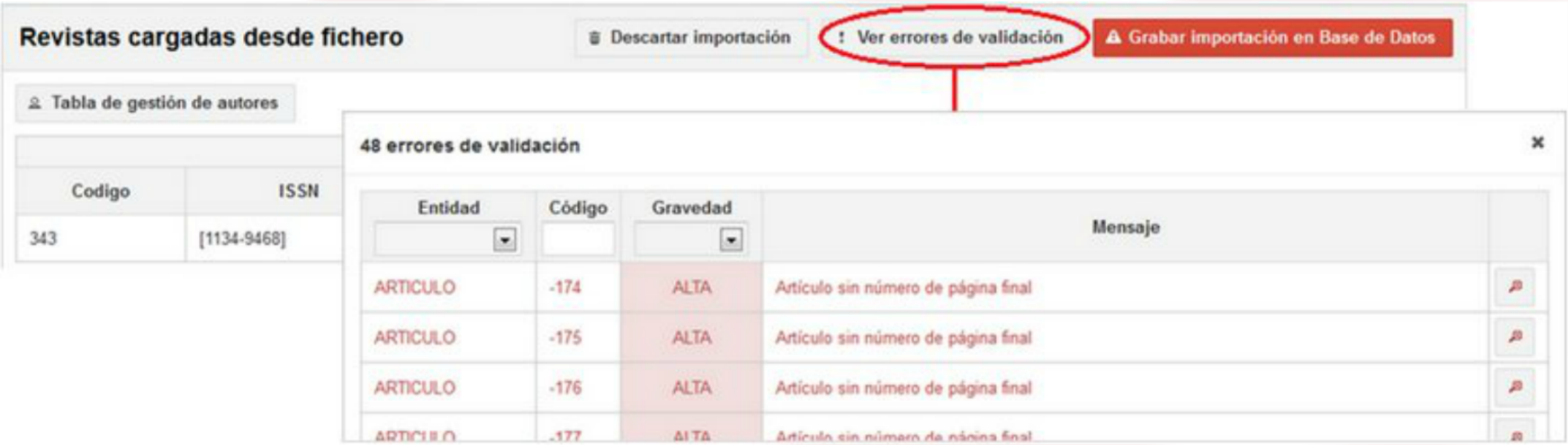
Prezi

Ficheros de gran tamaño

Prezi

Numerosos errores de validación de artículos

Prezi


Revistas cargadas desde fichero
Descartar importación
Ver errores de validación
Grabar importación en Base de Datos
Tabla de gestión de autores
Codigo
ISSN
343
[1134-9468]
48 errores de validación
Entidad
Código
Gravedad
Mensaje
ARTICULO
-174
ALTA
Artículo sin número de página final
ARTICULO
-175
ALTA
Artículo sin número de página final
ARTICULO
-176
ALTA
Artículo sin número de página final

Numerosos errores de validación de autores

Prezi

Editar autor

Nombre: * Antonio

Apellido1: Baraybar Fernández

Apellido2:

Sexo: Hombre

Tipo: **Persona** Entidad

Particula:

Formas de firma
Baraybar Fernández, Antonio; Universidad Rey Juan Carlos

Cancelar

Confirmar cambios

Prezi

Reedición de los ficheros de importación

- División de los ficheros grandes en otros más pequeños
- Corrección de errores previa a la importación

