

---

# Building a librarian e- infrastructure

---

Luis Zorita  
Pubman days 2011

---

# Analysis considerations

- -open source
- -e-collaborative environment
- -Service oriented architecture

---

# Managing Library Information with PubMan

- Why eSciDoc-Pubman
  - Use case Pubman implementation
  - Migration to escidoc(1.2.1)-pubman(6.2.8)

---

# Why eSciDoc-Pubman in Libraries

- ❑ Transversality

- ❑ Inside Institution
  - ❑ Among Institutions

- ❑ Have the libraries any role to play in e-science?

- ❑ Do they have to join the e-infrastructure?

- ❑ Library - research life cycle

- ❑ Library –data management

- ❑ New Services

- ❑ New Functions

---

---

# Why eSciDoc-Pubman in Libraries

## ➤ Transversality

### ➤ Inside Institution

#### ☐ Among Institutions

#### ☐ Have the libraries any role to play in e-science?

#### ☐ Do they have to join the e-infrastructure?

#### ☐ Library - research life cycle

#### ☐ Library –data management

#### ☐ New Services

#### ☐ New Functions

---

---

# Why eSciDoc-Pubman in Libraries

## ➤ Transversality

- ☐ Inside Institution

- Among Institutions

- ☐ Have the libraries any role to play in e-science?

- ☐ Do they have to join the e-infrastructure?

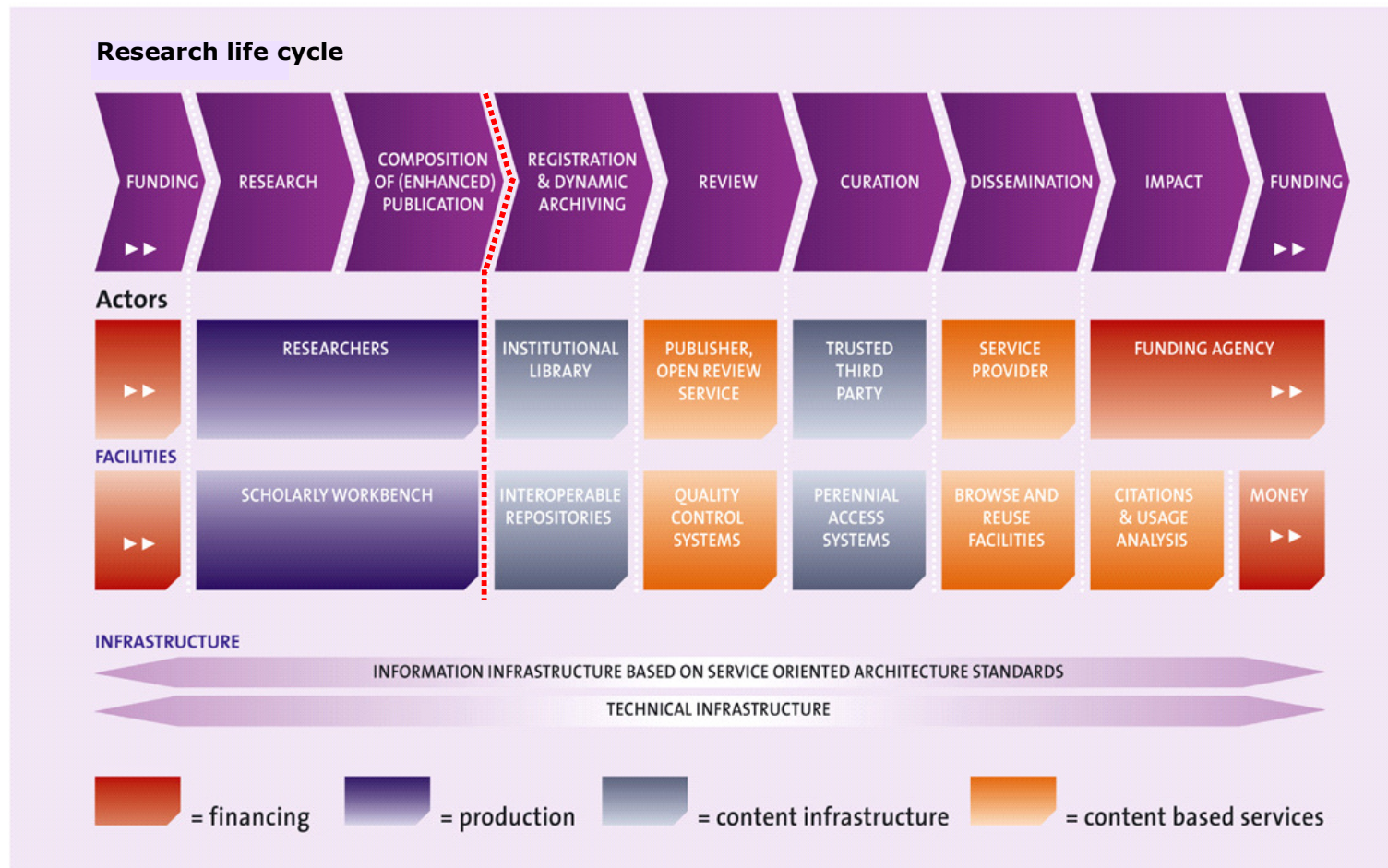
- ☐ Library - research life cycle

- ☐ Library –data management

- ☐ New Services

- ☐ New Functions

# Atlas



by Leo  
Waijers  
(SURF)

---

# Why eSciDoc-Pubman in Libraries

- ❑ Transversality

- ❑ Inside Institution
  - ❑ Among Institutions

- ❑ Have the libraries any role to play in e-science?

- ❑ Do they have to join the e-infrastructure?

