

Soporte a los nuevos desarrollos para el Sistema Integrado de Datos Institucionales

Tabla de contenido

<u>1. Introducción</u>	2
<u>1.1. Alcance</u>	2
<u>1.2. Definiciones y acrónimos</u>	2
<u>2. Pliego de prescripciones técnicas</u>	2
<u>2.1. Introducción</u>	2
<u>2.2. Objetivos del trabajo</u>	3
<u>2.3. Marco tecnológico</u>	5
<u>2.4. Metodología</u>	5
<u>2.5. Sucesión de hitos temporales</u>	6

1. Introducción

1.1. Alcance

El presente documento recoge el pliego de prescripciones técnicas que debe ser tenido en cuenta para la contratación de soporte técnico externo en el marco del proyecto del Sistema Integrado de Datos Institucionales (SIDI).

1.2. Definiciones y acrónimos

CII	Centro de Inteligencia Institucional de la UCM
<i>Data mart</i>	Subconjunto de información orientado a un área de negocio
<i>Data warehouse</i>	Almacén institucional de datos
<i>Data warehousing</i>	Técnicas empleadas en la construcción y explotación de un almacén de datos
ELT	Extract, Load and Transform. Procesos de transferencia y manipulación de datos
PAS	Personal de Administración y Servicios de la UCM
PDI	Personal Docente e Investigador de la UCM
RPT	Relación de Puestos de Trabajo de la UCM
SIDI	Sistema Integrado de Datos Institucionales
SIIU	Sistema Integrado de Información Universitaria del Ministerio de Educación
SSII	Servicios Informáticos de la UCM
UCM	Universidad Complutense de Madrid

2. Pliego de prescripciones técnicas

2.1. Introducción

Fruto de la colaboración entre el CII y los SSII, y a lo largo de los años 2016 y 2017, se ha implantado un sistema de inteligencia institucional (SIDI) cuyo objetivo es dar cobertura a las necesidades de información de los diferentes servicios de gestión de la UCM.

Durante esa primera etapa se ha construido mediante herramientas *software* de Oracle (Warehouse Builder 11 y Business Intelligence 12) un *data warehouse* que comprende la información básica de gestión estructurada en los siguientes *data marts*:

- Gestión académica.
 - Pruebas de acceso a la universidad y proceso de admisión.
 - Gestión de matrícula y rendimiento académico.
 - Egresados: Titulados y abandono.
- Indicadores de calidad y acreditación de titulaciones.

- Planificación docente.
- Recursos humanos.
- SIIU.

Esta información se ha ofertado fundamentalmente a los Vicerrectorados de Calidad y de Estudiantes y al Departamento de Estudios e Imagen Corporativa.

También se ha elaborado una relación de demandas de datos por parte de la Universidad, constatando la necesidad de nuevos desarrollos (para cubrir la expansión o cambio en las necesidades del usuario) y de un mantenimiento adaptativo (para que el sistema sea acorde a las nuevas versiones de software de *data warehousing* que surgen en el mercado).

Por lo que se refiere a los nuevos desarrollos, es fundamental tener en cuenta tres vertientes:

- El sistema debe ser capaz de satisfacer las nuevas demandas de información en las áreas de gestión académica, de planificación docente y de recursos humanos (que derivan de la extensión del acceso al SIDI a colectivos que no tomaron parte en la etapa inicial del proyecto, como otros Vicerrectorados, Centros y Servicios Centrales).
- Se deben incorporar al SIDI nuevos *data marts* (como pueden ser los de gestión económica, investigación y monitorización de uso del sistema).
- Hay que incluir cuadros de mando y permitir la generación automática de informes predeterminados que sean distribuidos a usuarios seleccionados y que ayuden al cumplimiento de tareas previsibles y periódicas (como elaboración de *rankings* y actividades de gestión de equipos decanales y de coordinadores de títulos académicos).

Por todo ello, y teniendo en cuenta la escasez de recursos humanos con que cuenta el equipo técnico del proyecto, se requiere la contratación de una empresa externa que proporcione un conocimiento de valor añadido en el desarrollo de los flujos de trabajo propios del CII.

La labor de dicha empresa consistirá en integrarse con los equipos de trabajo que tanto el CII como los SSII provean, de modo que el *know-how* metodológico, funcional y técnico de la empresa se vaya transfiriendo al equipo humano de la UCM a medida que vaya siendo efectiva su intervención.

A continuación se detallan los objetivos principales del trabajo que desempeñará la empresa contratada, y el marco tecnológico y metodológico al que deberá atenderse. Por último, se indica la sucesión de hitos temporales que deben ser satisfechos.