<startPage>5</startPage>
<endpage />

<startPage>5</startPage>
<endpage>5</endPage>

<sn>Pérez López</sn>

<sn>Pérez</sn>
<sn2>López</sn2>

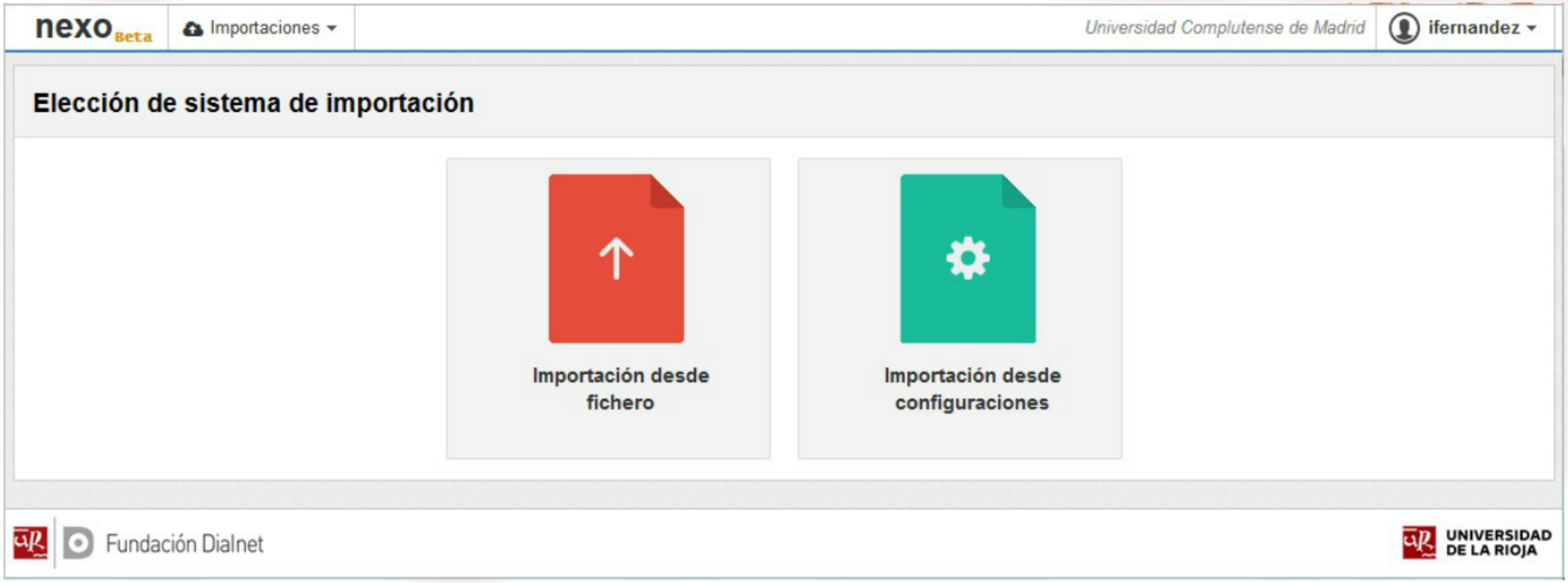
Prezi

Mejoras del programa

- Crear nuevo ejemplar
- Crear nuevo artículo
- Asignar un número de orden a los artículos sin paginar
- Pasar a minúsculas los títulos en mayúsculas, respetando las mayúsculas ortográficas
- Cambiar el orden entre título principal y título paralelo, para que se corresponda con el idioma del artículo
- Cambiar el orden de los autores de un artículo
- Añadir el DOI del artículo