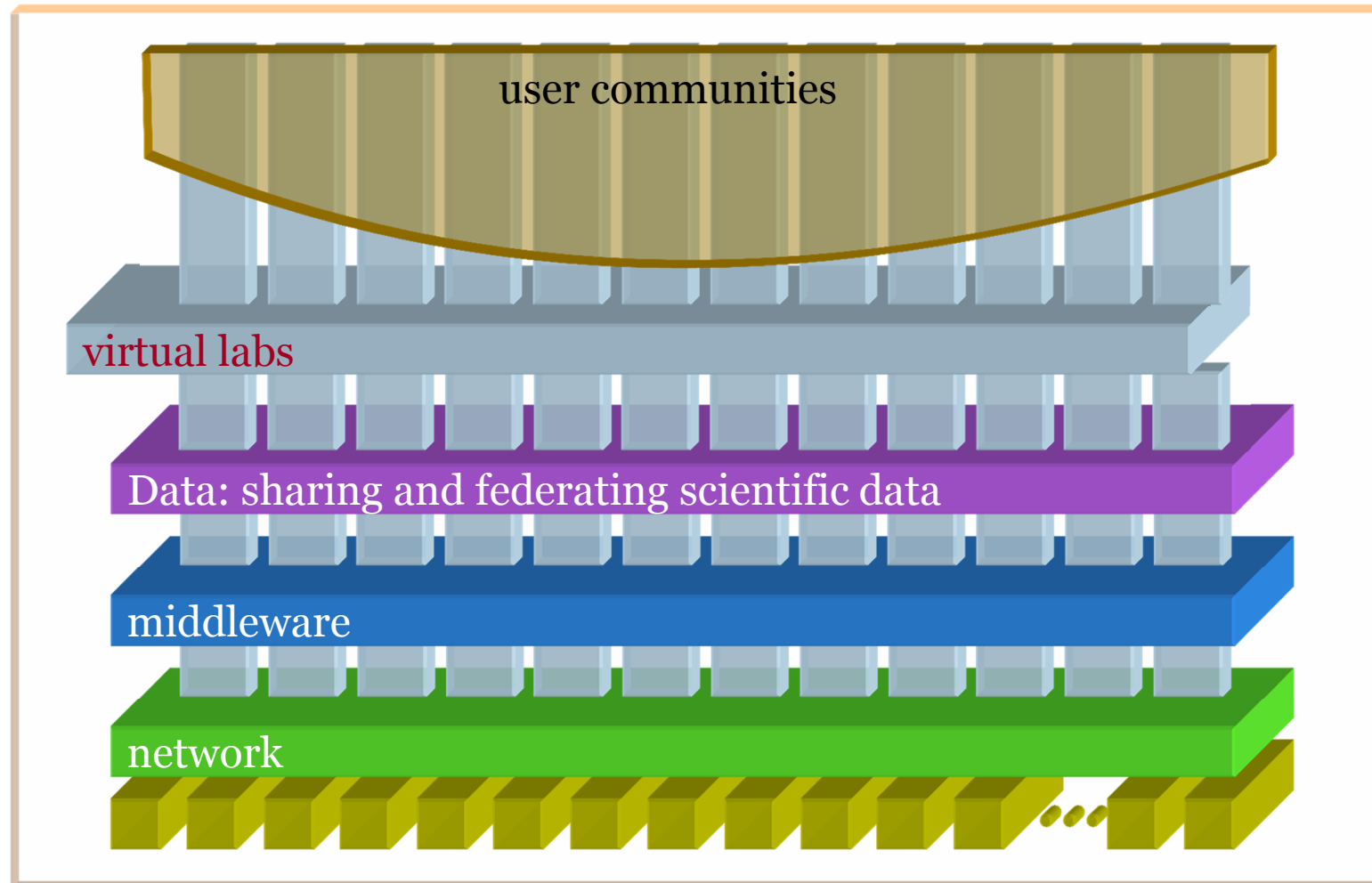
- ❑ Library - research life cycle

- **Library –data management**

- ❑ New Services

- ❑ New Functions

## e-Infrastructure - implementation



*Thanks to Mario Campolargo*

---

# Why eSciDoc-Pubman in Libraries

- ❑ Transversality

  - ❑ Inside Institution

  - ❑ Among Institutions

- ❑ Have the libraries any role to play in e-science?

  - ❑ Do they have to join the e-infraestructure?

- ❑ Library - research life cycle

- ❑ Library –data management

- **New Services**

- ❑ New Functions

---

# Why eSciDoc-Pubman in Libraries

- New future services
  - Exploitation of semantic knowledge
  - Data management, data retrieval, datamining
  - Reuse of data/information
  - New measures of impact and quality
  - Statistics
  - ???

---

# Why eSciDoc-Pubman in Libraries

- ❑ Transversality

  - ❑ Inside Institution

  - ❑ Among Institutions

- ❑ Have the libraries any role to play in e-publishing e-science?

  - ❑ Do they have to join the e-infrastructure?

