

***LA PROBLEMÁTICA DE LA OBTENCIÓN,  
PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN  
DE DATOS DE LA GESTIÓN UNIVERSITARIA.***

Junio 2016.

---

---

## Objetivo de esta presentación.

Con esta presentación se pretende mostrar la experiencia de la Universidad de Cádiz ante la problemática de la obtención, procesamiento, análisis y difusión de datos de la gestión universitaria.

El objetivo no es ser una guía exhaustiva sino enseñar algunos de los problemas que nos fuimos encontrando y las soluciones instaladas.



## Historia.

El proyecto del Sistema de Información parte como una necesidad expresada por parte del nuevo equipo rectoral. La universidad se encontraba en una situación en la que no había un órgano exclusivo al que solicitar datos de carácter general.

Se decide impulsar la creación del S. I. como un lugar centralizado al que dirigir las peticiones generales de información.

Para ello se cuenta con una directora general, una directora de secretariado, 3 técnicos informáticos, y una técnica de gestión. También se ha contado con el apoyo de 2 técnicos informáticos más que ya no están en el proyecto.

---

---

No se partía de vacío, ya que había habido anteriormente un proyecto de datawarehouse. Se tenía una estructura de base de datos de Oracle que se conectaba a las bases de datos de gestión para Académico y RR. HH.

Se valoró entonces usar la alternativa de Oracle (Oracle Business Intelligence) o Pentaho. Se optó por esta última para impulsar la política de software libre.

---

---

Por otro lado existe en la UCA una cierta costumbre por parte de algunas unidades funcionales importantes (Alumnos, Asuntos Económicos, RR. HH., etc.) de crear sus propios informes de gestión interna a través de la herramienta del Oracle Discoverer. La política era ir discriminando esta práctica impulsando la sencillez del Sistema de Información.

Ahora existe un proyecto paralelo para buscar un sustituto a dicha forma de trabajar.

---

---

## Máquinas.

La política de máquinas del Centro Integrado de Tecnologías de la Información de la UCA es usar SUSE para los servidores. En concreto SUSE Linux Enterprise Server 11 es en lo que se encuentra el servidor.

Se trata de dos máquinas virtuales. Una para producción y otra para pruebas. De 8GB de memoria y 50 GB de disco.

La virtualización de las máquinas facilita en gran medida sus copias de seguridad, su duplicación, el cambio de sus características, etc. .



## Almacén de datos.

La base de datos principal donde se registran los datos está en un servidor Oracle 11G.

Posee esquemas (usuarios) para las principales áreas temáticas. Cada una de ellas tiene sus propios procedimientos de extracción y manipulación de datos.

El servidor es un cluster de máquinas externas a la propia del S. I. .

## Instalación servidor y plugins.

El servidor actualmente instalado es **Pentaho Business Analytics** en la versión **6.0.1.0.386**, sin soporte de ningún tipo. Se descargó del github de pentaho y se compiló en la máquina local.

La primera vez se instaló la versión 5.0.7. Luego se hizo una nueva instalación en la versión actual y se traspasaron los datos. En este traspaso nos encontramos con algunas pegadas, como el cambio que se hizo en la gestión de permisos de directorios, o fallos en el código como el doble escapado HTML en los parámetros de los informes.

En la versión 5 se modificó el código para añadirle algunas funcionalidades. En la versión 6 esto ha sido cambiado por scripts externos.

---

---



Los plugins a instalar tienen su importancia. El principal es Sparkl.

En él se apoyan otros plugins que se han instalado o desarrollado:

- ACL Dashboards: para controlar lo accesible o no a determinados perfiles / usuarios.
- Log Manager: para acceder a los logs desde la consola de Pentaho.
- Artifact Catalog: para establecer una forma amigable de acceso al servidor por parte de los usuarios finales.


Un tema importante en nuestra situación era también no olvidarnos de descargar los drivers JDBC de Oracle y colocarlos en `tomcat/webapps/pentaho/WEB-INF/lib/`




















---

---

**LA PROBLEMÁTICA DE LA OBTENCIÓN, PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN DE DATOS DE LA GESTIÓN UNIVERSITARIA.**

Sparkl x

**SPARKL** Sort plugins by id 

	<b>ACL Dashboards</b> acldash   	<b>Curated Artifact Catalog</b> artifactCatalog   	<b>Spanish Language Pack In...</b> languagePack_es   
<b>Log Manager</b> logManager   	<b>Pentaho Repository Synch...</b> repositorySynchronizer   	<b>Sparkl</b> sparkl   	



## Integración con LDAP.

En la UCA se dispone de un sistema de autenticación general LDAP. El sistema está mantenido por el Área de Personal y las secretarías de los centros.

No obstante se necesitaban usuarios adicionales propios para el servidor. Por eso se optó por instalar OpenLDAP en la propia máquina a través del YAST2 de Suse.

Hay que configurar un directorio LDAP local y luego combinarlo con el de la UCA usando el backend “meta” de OpenLDAP.



Básicamente el fichero principal a editar es `/etc/openldap/slapd.conf`.

....

```
moduleload back_meta.la
```

....

```
database      bdb
suffix        "dc=sistemadeinformacion,dc=uca,dc=es"
checkpoint    1024 5
cachesize     10000
rootdn        "cn=root,dc=sistemadeinformacion,dc=uca,dc=es"
rootpw        {SSHA}*
directory     /var/lib/ldap
index         objectClass eq
```

```
# Metadirectory with local entries and entries from ldap.uca.es
```

```
database      meta
suffix        "dc=uca,dc=es"
uri           "ldap://ldap.uca.es/dc=uca,dc=es"
suffixmassage "dc=ldap,dc=uca,dc=es" "dc=uca,dc=es"
uri           "ldap://localhost/dc=sistemadeinformacion,dc=uca,dc=es"
```

....