Prezi

Trabajando con Nexo

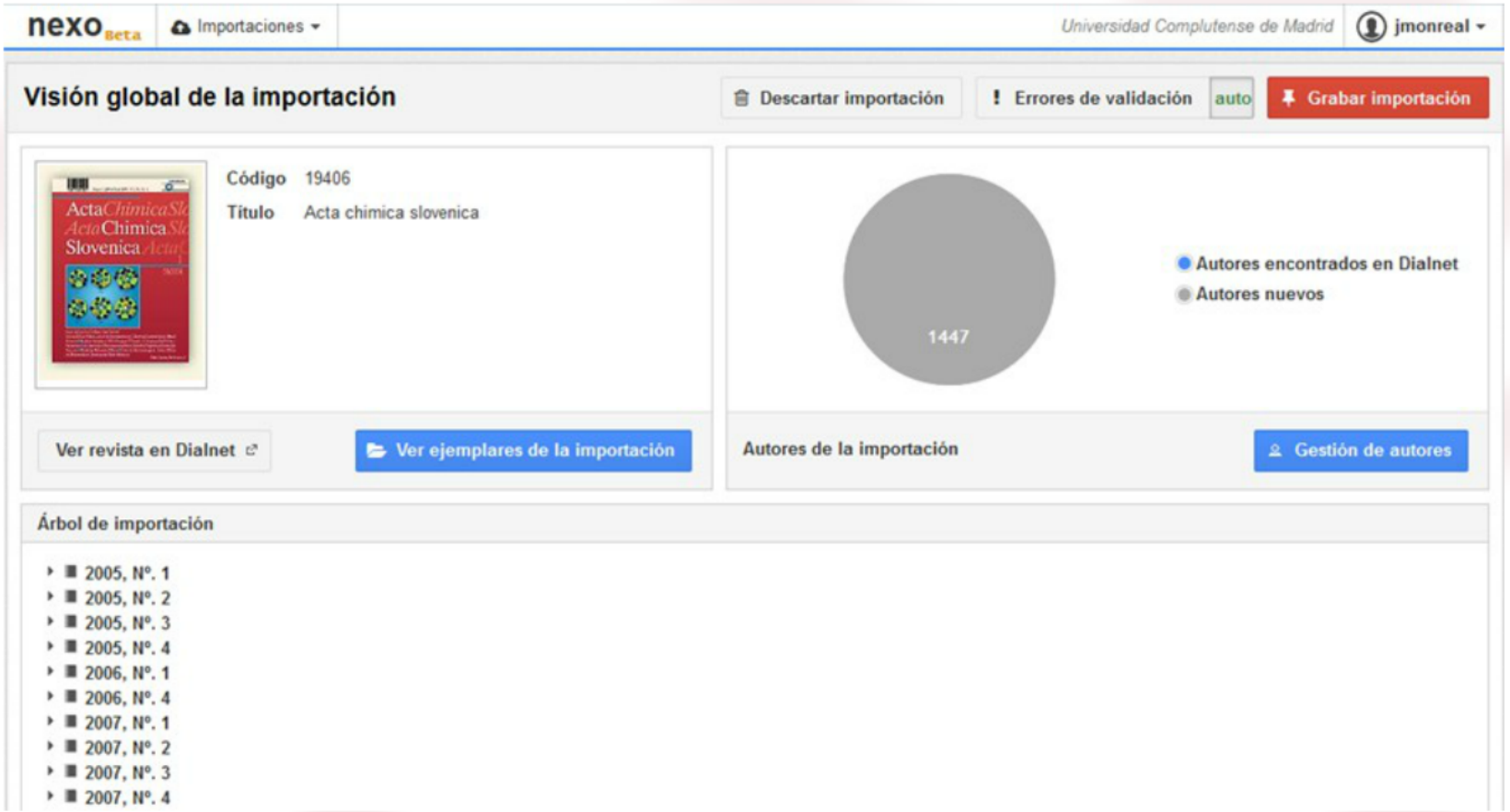


Prezi

Trabajando con Nexo

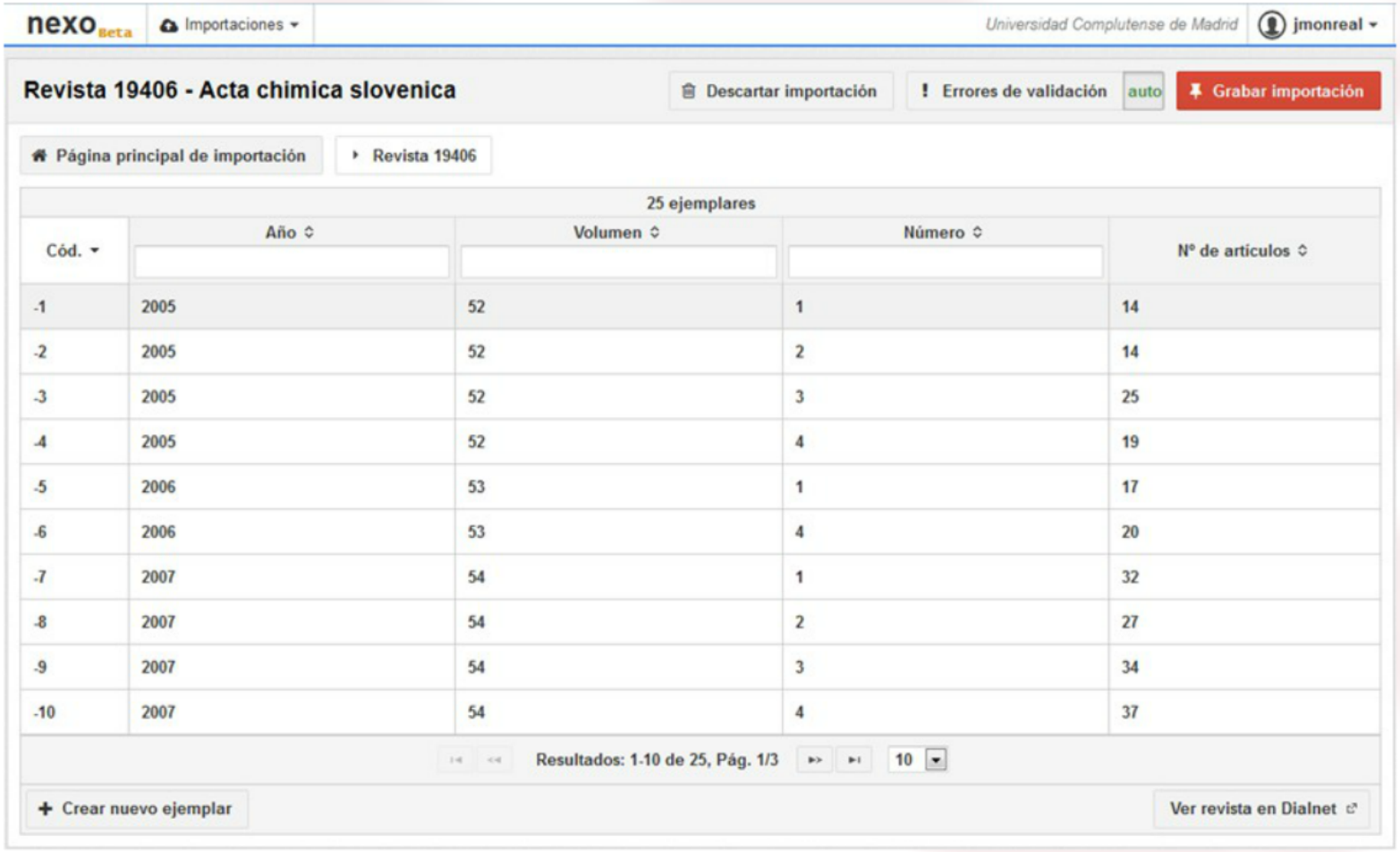
Prezi

Trabajando con Nexo



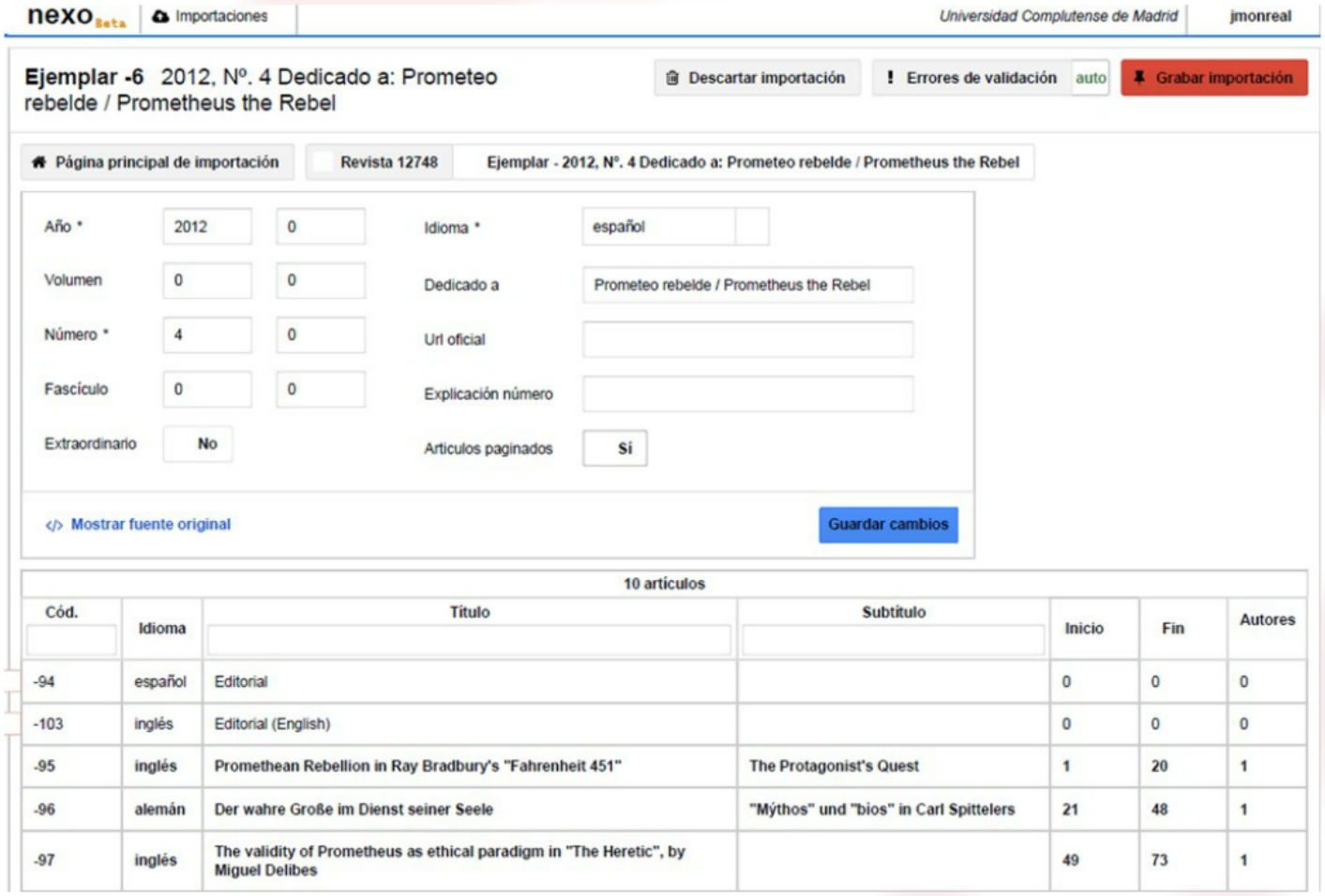
Prezi

Gestión de ejemplares y artículos

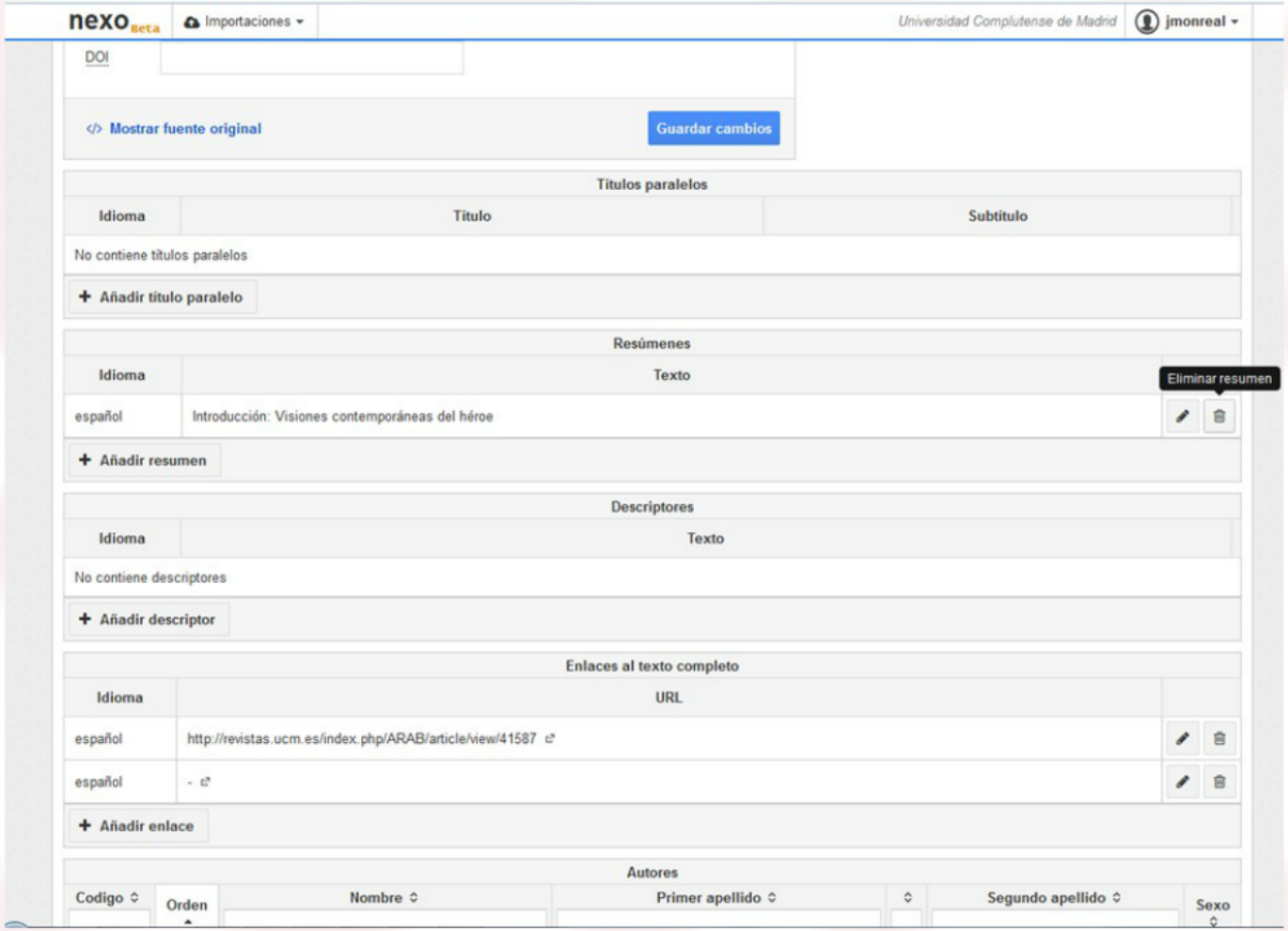


Prezi

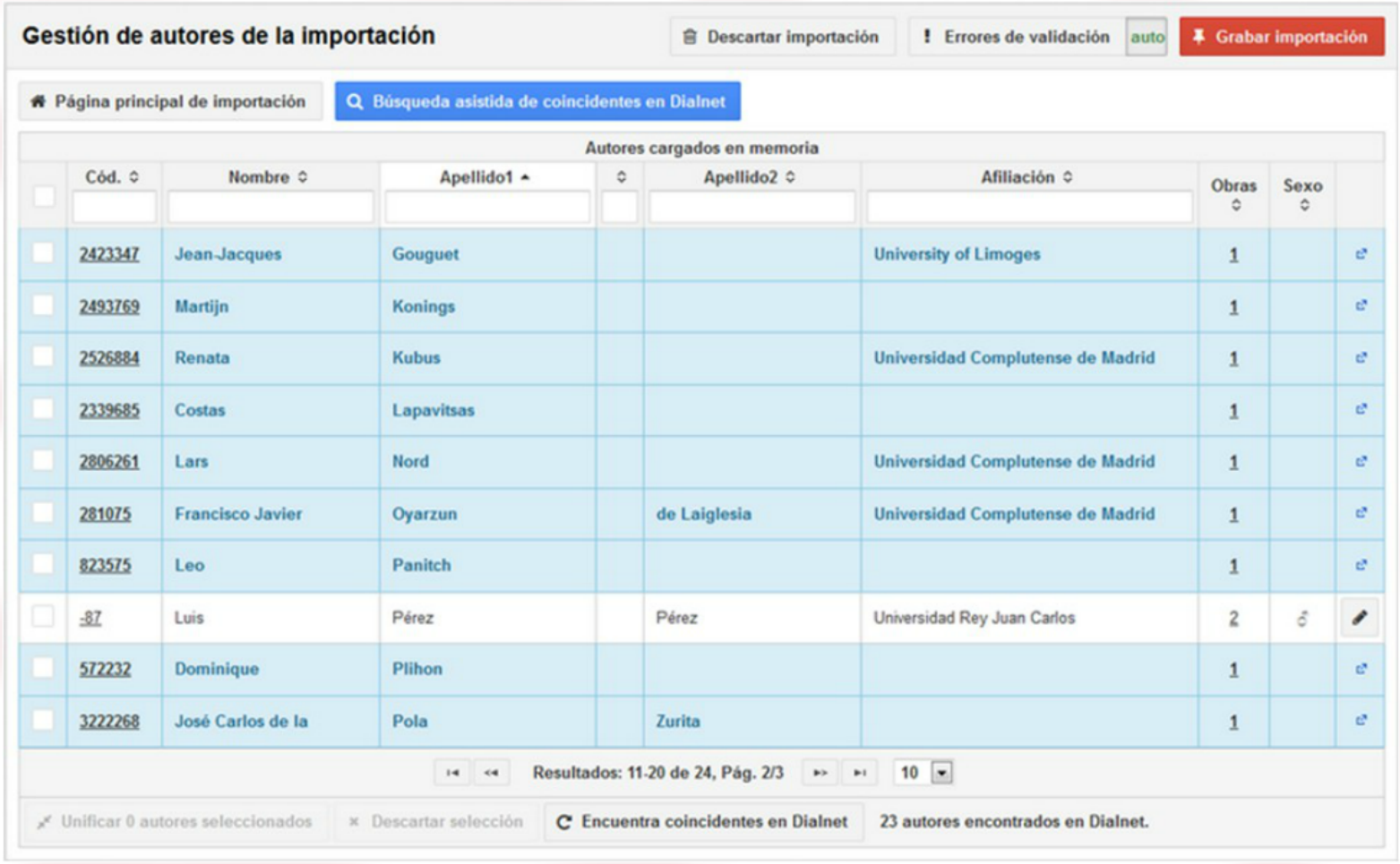
Gestión de ejemplares y artículos