- ❑ Library - research life cycle

- ❑ Library –data management

- ❑ New Services

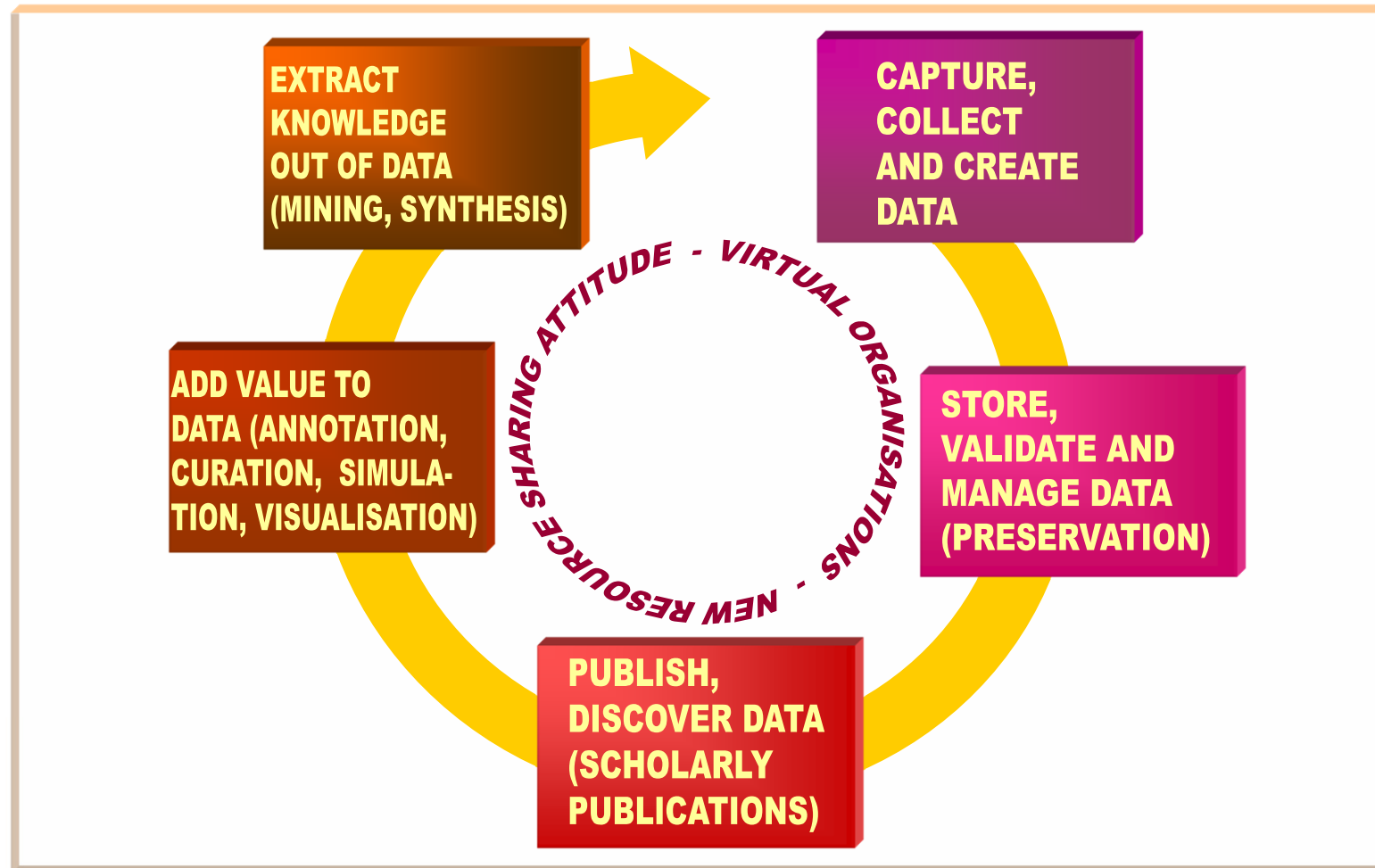
- New Functions

---

# Why eSciDoc-Pubman in Libraries

- ❑ Functions of the e-library
  - preservation (store, validate and manage data).
  - Innovation of the scientific publication cycle, including 'enhanced publications' and new models of peer review
  - Annotation, curation
  - Assessment, and impact measurements within an open access environment
  - Registration of research data
- Node in the cycle of data

# Why Pubman in Libraries



---

# Why eSciDoc-Pubman in Libraries

## □ Fedora

- Extensible metadata management
- Expressive inter-object relationships
- Relationship metadata is indexed and can be searched using semantic web query languages
- Dynamic content delivery
- Fedora repository runs as a service within a web server
- Version management

---

# Why eSciDoc-Pubman in Libraries

## □ eSciDoc

- eSciDoc is a system targeted at research organizations, universities
- eSciDoc has a organizational structure that allows collaboration among different organizations
- eSciDoc enables you to publish, visualize, manage, and work with complex digital objects
- eSciDoc provides a generic infrastructure
- eSciDoc addresses aspects of data curation
- eSciDoc supports semantic relations between objects
- eSciDoc infrastructure had a front-end web (PubMan) tested enough to be used in production environments

---

# Managing Library Information with PubMan

- Why eSciDoc-Pubman
  - Use case Pubman implementation
- Migration to escidoc(1.2.1)-pubman(6.2.8)

---

# Use case Pubman implementation

- ☐ Technical considerations
    - ☐ The Operating System had to be Centos that was not among those tested.
  - ☐ Institutional considerations
    - ☐ Pubman should be integrated in the institutional website
    - ☐ Spanish version
  - ☐ Look and feel considerations
    - ☐ Change radically the way that Pubman display the items
  - ☐ Develop a set of new statistical features
-

---

# Use case Pubman implementation

❑ The tasks carried out belong to 5 different categories:

➤ customization of css (this year's topic?)

❑ Implementing built-in functionalities

❑ adding spanish language support

❑ modification of existing functionalities

❑ creation of new statistical functionalities

---

# Use case Pubman implementation

## Implementing built-in functionalities

- ❑ Importing authors
- ❑ Adding support for the new entity 'keywords'
- ❑ Creation of new CONE entity
- ❑ Provide individual support to each keyword in the professional user interface

# Use case Pubman implementation

## Implementing built-in functionalities

- Importing authors

The screenshot shows the CoNE - Control of Named Entities web application running in a Mozilla Firefox browser. The browser's address bar displays the URL: `http://amaa.eni.es/cv/ajrone/?eScDnDnIesHndIesRvNDStIFCYNtMgSMkSMTHNjkw`. The application's header includes the title "CoNE - Control of Named Entities" and navigation links for "Home", "Search", "Enter New Entity", and "Import". Below the header, there is a table listing various entity types and their associated data formats and query options.