---

---

## Base de datos de usuarios.

Pero necesitábamos combinar esta autenticación con permisos específicos del tipo Director de Departamento, Vicerrector, etc.. Para ello se combina nuestro LDAP con una base de datos propia local en Postgresql. Estos cargos se toman de UXXI-RRHH en el que se ha definido una vista con los cargos actuales.

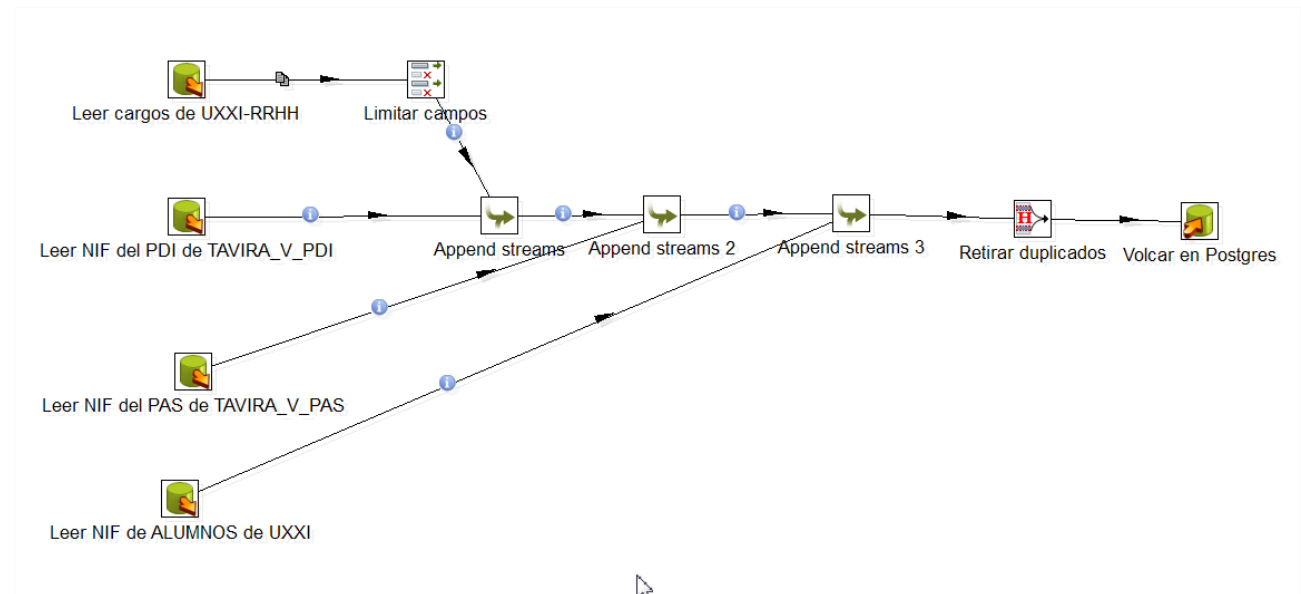
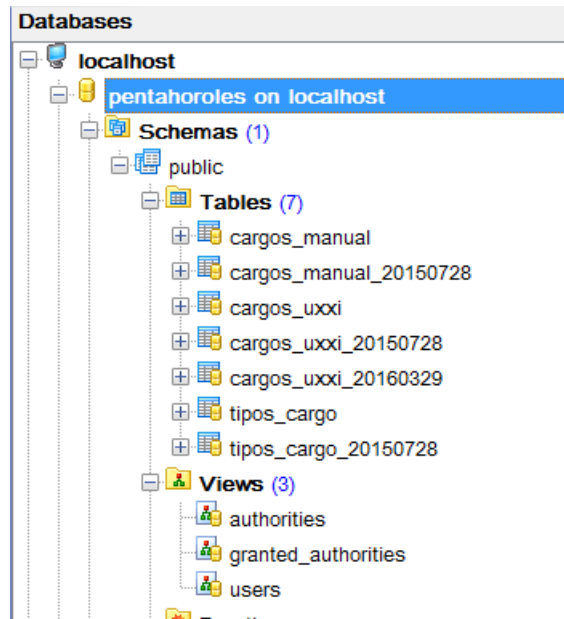
En la base de datos local que llamamos *pentahoroles* tenemos tres tablas: *cargos\_uxxi*, *cargos\_manual* y *tipos cargo*. Después hay tres vistas sobre estas tablas que son las que usa Spring Security para el tema de roles: “**granted\_authorities**” es la principal, y “**users**” y “**authorities**” se derivan de ella.

Mediante el `pg_hba.conf` limitamos el acceso a localhost, pero para poder trabajar con ella lo hacemos en local usando un túnel.

---

---

La tabla *cargos\_uxxi* se refresca diariamente a través de la ejecución programada de un trabajo kettle.