2.2. Objetivos del trabajo

La necesidad fundamental que subyace a esta contratación es la de hacer evolucionar el presente sistema de inteligencia institucional, SIDI (basado en las técnicas y conceptos habituales de *data warehousing*) de modo que contemple en profundidad las grandes áreas de gestión de nuestra Universidad. Tomando éstas como referencia, se describen a continuación los componentes del sistema o *data marts* cuyo desarrollo o evolución deben abordarse, y que formarán parte del *data warehouse* o almacén institucional de datos generales de la UCM.

Hay que tener en cuenta que parte de los *data marts* ya están implementados y el presente contrato supone una evolución de los mismos, por lo que se hace necesario un conocimiento previo del sistema SIDI para optimizar el rendimiento del proyecto.

- Gestión académica. Este *data mart*, ya implantado, debe profundizar en la inclusión de todos aquellos nuevos indicadores que no estén presentes en la actualidad, y que se determinarán tras las oportunas reuniones con los nuevos grupos funcionales implicados.
- Planificación docente. Este *data mart* comprende los indicadores de información relativos a la plantilla de personal docente, los departamentos en los que presta su actividad y las cargas docentes del personal. Los datos que ya ofrece deben someterse a un proceso continuo de revisión que contemple las nuevas reglas de negocio que se consideren oportunas por parte de los usuarios funcionales del proyecto.
- Recursos humanos. El *data mart* de recursos humanos contempla a todo el personal de la UCM: PAS y PDI. Para el proyecto objeto de la presente contratación, se considera apropiado incluir aspectos adicionales a los ya implantados, tales como (por citar algunos) evolución de la carrera profesional de los diferentes colectivos, adecuación del conjunto de personal a la RPT en vigor o a posibles modificaciones de la misma y análisis del grado de envejecimiento de la plantilla relacionado con el relevo generacional.
- SIIU. La UCM suministra periódicamente información al SIIU relacionada con las áreas de gestión académica, económica y de recursos humanos. El *data mart* ya implementado permite establecer los oportunos controles de calidad antes del envío de los datos. Para el presente contrato se plantea establecer un mecanismo de captura de la información publicada en la *web* del SIIU para realizar una comparativa con la que el *data mart* almacena.

NOTA: Para más información sobre el SIIU, consúltese la página correspondiente del Ministerio de Educación, Cultura y Deporte (<http://www.mecd.gob.es/educacion-mecd/areas-educacion/universidades/estadisticas-informes/siiu.html>).

- Gestión económica. Este *data mart* debe incluir al menos aquellos indicadores asociados a la gestión del personal implicado en procesos de docencia. Su finalidad fundamental en un primer estadio es la de conocer el coste por estudiante de cada titulación. Para ello debe englobar información relativa a tasas académicas, transferencias nominativas que llegan a la UCM procedentes de la Comunidad Autónoma de Madrid y costes de personal. En la actualidad se está implementando un sistema de inteligencia de negocio propio del área de gestión económica, que queda fuera del alcance del presente proyecto. El *data mart* que se propone en esta contratación no tendrá la envergadura de aquél, pero deberá contemplar la posibilidad de nutrirse de los indicadores de actividad que ese sistema vaya generando. De este modo, y a futuro, su entidad irá viéndose incrementada, facilitando además una integración con el *data warehouse* general que el presente proyecto irá construyendo.
- Investigación. Este subconjunto de información debe centrarse en la relación del profesorado, de los grupos de investigación y de los departamentos con los resultados de la investigación desarrollada en la UCM. Para ello contará con el acceso a una base de datos relacional propia del área de investigación (actualmente en desarrollo), que podrá ser almacenada en el *data warehouse*. No se contempla un análisis avanzado de la información, sino simplemente un sistema de visualización de los indicadores construidos sobre la base de datos de investigación.
- Monitorización de uso de la plataforma. Se desarrollará un *data mart* que capture información del sistema SIDI y que permita determinar aspectos de uso del mismo

tales como informes top, informes que no se usan, usuarios top, usuarios que no acceden a la plataforma, informes de rendimiento lento, etc.

Adicionalmente a los *data marts* enumerados, durante el proyecto deberá contemplarse tanto la formación de usuarios de SIDI (bien sean consultores o bien creadores de informes) como la adaptación de las herramientas actuales a los nuevos estándares del mercado, en concreto el cambio de Oracle Warehouse Builder a Oracle Data Integrator como plataforma de ELT en los desarrollos realizados para captura y procesado de la información.

2.3. Marco tecnológico

La plataforma disponible en los SSII para desarrollar el proyecto está basada en Oracle (*suites* de productos en versiones 12 instalados en la infraestructura de la UCM). Por tanto, la arquitectura general consta de los siguientes componentes:

- Base de datos Oracle Enterprise 12.
- Herramienta de ELT Oracle Data Integrator 12.
- Herramienta de inteligencia de negocio Oracle Business Intelligence Enterprise Edition 12.