Entity Type	Description	Allowed eScDnDnIesHndIesRvNDStIFCYNtMgSMkSMTHNjkw
esdoomintypes	HTML	query all details
	RDF/XML	query all details
	.SDN	query all details
	Options	query all details
languages	HTML	query all details
	RDF/XML	query all details
	.SDN	query all details
	Options	query all details
esdoomintypes	HTML	query all details
	RDF/XML	query all details
	.SDN	query all details
	Options	query all details
ddc	HTML	query all details
	RDF/XML	query all details

---

# Use case Pubman implementation

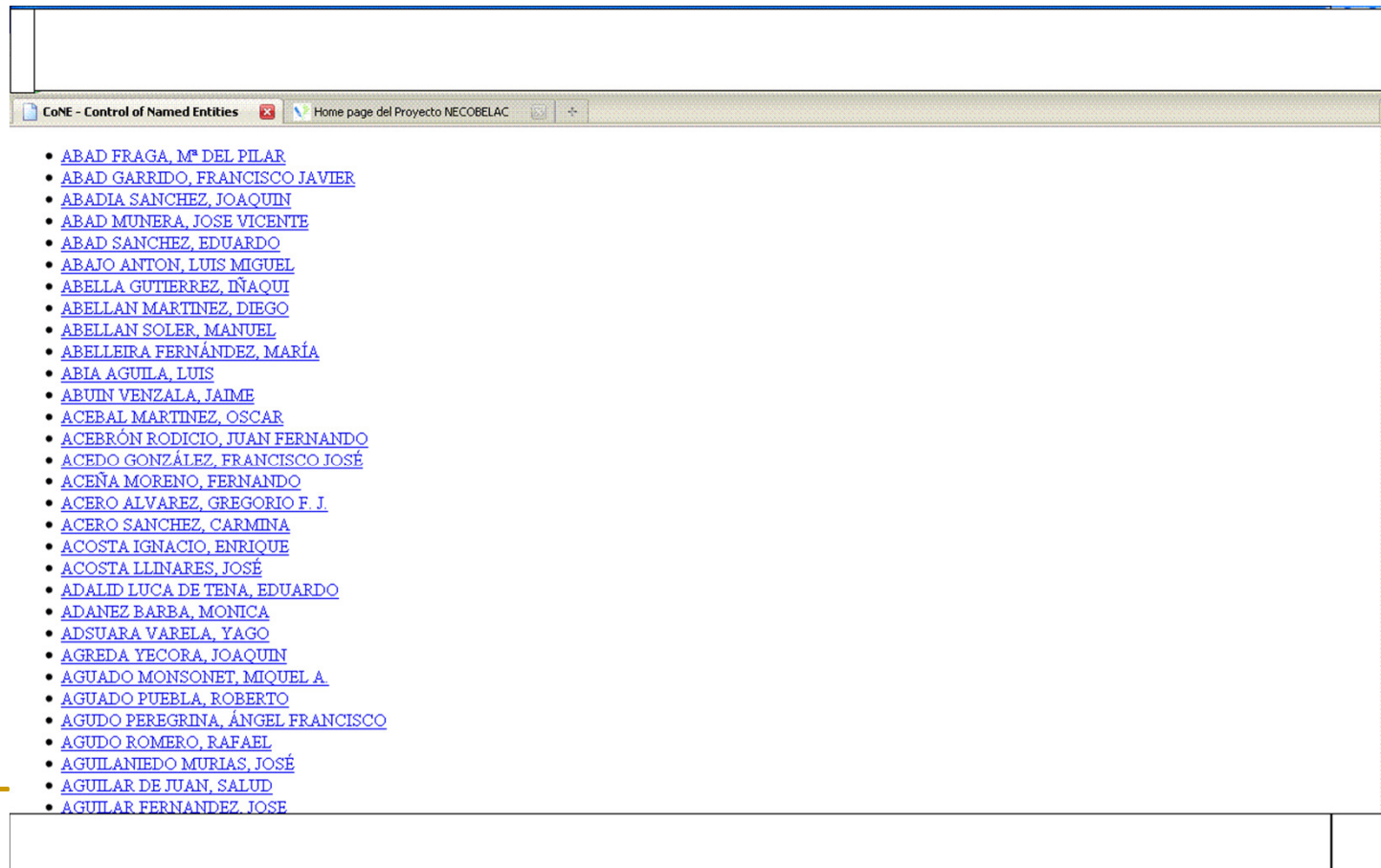
Implementing built-in functionalities-importing authors

```
<rdf:RDF>
<rdf:Description rdf:about="urn:cone:persons4457">
<dc:title>ABAD FRAGA, Ma DEL PILAR</dc:title>
</rdf:Description>
<rdf:Description rdf:about="urn:cone:persons3150">
<dc:title>ABAD GARRIDO, FRANCISCO JAVIER</dc:title>
</rdf:Description>
<rdf:Description rdf:about="urn:cone:persons2108">
<dc:title>ABADIA SANCHEZ, JOAQUIN</dc:title>
</rdf:Description>
<rdf:Description rdf:about="urn:cone:persons4142">
<dc:title>ABAD MUNERA, JOSE VICENTE </dc:title>
.....
```