Para hacer los cambios necesarios hay que irse a pentaho-solutions/system para cambiar algunos ficheros:

- **“security.properties”** para que diga que vamos a usar LDAP para autenticación y una BD interna para autorización.
- **“pentahoObjects.spring.xml”** para que utilice esta nueva propiedad **“role.provider”**.
- Configurar **“applicationContext-security-ldap.properties”** para que Pentaho pueda hablar con el LDAP local.
- Cambiar **“applicationContext-spring-security-jdbc.properties”** para darle los datos de nuestra base de datos local.
- Hay que cambiar el bean "populator" de Spring que por omisión consulta los grupos del directorio LDAP, para que se fije en los grupos de la BD interna. Para eso, hay que cambiar **“applicationContext-spring-security-ldap.xml”**.
- Ocultar eso de "Login as Evaluator", cambiando **“pentaho.xml”**

## Página de Inicio y tema.

Para continuar con la personalización del servidor se decide crear un tema propio para la UCA. Básicamente lo que se ha hecho es copiar el tema *Crystal* como *temauca* y hacer algunos cambios.

Algunos de los ficheros a cambiar están en `biserver-ce\pentaho-solutions\system\common-ui\`. Por ejemplo:

- **login-crystal-bg.jpeg** para el fondo de la pantalla de login
- **themes.xml** para meter el tema UCA y definir su CSS principal.
- **puc-login-logo.png** para cambiar el logo de la pantalla de login.

Otros están en `\biserver-ce\tomcat\webapps\pentaho\mantle\themes`.

Por ejemplo:

- **themes.xml** para meter el tema UCA y definir su CSS principal.





Por último:

- Editar el fichero **/pentaho-solutions/system/pentaho.xml** para poner temauca como tema por defecto.
- Adaptar el fichero **/tomcat/webapps/pentaho/jsp/PUCLogin.jsp** para que la pantalla de inicio de sesión esté en español.



## Apache y acceso anónimo.

Antes del pentaho colocamos un apache. Esto nos permite un control más fino sobre determinados aspectos.

Por ejemplo podemos definir el tiempo que permanecerán en caché los ficheros de imágenes, css, etc..

Nos permite poner el servidor bajo SSL y redirigir cualquier petición al puerto 80 hacia el SSL.

También podemos filtrar el acceso a la consola a determinados usuarios y obligar a los demás a pasar por el catálogo.

Y es una forma sencilla de hacer redirecciones y alias.

---

---

Otra particularidad que tenemos es que se permite el acceso sin autenticar o acceso anónimo. Para ello lo que se hace es redireccionar desde la dirección <https://sistemadeinformacion.uca.es/publico> a un script en PHP.

Este script autentica al usuario como un pseudousuario de nombre “*publico*” y redirige al catálogo.

---

---

## Refresco de datos.

Periódicamente se van haciendo refresco de los datos desde las fuentes de origen. Casi todos esos datos van a la base de datos de Oracle donde tenemos el almacén de datos.

Se recogen de dos formas principalmente. Una es mediante jobs programados en Oracle, que suelen llamar a procedimientos PL/SQL de carga de datos. La unión con las fuentes externas se hace por dblinks.

La segunda, usada fundamentalmente con bases de datos no Oracle u otro tipo de fuentes de datos, es usar trabajos kettle que son llamados por el cron del servidor. Para ello necesitamos tener **Pentaho Data Integration** también en el servidor. Y sincronizar el fichero con los JNDIs del servidor, con nuestros PCs locales donde desarrollamos y probamos los trabajos.

---

---

## Catálogo.

Desde la UCA se vio la necesidad de desarrollar una interfaz para el usuario que de forma general accede al Sistema de Información. Este usuario no debía acceder a la consola. Y debe disponer de una experiencia individualizada, de forma que lo que pueden ver por ejemplo los Coordinadores de Grado no tenga que ser lo mismo que lo que vean los Decanos, por ejemplo.

Para ello se usó Sparkl para desarrollar un plugin. El proyecto está en el github de uno de nuestros técnicos. Básicamente se usan cuadros de mando para mostrar y manipular la lista de informes y cuadros de mando disponibles.

---

---

# LA PROBLEMÁTICA DE LA OBTENCIÓN, PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN DE DATOS DE LA GESTIÓN UNIVERSITARIA.

Para el flujo de información usan transformaciones kettle. La información base propia del catálogo la guardamos en otra base de datos del Postgresql local llamada *artifact\_catalog*, que principalmente tiene 4 tablas:

- *resources*: los recursos finales (normalmente cuadros de mando e informes )
- *sections*: las secciones en las que se estructura de forma de árbol.
- *sections\_resources*.
- *sections\_roles*: permisos sobre las secciones



The screenshot shows the 'Sistema de Información de la Universidad de Cádiz' interface. On the left is a sidebar menu with categories like 'Inicio', 'Listado completo', 'CEIMAR', 'Censos', 'Datos académicos', 'Datos económicos', 'Datos para representantes', 'Indicadores Claves', 'Indicadores SGC', 'Infraestructuras', and 'Innovación docente'. The main content area is divided into several sections: 'Inicio' (welcome message), 'CEIMAR' (Grupos del CEIMAR), 'Censos' (Censos de la Universidad de Cádiz), 'Datos académicos' (Distribución de los alumnos matriculados), and 'Datos económicos' (Relación de Acreditados). Each section contains data tables and charts.

## Entradas de peticiones de información.

Hoy en día el procedimiento para solicitar un nuevo listado/cuadro de mando es a través de un CAU propio. Estas peticiones se centralizan en las directoras, que junto con los técnicos evalúan la viabilidad de las mismas. Además, junto al Equipo de Gobierno establecen la estrategia general a seguir.

El Rector en persona, en su empeño por impulsar el tema, ha realizado reuniones del Sistema de Información con representantes de las principales unidades funcionales de la UCA, en las que se les ha solicitado su colaboración e implicación en el tema.