En la actualidad, los procesos de ELT están implementados mediante Oracle Warehouse Builder. Siguiendo las líneas de desarrollo de los productos de Oracle, se realizará una migración en SIDI hacia la herramienta Oracle Data Integrator, para lo que se requiere un análisis previo de los desarrollos ya implementados.

La instalación dentro de la infraestructura de seguridad y comunicaciones de la UCM requiere que los productos utilizados por los usuarios finales estén integrados con el servidor LDAP institucional (Oracle Directory Server, en proceso de migración a OpenLDAP).

La instalación del entorno de trabajo es responsabilidad de la UCM. No obstante, la empresa podrá ser requerida para aportar conocimiento tecnológico de cara a conseguir la mejor configuración según los fines del proyecto.

Los sistemas fuente de que se nutre el *data warehouse* abarcan diversos sistemas de bases de datos relacionales (Oracle, Microsoft SQL Server, MySQL, PostgreSQL), repositorios de ficheros (de texto plano y XML) y servicios *web*.

Los productos de salida del sistema son los facilitados por las herramientas *software* utilizadas (acceso vía *web* a informes predeterminados, a consultas específicas, a cuadros de mando y a navegación dirigida; envío por e-mail o a servidores de gestión y documentación de datos en formatos Excel, PDF, CSV, XML o de texto plano; publicación de servicios *web*).

2.4. Metodología

El proyecto se estructurará mediante la creación de los *data marts* referidos en el apartado de "Objetivos del trabajo" centrados en áreas concretas, entregables en periodos relativamente breves de tiempo, y que, a medida que vayan siendo creados, se integren en el *data warehouse* institucional.

El orden en el que los diferentes *data marts* deben ser desarrollados es el indicado más adelante (ver apartado de "Sucesión de hitos temporales"). La implementación secuencial de los diferentes sectores a abordar permitirá acotar el ritmo de producción previsto y ajustarlo a las necesidades de información de los usuarios finales del CII.

Para cada *data mart* a implantar se deberán contemplar las siguientes fases. Cada una de ellas conlleva la entrega, además de los productos que constituyen su objetivo principal, de la pertinente documentación. En todas ellas, la empresa contará con la participación activa del equipo humano de la UCM, tanto para gestionar las necesarias reuniones con los grupos de usuarios funcionales implicados como para la implementación y desarrollo de las fases en cuestión.

- Análisis de los objetivos de negocio. Determinación de fuentes de datos, colectivos de usuarios destinatarios de la información, demanda de información a satisfacer y, en su caso, mecanismos actualmente implementados.
- Activos de información que sustentarán el subconjunto de información. Recopilación de definiciones de negocio a considerar, diseño preliminar de los productos de datos a ofrecer a los usuarios finales y aspectos relacionados como formatos, granularidad, temporalidad, seguridad, etc.
- Modelado multidimensional. Diseño lógico y físico de la parte del *data warehouse* que sustentará el *data mart* considerado.
- Implementación. Utilización de metodologías ágiles para la implementación del *data mart* en sucesivas iteraciones que contemplen su entrega y validación por parte de los usuarios funcionales. Como entregable final se obtendrá el producto de datos plenamente operativo.

2.5. Sucesión de hitos temporales

Se consideran como hitos fundamentales, a cumplir durante el año 2018 y en el orden que se indica, los siguientes:

- Hito 1: Inclusión de nuevos desarrollos en el *data mart* de recursos humanos (del 1 de Enero de 2018 al 28 de Febrero de 2018).
- Hito 2: Creación del *data mart* de investigación (del 1 de Marzo de 2018 al 30 de Junio de 2018).
- Hito 3: Creación del *data mart* de gestión económica (del 1 de Julio de 2018 al 15 de Noviembre de 2018).
- Hito 4: Creación del *data mart* de monitorización de uso de la plataforma (del 15 de Noviembre de 2018 al 31 de Diciembre de 2018).
- Hito 5: Cursos de formación de los usuarios en el uso de la plataforma (se estiman unas 5 sesiones formativas de 4 horas de duración cada una repartidas a lo largo del año 2018).
- Hito 6: Generación automática de informes basados en los nuevos *data marts* (en paralelo a los objetivos precedentes) e incorporación de la información en los cuadros de mando que correspondan.
- Hito 7: Migración a la plataforma Oracle Data Integrator de los desarrollos existentes (a medida que vayan estando disponibles y en paralelo a los objetivos precedentes).

Cada uno de los cuatro primeros hitos podrá estar asociado a una facturación parcial del importe total, según se indica a continuación. La suma de los importes de los cuatro primeros hitos, dado que se extienden durante toda la duración del contrato, supondrá el 100% de total del importe. La facturación del cuarto y último hito sólo será tramitada una vez que se compruebe que el resto de hitos no facturables han sido alcanzados.

Importe facturable por hitos, en porcentaje respecto al total del importe de la contratación:

- Hito 1: 18%
- Hito 2: 40%
- Hito 3: 30%
- Hito 4: 12%