# Use case Pubman implementation

## Implementing built-in functionalities

- Importing authors



---

# Use case Pubman implementation

## Implementing built-in functionalities

### ❑ Adding support for the new entity 'keywords'

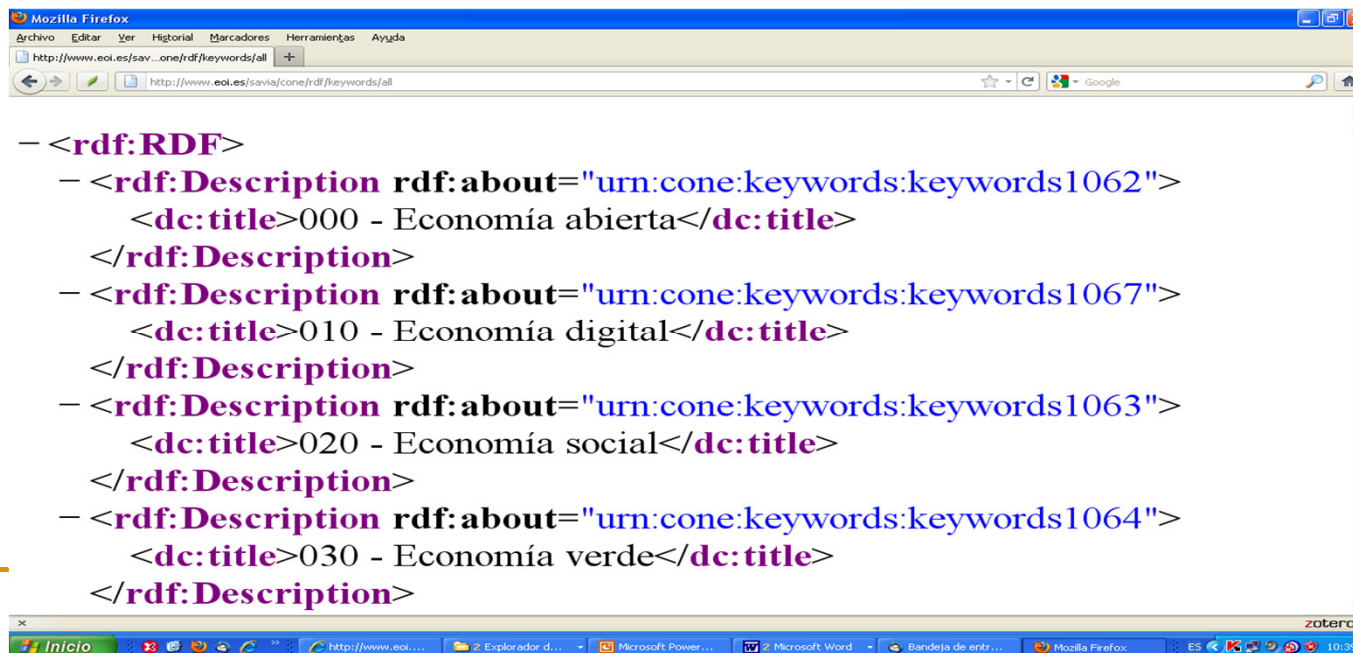
- ❑ Creation of new CONE entity

- ❑ Provide individual support to each keyword in the professional user interface

# Use case Pubman implementation

## Implementing built-in functionalities

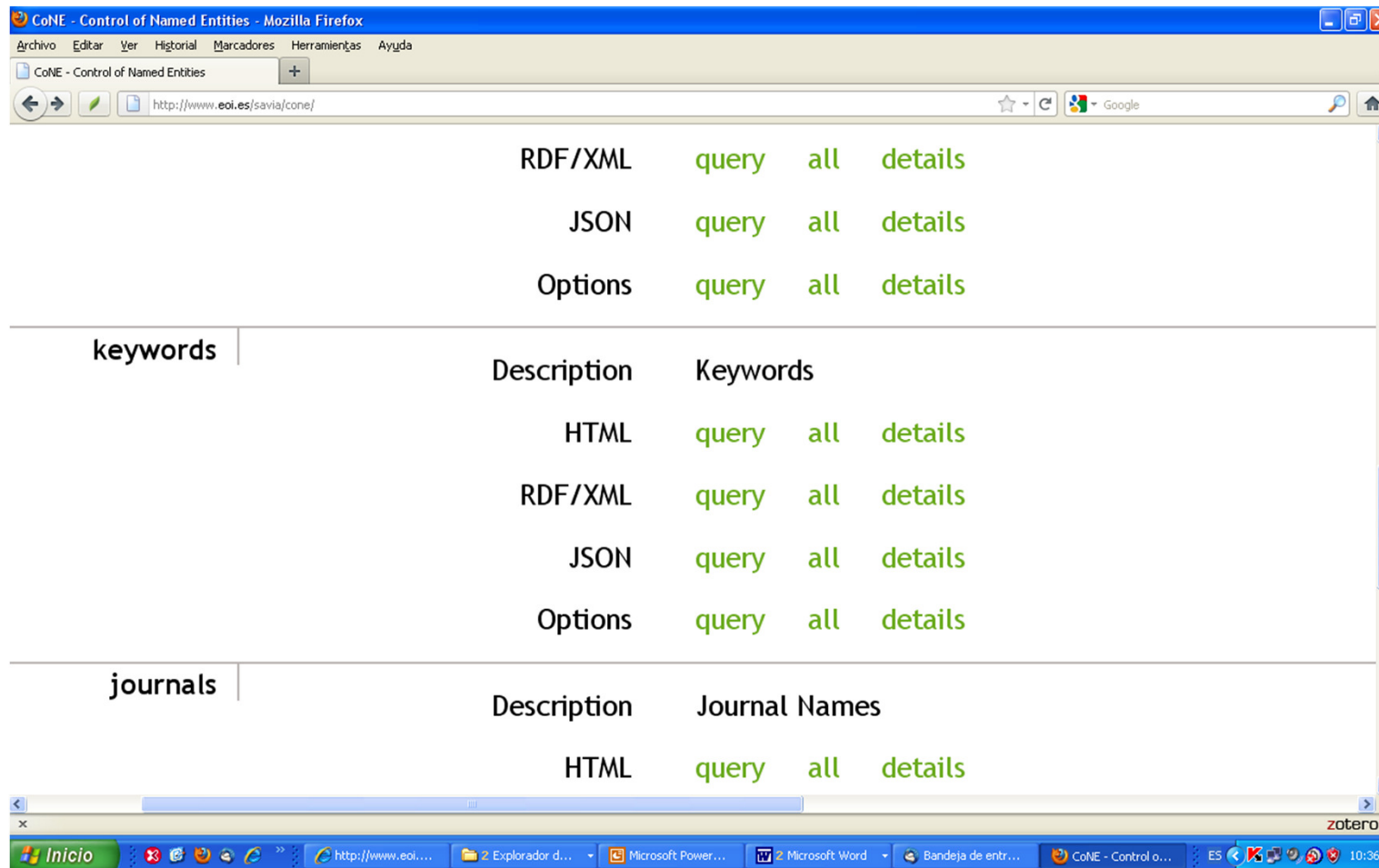
- Adding support for the new entity 'keywords'
  - Creation of new CONE entity
- Provide individual support to each keyword in the professional user interface



```
- <rdf:RDF>
  - <rdf:Description rdf:about="urn:cone:keywords:keywords1062">
    <dc:title>000 - Economía abierta</dc:title>
  </rdf:Description>
  - <rdf:Description rdf:about="urn:cone:keywords:keywords1067">
    <dc:title>010 - Economía digital</dc:title>
  </rdf:Description>
  - <rdf:Description rdf:about="urn:cone:keywords:keywords1063">
    <dc:title>020 - Economía social</dc:title>
  </rdf:Description>
  - <rdf:Description rdf:about="urn:cone:keywords:keywords1064">
    <dc:title>030 - Economía verde</dc:title>
  </rdf:Description>
```