En ese sentido se han desarrollado reuniones con las unidades funcionales en las que se les ha pedido un interlocutor con el sistema, y recabado información sobre sus necesidades de información prioritizadas.

---

---

Ese interlocutor tiene como atribuciones la validación de los datos que se obtienen, y el establecimiento de los criterios que cada salida de información tiene.

Éste es uno de los principales problemas que tiene un Sistema de Información. La credibilidad en los datos es lo que lo sustenta. Un problema generalizado es establecer un criterio único en los cálculos. Por ejemplo, puede ser muy difícil establecer el cálculo del número de alumnos matriculados:

- ¿Se cuentan los erasmus entrantes o no se cuentan?
  - ¿Se cuentan las dobles titulaciones como una titulación, como las dos que la componen, como las 3?
  - Etc.
- 
-



Además de eso el Sistema de Información ha reproducido y reemplazado otros sistemas y salidas de información que se usaban con anterioridad.

La veracidad en los datos, la facilidad de uso, y el impulso del Equipo de Gobierno debe servir para evitar otro gran problema de un sistema de este tipo, que es que se trabaje en ello pero no se consiga cambiar la costumbre de la gente de acceder a otras fuentes de información.

En ese sentido es de gran ayuda que cualquier cosa nueva sea evaluada por el Sistema de Información en primer lugar.

---

---

## Uso de Informes.

Gran parte de las salidas de información son Informes hechos con **Pentaho Report Designer**. Se suele trabajar con copias descargadas en local que luego se suben al servidor, bien directamente o bien a través del propio PRD.

Es muy importante que la versión local del PRD sea igual que la versión del servidor de producción. Y como siempre es preferible trabajar con JNDIs en lugar de con conexiones directas. Estos JNDIs deben coincidir en local y en el servidor.

Se trata de un método más fácil y rápido que el de los cuadros de mando. Como contrapartida su estructura es más fija y su interactividad muy limitada. Se restringe al cambio de parámetros de los informes, y a introducir enlaces dentro de ellos.

---

---

Pero estamos hablando de una herramienta muy potente, que provee campos agregados automáticos, numeración de páginas, una estructura para agrupar los resultados por los campos y en los niveles que se elija, etc. .

Sus salidas son en CSV, Excel, PDF, HTML y Texto Plano; y además es posible que el informe sea diferente según el formato elegido.

Se pueden introducir subinformes comunicados mediante los campos indicados, con el informe padre.

Y por supuesto se puede parametrizar los informes.

---

---

# LA PROBLEMÁTICA DE LA OBTENCIÓN, PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN DE DATOS DE LA GESTIÓN UNIVERSITARIA.

Pentaho Report Designer - pld\_actiprof - D:\Trabajo\Harpo\Ordenación Académica\Actividades\_Académicas\svn\pld\_actiprof.prpt

File Edit View Insert Format Data Extras Window Help

pld\_actiprof

Ab Arial Narrow 6 B I U white

206%

UCA Universidad de Cádiz

Page Header

Report Header

Group Header

Group Header

Details Header

Details

Group Footer

Group Footer

Report Footer

Valoración de actividades del profesorado. Curso \$(pcurso).  
Informe preliminar.

Departamento: \$(DEPARTAMENTO)

Área: \$(AREACONOC)

PDI	Categoría	Nº de Actividades docentes	Créditos en actividades docentes	Nº de Actividades de gestión	Créditos en actividades de gestión	Nº de actividades de Inves
\$(LL1PRS) \$(LL2PRS),\$(NOMPRS)	CATEGORIA	N_REDUC_1	S_CRED_1	N_REDUC_3	S_CRED_3	N_REC
Totales del Área		TotalAreaAD	TotalAreaAD	TotalAreaAG	TotalAreaAG	TotalAr
Totales de las Áreas mostradas de este Departamento		TotalDptoAD	TotalDptoAD	TotalDptoAG	TotalDptoAG	TotalDp

Este informe integra al profesorado que tiene al menos una actividad valorada.  
Pinche sobre el nombre del PDI para ir al detalle.

En las Actividades de Investigación, al Profesorado Sustituto Interino se le podrán reconocer hasta un máximo de 6 créditos, según la Instrucción del Vicerrector de Ordenación Académica y Personal y del Vicerrector de Investigación UCA/101VOAP-VI/2015, de 28 de mayo de 2015, sobre actividades del Profesorado

Structure Data

- Master Report
  - Page Header
  - Report Header
  - Group: DEP\_CODALF
    - Group Header
    - Group Body
      - Group: ACO\_CODALF
        - Group Header
        - Details Body
          - Details Header
          - band
            - Ab label: PDI
            - Ab label: Categoría
            - Ab label: Nº de Actividades
            - Ab label: Créditos en activid
            - Ab label: Nº de Actividades
            - Ab label: Créditos en activid
            - Ab label: Nº de actividades
            - Ab label: Créditos en activid

Style Attributes

Name	Inh	Value	Expr
size & posit...			
x		289.0	
width		47.0	
y		0.0	
height		100.0%	
invisible-consum...	<input checked="" type="checkbox"/>		
visible	<input checked="" type="checkbox"/>	true	
dynamic-height	<input checked="" type="checkbox"/>		
font			
family		Arial Narrow	
font-size		6	
bold	<input checked="" type="checkbox"/>	true	
italics	<input checked="" type="checkbox"/>	false	
underline	<input checked="" type="checkbox"/>	false	
strike through	<input checked="" type="checkbox"/>	false	
smooth	<input checked="" type="checkbox"/>	auto	
embed	<input checked="" type="checkbox"/>	false	
text			

206,5 of 284,0 MB

# LA PROBLEMÁTICA DE LA OBTENCIÓN, PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN DE DATOS DE LA GESTIÓN UNIVERSITARIA.