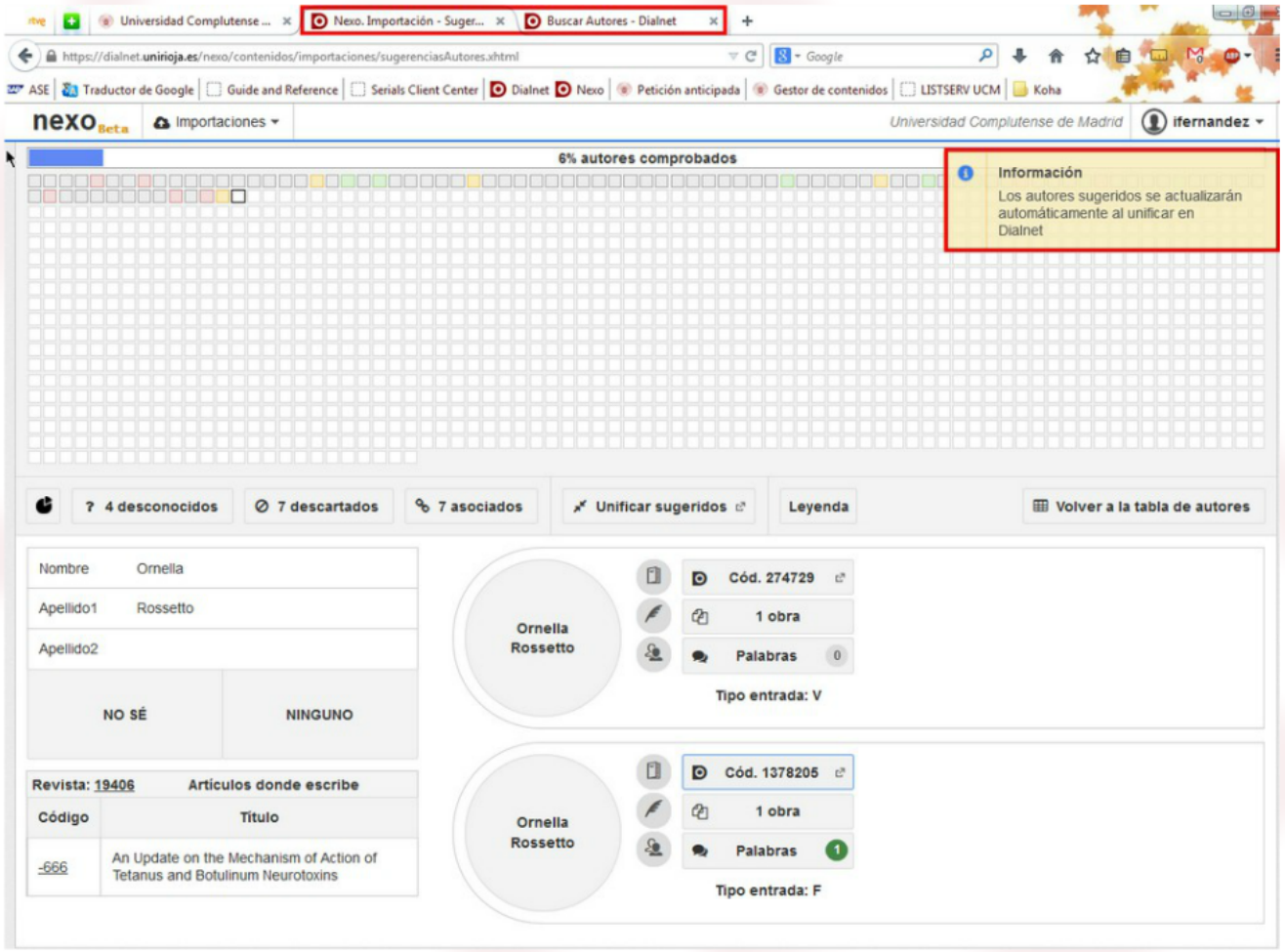
Gestión de ejemplares y artículos



Gestión de autores



Gestión de autores



Identificación de autores: dificultades

- Los datos llegan a Dialnet en forma de avalancha
- Aportados por múltiples instituciones con métodos de trabajo distintos
- Llegan por varias vías: manual, masiva y semiautomática (Nexo)

Prezi

Identificación de autores: dificultades

- Distintas tipologías de autor: personas, entidades
- Distintos tipos documentales: libros, tesis, congresos, artículos...
- Carencia de fuentes de información auxiliares
- Distintas formas del nombre por imposiciones editoriales

Prezi

Opiniones favorables

Prezi

"Me ha parecido un buen programa. He trabajado con más comodidad en Nexo que en Urraca".

"[...] Mucho mejor que Urraca. Deberían trabajar en algo similar a Nexo para sustituirla. Más amigable e intuitivo [...]".

"Actualmente, se trabaja muy bien en la parte de gestión de ejemplares: los errores se localizan con facilidad y se corrigen de forma rápida".

Prezi

Opiniones críticas

Prezi

"[...] cuando en una revista hay artículos coincidentes en su paginación [...] el programa lo califica como error grave y no permite el volcado final de datos".