# Use case Pubman implementation

## Implementing built-in functionalities



# Use case Pubman implementation

## Implementing built-in functionalities

- Adding support for the new entity 'keywords'
- Creation of new CONE entity
- Provide individual support to each keyword in the professional user interface

Publication Manager 5.0.3 build date 2009-08-12 16:43:54 - Mozilla Firefox

Archivo Editar Ver Historial Marcadores Herramientas Ayuda

Publication Manager 5.0.3 build date 2009-0... +

http://www.eoi.es/savia/pubman/faces/viewItemFullPage.jsp

Google

ocultar Personas y Organizaciones

palabras clave

030 - Economía verde 200 - Medio ambiente, energía y sostenibilidad 220 - Energía 222 - Energías renovables

DDC Materia

118 - Force & energy

Resumen

ocultar Contenido

Lengua de Publicación

es - Spanish; Castilian

Fecha de publicación Imp

2010-01-01

Fecha de publicación en l

2011-03-23

Fecha de aceptación

Fecha de envío

Fecha de modificación

Fecha de creación

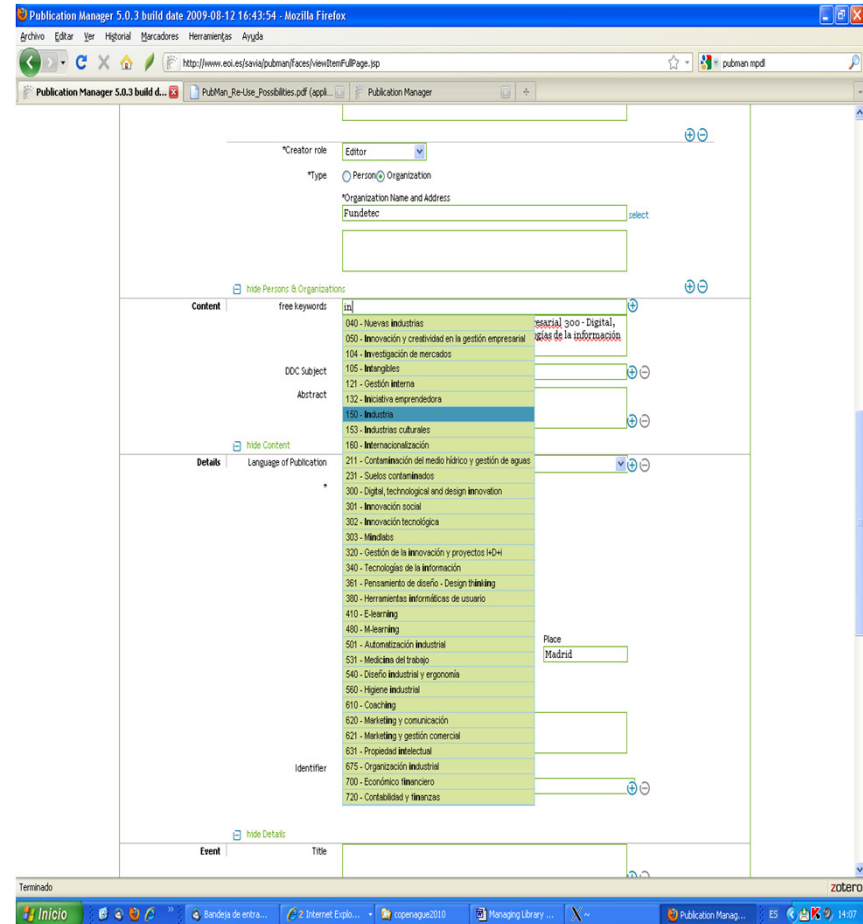
Num. Total de páginas

119

# Use case Pubman implementation

## Implementing built-in functionalities

- Adding support for the new entity 'keywords'
- Add ability to auto-suggest when adding a keyword on the new graphic interface



# Use case Pubman implementation

## Adding spanish language support

Translation files: Label\_es.properties, Tooltip\_es.properties Messages\_es.properties to spanish

Editing the contents of the faces-config.xml file and change the default local and add support for new locale

```
<locale-config>
    <default-locale>es</default-locale>
    <supported-locale>en</supported-locale>
    <supported-locale>es</supported-locale>
    <supported-locale>de</supported-locale>
    <supported-locale>ja</supported-locale>
</locale-config>
```

Add to the class folder the files. properties that contain the elements language dependent of the application: Label\_es.properties, and Tooltip\_es.properties Messages\_es.properties .

In order to get the spanish language in the selection box, we had to add the new option :

```
<f:selectItem itemValue="es" itemLabel="#{lbl.ENUM_LANGUAGE_ES}"/>
to MetaMenu.jspf
```

ENUM\_LANGUAGE\_ES = Español  
in all the files Label\_xx.properties of the classes folder.