< > 1 / 1
☰

**Grado / Máster** Centro:   
 GRADO ESCUELA DE INGENIERÍA NAVAL Y OCEÁNICA

**Tipo de Salida**   
 PDF

Ver Informe  Auto-Enviar

🏠 🔍 ↶ ↷
Página: 1 de 1
− + 80%
🖨️ 📄 📌 ➡



*RSGC-P08-01: INFORME DE RESULTADOS DE ANÁLISIS DE LA SATISFACCIÓN SEGÚN GRUPO DE INTERÉS.*  
*Tipo de estudio: GRADO*

Centro: ESCUELA DE INGENIERÍA NAVAL Y OCEÁNICA

*RSGC-P08-01: INFORME DE RESULTADOS DE ANÁLISIS DE LA SATISFACCIÓN SEGÚN GRUPO DE INTERÉS.*

TÍTULO	INDICADOR	2012-13	2013-14	2014-15	2015-16	
GRADO EN ARQUITECTURA NAVAL E INGENIERIA MARITIMA - 2501760	<b>ISGC-P08-01: Tasa de respuesta de la encuesta para el análisis de la satisfacción.</b>	Alumnado - Título				
		Alumnado - Centro				
		Alumnado - UCA				
		PDI - Título				
		PDI - Centro				
		PDI - UCA				
	<b>ISGC-P08-02: Grado de satisfacción global del alumnado con el título</b>	Grado de satisfacción global con la titulación.				
		Grado de satisfacción global con el Centro.				
		Grado de satisfacción global con la Universidad.				
		Aspectos relacionados con la información que dispone sobre el título.				
		Aspectos relacionados con la organización y desarrollo de la titulación que cursa actualmente.				
		Aspectos entorno a recursos materiales y de infraestructuras puestos a disposición del título.				
	<b>ISGC-P08-03: Grado de satisfacción global del PDI con el título</b>	Grado de satisfacción global con Satisfacción global con las asignaturas que imparte en el título.				
		Grado de satisfacción global con Grado de satisfacción global con la titulación.				
Grado de satisfacción global con Grado de satisfacción global con el Centro.						
Grado de satisfacción global con Grado de satisfacción global con la Universidad.						
Alumnado.						
Organización y desarrollo de la docencia.						
Recursos de infraestructuras.						
	<b>ISGC-P08-04: Grado de satisfacción global del PAS con el título</b>	<i>NOTA: No se realiza la encuesta por la complejidad de la asignación al título del centro.</i>				

1 / 1
22 Junio 2016

## Uso de Cuadros de Mando.

Mediante la librería CDE de Pentaho introducimos cuadros de mando en el Sistema de Información. ( <http://www.webdetails.pt/ctools/cde/> ).

Los cuadros de mando proporcionan mayor interactividad con el usuario, que puede ver evolucionar los datos, eliminar / poner series en las gráficas, cambiar de tipo de representación, parametrizar la salida de datos, etc..

Además una cualidad que les da mucha potencia es que los orígenes de datos pueden ser transformaciones kettle.



La contrapartida es que no están pensados para una salida impresa. De hecho, desde la UCA se ha mantenido la política de que los cuadros de mando quepan en una sola pantalla, sin necesidad de hacer scroll.

Lógicamente también suelen tener listados, pero estos están paginados. Así que para solventar este hecho muchas veces se colocan botones para enlazar con listados hechos con el PRD de los mismos datos, o para proporcionar salidas en Excel / CSV.

También son más laboriosos de realizar.

---

---

# LA PROBLEMÁTICA DE LA OBTENCIÓN, PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN DE DATOS DE LA GESTIÓN UNIVERSITARIA.

**CDE** New Save Save as... Reload Settings Distribución geográfica de alumnos matric...

Charts Others Generic Scripts Selects Custom Community Contributions CDF Core Functionality Widgets Saiku CFR Pivot4]

Type	Name
Group	Charts
CCC Pie Chart	Grafico1
CCC Bar Chart	Grafico2
Group	Generic
Group	Others
Group	Selects
Select Component	SelectTipo
Select Component	SelectTasa
Select Component	SelectCampus
Select Component	SelectTipoestudio
Select Component	SelectErasmus
Group	Custom
Map Component	MapTilesets2

Properties / Advanced Properties

Property	Value
Name	MapTilesets2
Listeners	['param_tipo', 'para (...)
Parameters	[["param_anomes", "pa (...)
Center Latitude	40.0028028
Center Longitude	-4.003104
Colormap	[]
Default zoom Level	3
Map Engine	openlayers
Operation Mode	Markers
Shape definitions	-
Tilesets to display as layers	default
Datasource	DSMapa
HtmlObject	DIV_Mapa



# LA PROBLEMÁTICA DE LA OBTENCIÓN, PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN DE DATOS DE LA GESTIÓN UNIVERSITARIA.