Prezi

"Para corregir los posibles errores en las autorías, muchas veces es necesario ir a la página de detalle de autor [...] [Al volver] a la tabla de autores [...] el programa no te devuelve a ese punto, sino que te manda al inicio de la tabla, de manera que si estabas trabajando con un autor de la página 18 de la tabla, tienes que volver a esa posición desde el principio".

Prezi

"[...] el programa no te da la opción de eliminar directamente un autor que tiene asociado más de un artículo. Me encontré con un supuesto autor llamado "Presentación", con ocho artículos. En vez de eliminarlo de una vez desde la tabla de autores, tuve que ir entrando en cada uno de los artículos para eliminar la autoría".

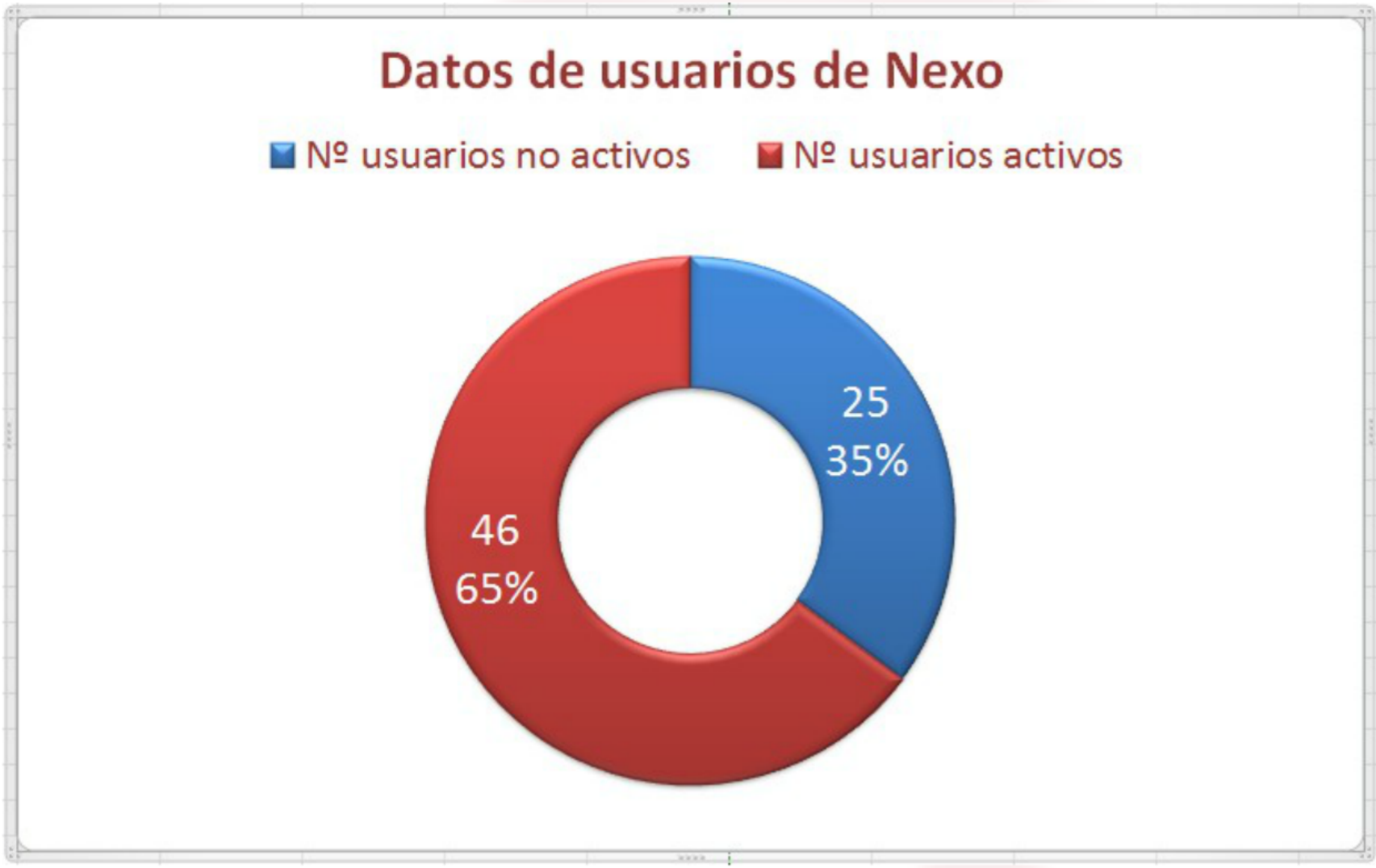
Prezi

Nexo *vs.* Urraca

- Es más rápido, ágil y amigable
- La navegación entre las distintas herramientas y ámbitos es más fácil
- Ofrece más ayudas contextuales (el bendito botón derecho)
- Es más dúctil a nuevos desarrollos y mejoras
- Permite la integración de operaciones en sí mismo y en paralelo con Dialnet-Plus
- Potencialmente, es más adaptable a distintios perfiles de vaciadores
- Podrá aplicarse tanto al vaciado por importación como a la creación de registros sueltos