---

# Use case Pubman implementation

## Modification of existing functionalities

### ❑ Modification of existing functionalities

- ❑ Choice a document as outstanding and show this in initial page via AJAX
- ❑ Item display and related documents

# Use case Pubman implementation

## Modification of existing functionalities

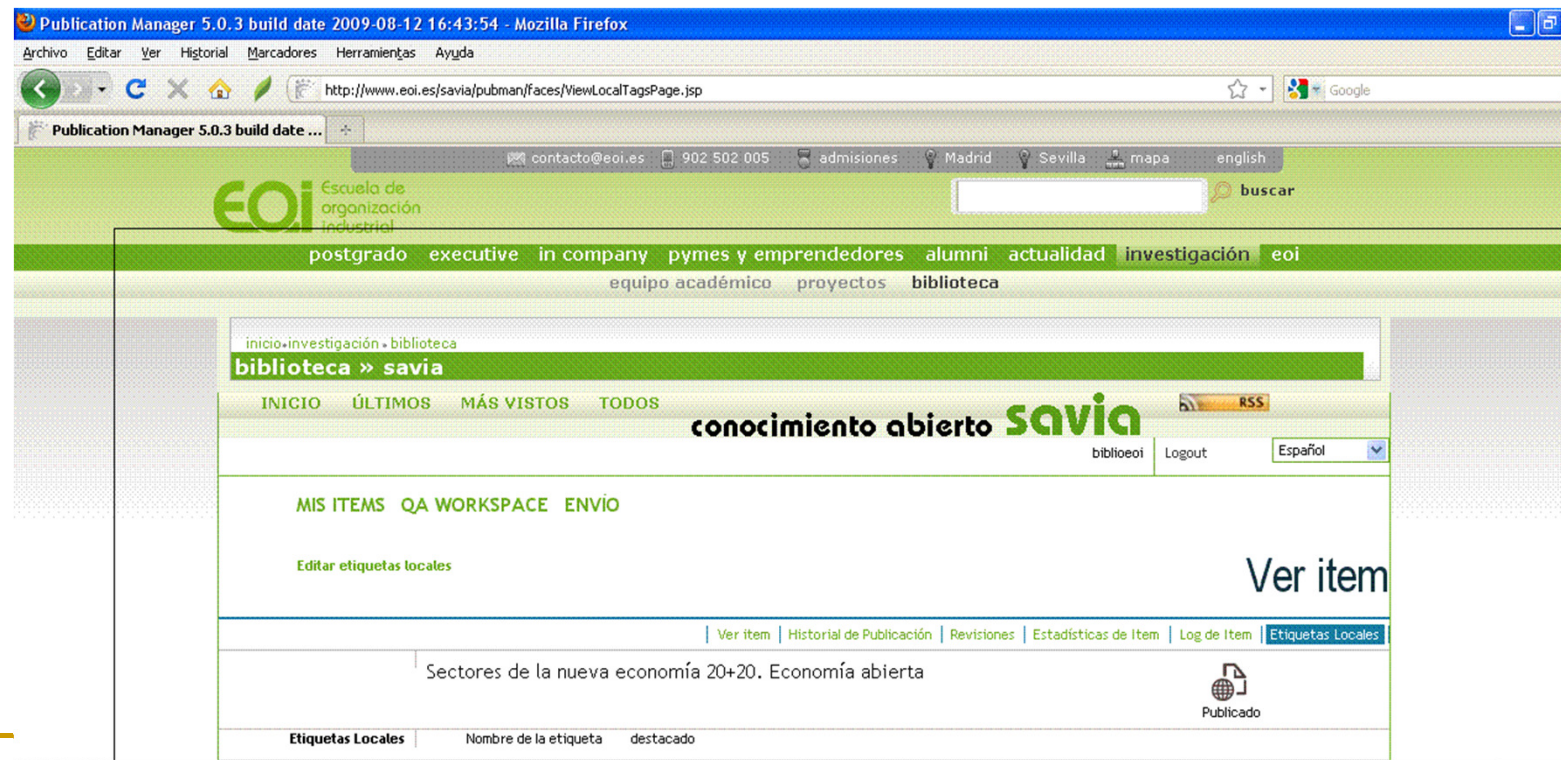
- Defining the functionality to set a document as outstanding and show this in initial page via AJAX



# Use case Pubman implementation

## Modification of existing functionalities

- Defining the functionality to set a document as outstanding and show this in initial page via AJAX



# Use case Pubman implementation

## Modification of existing functionalities

- ❑ Item display and related documents
  - ❑ add an image with name portada.gif to each document
  - ❑ include links to 4 related items
  - ❑ Bibliographic description