## Distribución de los alumnos matriculados por lugar de procedencia

Procedencia de los alumnos:

Alumnos de otros países de la U. E. ▾

CURSO:

2015-16 ▾

CAMPUS:

Todos los campus ▾

Estudios:

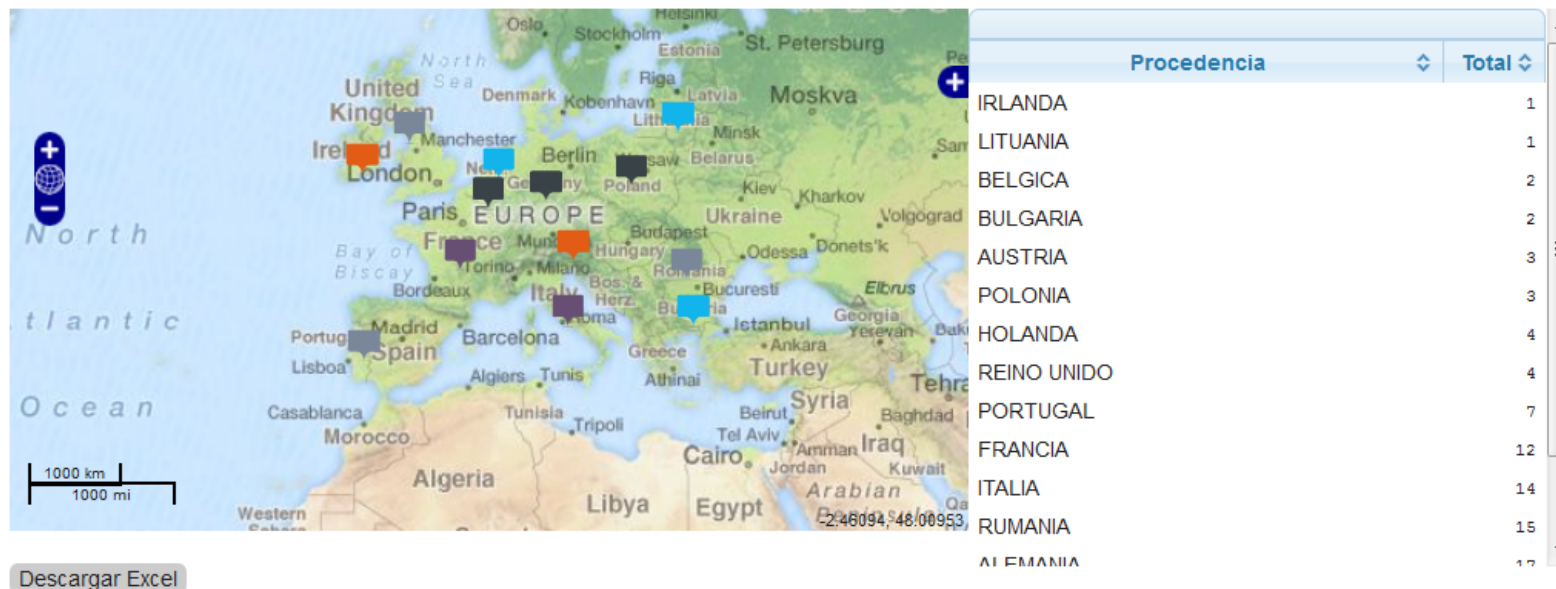
Grados ▾

Incluir Erasmus/Séneca entrantes:

No ▾

Cambiar Gráfico

Alumnos de otros países, curso 2015-16:



- Coloquese sobre los valores en los gráficos para obtener la leyenda y el dato del valor dado.
- Pulse en el botón "Cambiar Gráfico" para alternar el tipo de gráfico.

## Grandes áreas.

Las grandes áreas que abordamos en el Sistema de Información de la UCA a día de hoy son:

- Campus de Excelencia
  - Censos
  - Datos Académicos
  - Datos Económicos
  - Resultados de encuestas para el Sistema de Garantía de Calidad
  - Innovación Docente
  - Área de Personal
  - Investigación
  - Movilidad Internacional
  - Ordenación Académica
- 
-

Mención aparte merece un apartado que llamamos “**Indicadores clave**”. En él establecemos unos cuadros de mando pensados para el Equipo de Gobierno, que proporcionan un vistazo general de los grandes aspectos de la Universidad. La idea es tomar el pulso a la situación definiendo una serie de menos de 50 indicadores y ver su evolución en el tiempo mediante gráficas.

Un gran empuje al proyecto ha sido la ayuda a la elaboración de la Memoria de Responsabilidad Social. Este documento que detalla la actuación de la Universidad el curso pasado engloba casi todos los aspectos de la misma, y se complementa con listados e indicadores que en la medida de lo posible se han intentado automatizar en el Sistema de Información.

---

---

# LA PROBLEMÁTICA DE LA OBTENCIÓN, PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN DE DATOS DE LA GESTIÓN UNIVERSITARIA.

## Indicadores Claves Universidad de Cádiz.

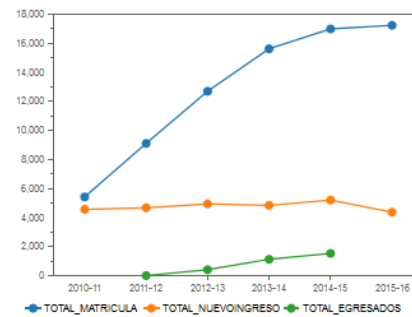
Seleccione un año para obtener sus indicadores correspondientes.