Prezi

Estadísticas

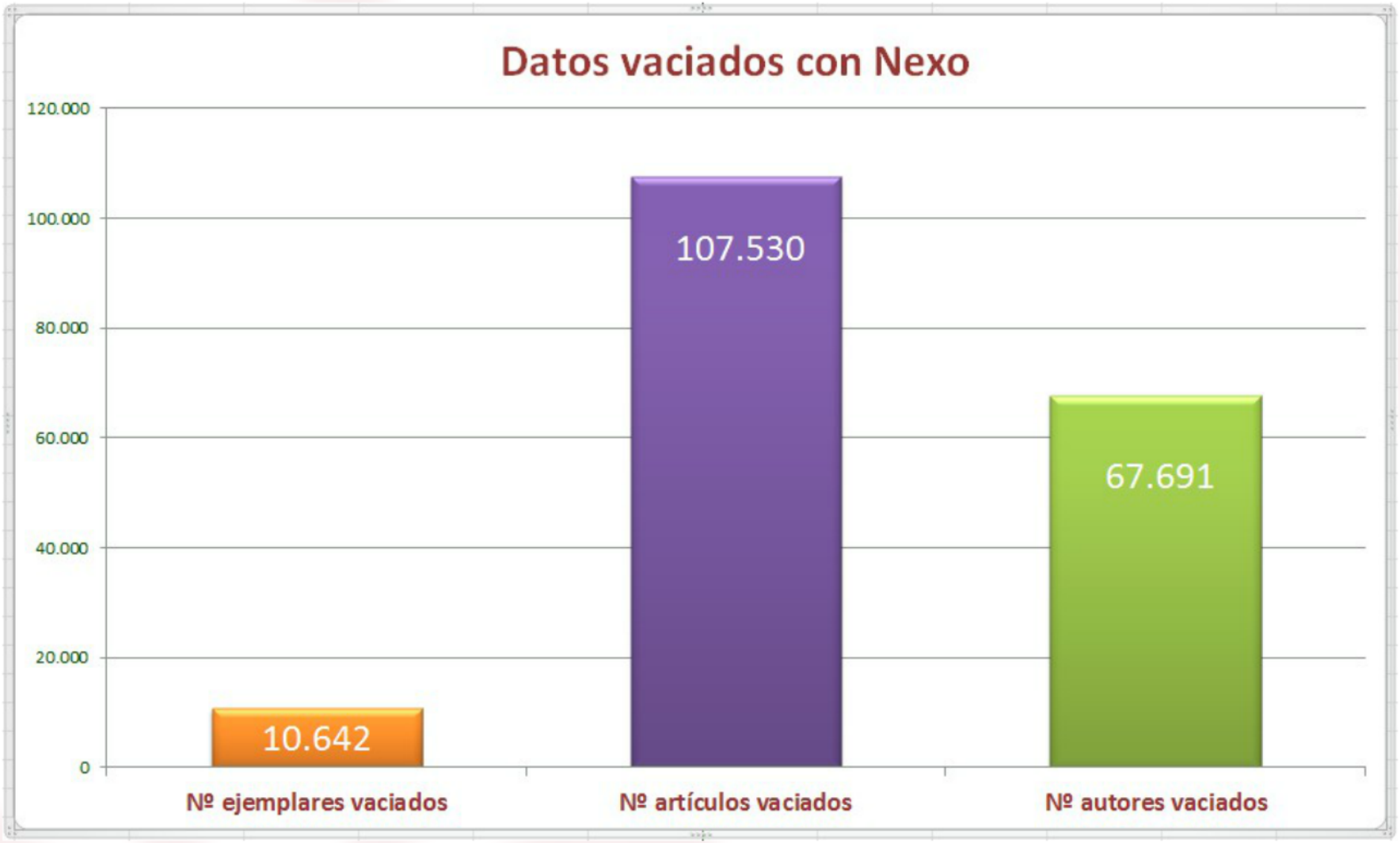


Prezi

Estadísticas

Prezi

Estadísticas



Prezi

Propuestas de mejora

Prezi

Generales

Mejorar la capacidad de *backup* de las importaciones en curso

Habilitar la posibilidad de cargas de registros de autores procedentes de otra institución o de otra base de datos

Posibilidad de aplicar Nexo al resto de los tipos documentales

Prezi

Gestión de autores

Corregir errores en el momento de la importación: separar primer y segundo apellido que vienen juntos

Mejorar la tabla de autores: volver al autor del que se parte cuando este se edita, respetar la selección de número de autores por página

Posibilidad de eliminar de una vez autores con más de un artículo

Prezi

Gestión de autores

Mayor autonomía en la edición de autores

Ampliación del ámbito de cotejo de los autores de la importación, más allá de la coincidencia de la materia

No considerar un posible error los apellidos de menos de cuatro caracteres

Prezi

Ejemplares y artículos

Ayuda contextual sobre los códigos de materia de un autor añadido ya existente en Dialnet

Mejorar el tratamiento de la variopinta paginación de los artículos para admitir: varias secuencias de páginas, paginación no numérica, solapamientos, ausencia de página final...

Integrar correctores ortográficos para distintos idiomas

Prezi

Conclusión

- 46 personas sin dedicación preferente...
- Una media de 40 horas de trabajo por persona durante 10 meses...
- Completada la importación del 57 % de 600 revistas...
- Completada la importación del 53 % de más de 200.000 artículos...

Prezi

Conclusión

...sin Nexo
no hubiera
sido posible

Prezi

Nexo: experiencia de la Biblioteca Complutense

Muchas gracias

Fundación Dialnet

Prezi