---

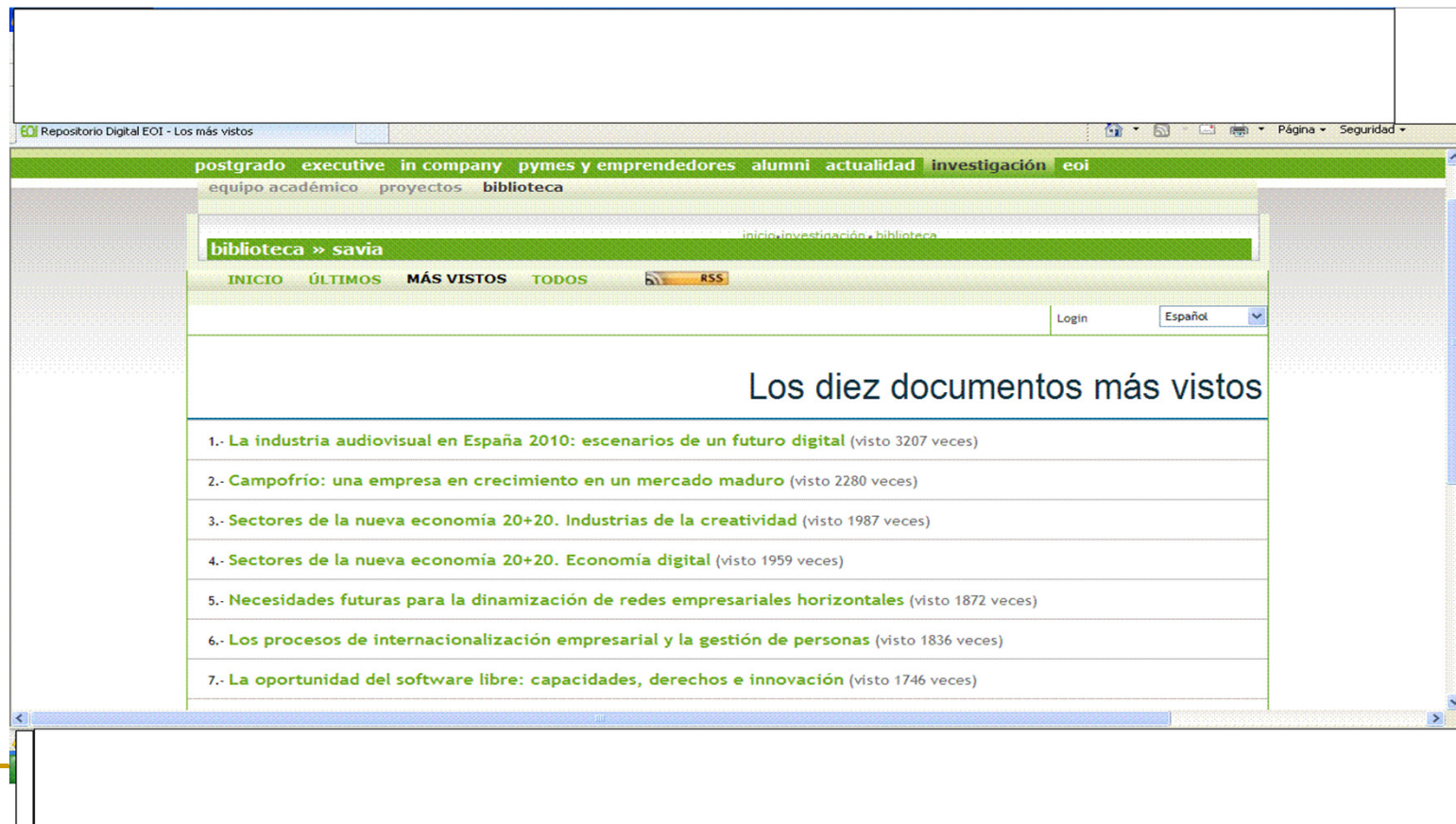
# Use case Pubman implementation

- ❑ The tasks carried out belong to 5 different categories:
    - ❑ customization of css (this year's topic?)
    - ❑ Implementing built-in functionalities
    - ❑ adding spanish language support
    - ❑ modification of existing functionalities
    - creation of new statistical functionalities
      - Top ten
      - ❑ Tags cloud
-

# Use case Pubman implementation

## Creation of new statistical functionalities

- Top ten



# Use case Pubman implementation

## Creation of new statistical functionalities

- Development of new statistical functionalities

- Top ten

- eSciDoc core framework that exposes services to get and aggregate statistical data
    - to get statistical data matching many objects (the ten most visited items), we used a eSciDoc-core persistent layer to access with sql statements to statistical data
    - report-definition-list.xml}

```
• {{{
<report-definition xmlns="http://www.eSciDoc.de/schemas/reportdefinition/0.3" objid="10">
<name>List items retrievals</name>
<scope objid="1"/>
<sql>
SELECT substring(object_id, '(escidoc:([0-9]+))') AS itemId, sum(requests) AS itemRequests
FROM _1_object_statistics
WHERE handler='de.eSciDoc.core.om.service.ItemHandler'
AND request='retrieve'
AND object_id LIKE 'escidoc:%'
GROUP BY itemId
ORDER BY itemRequests DESC
LIMIT 10 OFFSET {desde};
</sql>
</report-definition>
• }}}
•
```

- new daily scheduled task where pubman is installed and managed by the pubman framework to create every day the topten and the list of the topten tag

# Use case Pubman implementation

## Creation of new statistical functionalities

Tag cloud



---

# Use case Pubman implementation

## Creation of new statistical functionalities

- Tag cloud
  - select the tags from the keywords belonging to most visited items and it's weight is in decreasing order
  - The relative weight of each tag is given by the formula
- {{{
- pesoDec = (float)((((float)(listaPesos.get(i)-minRetrievals)/(float)diferencia)\*2);
- }}}}

---

# Managing Library Information with PubMan

- Why eSciDoc-Pubman
- Use case Pubman implementation
- Upgrading to escidoc(1.2.1)-pubman(6.2.8)

---

# Upgrading to escidoc(1.2.1)-pubman(6.2.8)

- Maybe a upgrading guide should be required
- New framework easier to install-implement
  - One server
  - One jboss
- New AdminTool under Tomcat
- Cone improved
- Specify a URL to external HTML content which will be displayed on Pubman's homepage instead of the standard content .
- Data migration

# Conclusions

- ❑ The support from eSciDoc-Pubman, through the mailing list, has been valuable.
- ❑ eSciDoc is a platform capable of creating e-publishing / e-research infrastructure, so it's a very good tool to support the e-infraestructure to participate in the workflows of the e-publishing / e-science cycle .
- ❑ It's possible to customize and configure Pubman to fit your needs
- ❑ We hope the eSciDoc-Pubman tandem continues to evolve.
- ❑ Pubman is, for university libraries, a commitment to the future.

---

# Muchas gracias por su atención

- Thank you
- Danke schön

...