2014 ▼

### Vicerrectorado de Alumnado

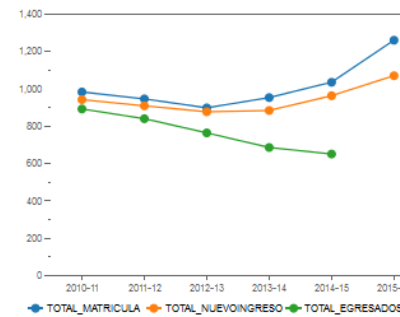
#### Alumnos en estudios de grados

Datos para el curso 2014 :  
 - Nº de matrículas: 16977  
 - Nº de nuevos ingresos: 5201  
 - Nº de egresos: 1533



#### Alumnos en estudios de másteres

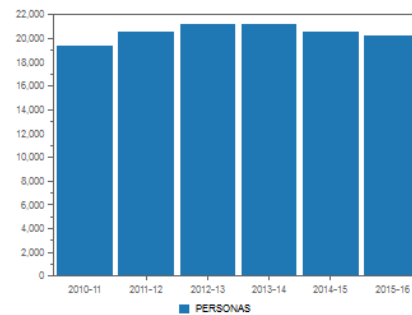
Datos para el curso 2014 :  
 - Nº de matrículas: 1035  
 - Nº de nuevos ingresos: 963  
 - Nº de egresos: 651



#### Personas matriculadas en la UCA

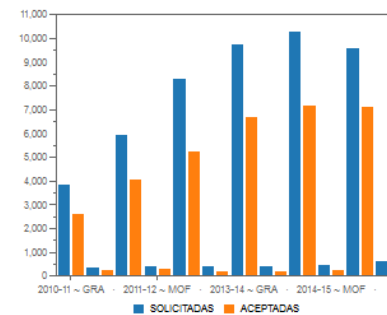
Datos para el curso 2014 :  
 - Nº de personas matriculadas en la UCA: 20618

Incluye a alumnos matriculados en Primer y Segundo Ciclo, Grado, Máster, y Doctorado



#### Solicitudes de becas MEC

Datos para el curso 2014 :  
 - Nº de solicitudes presentadas en Grado: 10308  
 - Nº de solicitudes concedidas en Grado: 7160  
 - Nº de solicitudes presentadas en Máster: 466  
 - Nº de solicitudes concedidas en Máster: 267



Este empeño en abordar las grandes áreas de la Universidad ha supuesto un reto. En el origen el S. I. estaba circunscrito a las bases de datos de operación principales (Asuntos Económicos, Académicos, y de RR. HH. ). Esto hacía que se fuera sobre bases de datos estructuradas y potentes y con un mantenimiento de datos exhaustivo.

Pero poco a poco se han ido incorporando diversas fuentes de datos (Administración electrónica, Investigación y SICA, Posgrado, becas, Oficina de Relaciones Internacionales, etc.) cuya consistencia podría no ser tan firme. Es aún más importante en estos casos contar con la ayuda del interlocutor de las Unidades funcionales.

Todo este empeño puede desembocar en un peligro que yo llamo personalmente “*matar al mensajero*”. Al hacer más fácil el acceso a datos que hasta entonces no lo eran tanto, los errores en los datos que puedan haber, están más al descubierto.

Así por ejemplo un profesor puede elevar una petición para indicar que él no da determinada asignatura, o un PAS indicar que no se encuentra en determinado lugar para el censo.

La tendencia natural es decir que el S. I. falla, aunque lo que en realidad ocurre es que los datos están incorrectos en origen. Es importante entonces contar con la ayuda de la Unidad Funcional implicada, y dar una respuesta rápida, y que sirva en lo posible para evitar repetir la situación. Una forma de hacer esto es informar al usuario que dirija su petición a la unidad encargada (en nuestro ejemplo del PAS en el censo que se dirija a Secretaría General, etc.).

---

---

## Estadísticas de uso.

Un apartado del catálogo son las métricas de uso. Son unos cuadros de mando que atacan los datos de uso del resto del S. I. Nos sirven para valorar los esfuerzos e intentar dinamizar las consultas más utilizadas y mejorar la experiencia del usuario.

Estos datos se captan de los logs de uso del tomcat y mediante transformaciones kettle se introducen en tablas en el almacén de datos.

Así podemos saber por ejemplo:

- Los 25 recursos que han supuesto un mayor tiempo de procesamiento, para el intervalo de fechas especificado.
  - Los 25 usuarios que han realizado más accesos.
  - Los 25 recursos que han sido más utilizados.
  - La distribución por día de la semana y hora del día de los accesos a los recursos
  - Etc.
- 
-



## Accesos por día semana/hora

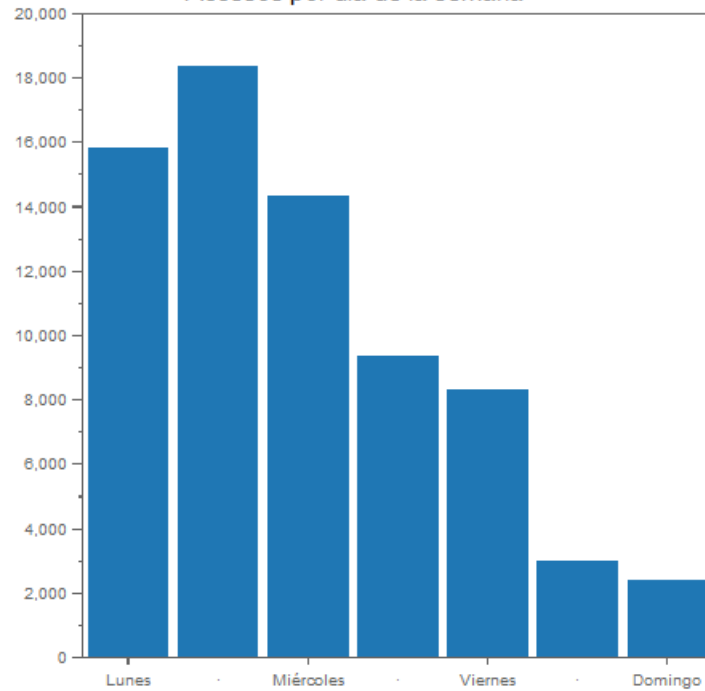
Selecciona un rango de fechas:

2016-01-01 > 2016-06-22

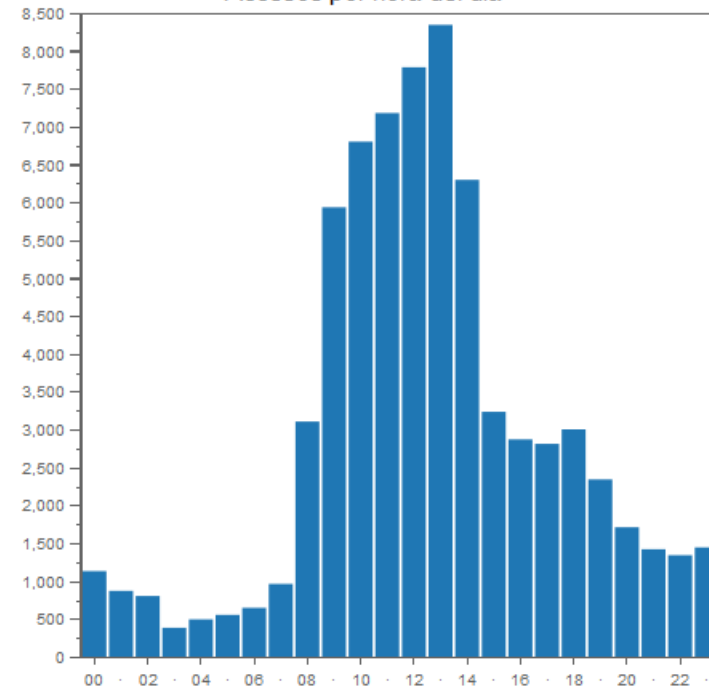
Exportar días

Exportar horas

Accesos por día de la semana



Accesos por hora del día





## Pruebas de Selenium y scripts de mantenimiento.

Se mantiene un seguimiento del servidor del S. I. de forma automática. Esta bajo supervisión Zabbix que nos avisa de inmediato ante caídas del mismo.

También se realizan pruebas periódicas del contenido. Para ello se han hecho unos scripts utilizando selenium.

Los scripts de refresco de datos también guardan logs y envían correos en caso de producirse algún fallo.



Además disponemos de un script que realiza dos acciones que habíamos perdido al actualizar el servidor. Dicho script:

- Borra los directorios home de los usuarios que hayan entrado en el sistema, menos los de los técnicos del S. I. .
- Sustituye los ficheros creados por programaciones por sus versiones actualizadas.

Para ello usa el propio API que proporciona Pentaho.

---

---

## LA PROBLEMÁTICA DE LA OBTENCIÓN, PROCESAMIENTO, ANÁLISIS Y DIFUSIÓN DE DATOS DE LA GESTIÓN UNIVERSITARIA.

```
# coding: utf-8

import requests
import xml.etree.ElementTree as ET
import logging
import itertools
logging.basicConfig(filename='/srv/pentaho/biserver-ce/tomcat/logs/script_pentaho_6.log', level=logging.INFO, format='%(asctime)s - %(levelname)s - %(message)s')

# Datos del usuario con el que acceder a la API
ACCESS_USER = 'admin'
ACCESS_PASSW = '*'

# Si se pone a False, las carpetas se moverán a la papelera
DELETE_PERMANENTLY = True

# URLs de la API
ENDPOINTS = {
    'getFolders': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/:home/children?depth=-1&showHidden=false&filter=u{0}{1}*%7CFOLDERS',
    'delete': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/delete',
    'deletePermanent': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/deletepermanent',
    'getPldAsignaturasCanceladas': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/:public:ordacad:actividades/children?depth=-1&showHidden=false&filter=pld_asignaturas_canceladas',
    'getPldAsignaturasCanceladasXLS': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/:public:ordacad:actividades/children?depth=-1&showHidden=false&filter=pld_asignaturas_canceladas_xls',
    'renombrarPldAsignaturasCanceladas': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/:public:ordacad:actividades:{0}/rename?newName=pld_asignaturas_canceladas',
    'getContratoProgramaAAMDS': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/:public:Innovación%20Docente/children?depth=-1&showHidden=false&filter=contratoProgramaAAMDS',
    'getContratoProgramaAAMDSPDF': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/:public:Innovación%20Docente/children?depth=-1&showHidden=false&filter=contratoProgramaAAMDS_PDF',
    'renombrarContratoProgramaAAMDS': 'https://sistemadeinformacion.uca.es/pentaho/api/repo/files/:public:Innovación%20Docente:{0}/rename?newName=contratoProgramaAAMDS',
}

# Carpetas de usuarios que se mantendrán
IGNORE_USERS = [
    'u31-----',
    'u32-----',
    'u31-----'
]
```

Correos de contacto:

[dgsi@uca.es](mailto:dgsi@uca.es)

[dsisi@uca.es](mailto:dsisi@uca.es)