

**ANEXO III**
**REF.: 2015/CP/144**
**MODELO DE CONVOCATORIA DE CONTRATACIÓN**
**AXUDA**

Referencia: Proxecto MINECO ESP2014-55996-C2-2-R			
Título: Técnicas de I.A. para la explotación del catálogo de Gaia: Análisis, validación y visualización.			
CENTRO: Escuela Técnica Superior de Náutica y Máquinas			
OBJECTO DO CONTRATO: Desenvolvemento de aplicacións JAVA en entornos de cómputo intensivo (Apache Hadoop, Spark, CUDA, OpenCL) no campo da astrofísica computaciol (prox. espacial Gaia - ESA). Colaborador/a nas tarefas do (proxecto/convenio) de investigación que se indica.			
LUGAR E HORARIO: Laboratorio 3.2. Grupo de Investigación LIA[2]. Horario 9h-14h e 16h-18:30h de luns a venres			
GRUPO DE COTIZACIÓN NO RÉXIME XERAL DA SEGURIDADE SOCIAL: Grupo I			
CATEGORÍA PROFESIONAL *(ACORDO DO 17 DE NOVEMBRO DE 2014 SOBRE O PERSOAL CONTRATADO CON CARGO AO CAPÍTULO VI DOS ORZAMENTOS DA UDC)			
Investigador asociado <input type="checkbox"/>		Investigador en formación <input checked="" type="checkbox"/>	
Técnico de apoio á investigación <input type="checkbox"/>		Axudante de apoio á investigación <input type="checkbox"/>	
Técnico administrativo <input type="checkbox"/>		Axudante administrativo <input type="checkbox"/>	
XORNADA	Tempo Completo: <input checked="" type="checkbox"/>	Tempo Parcial: <input type="checkbox"/>	
DURACIÓN	13 meses prorrogables	DATA APROX. DE INICIO	01/12/2015
RETRIBUCIÓNS BRUTAS : 1368,5			
PARTIDA ORZAMENTARIA: 6310210517-541A 6490200			

\* O pagamento da retribución realizarase en 12 mensualidades (inclúese en cada mensualidade a parte proporcional de paga extraordinaria correspondente)

**REQUISITOS**

TITULACIÓN: Enxeñeiro en Informática / Graduado en Informática
EXPERIENCIA: Ver detalles na Memoria
OUTROS REQUISITOS: Como investigador en formación o candidato/a debe estra matriculado nun programa de doutoramento do SUG.

**LUGAR DE PRESENTACIÓN DE SOLICITUDES E DOCUMENTACIÓN**

LUGAR: Por correo electrónico a Carlos Dafonte (dafonte@udc.es) ou no laboratorio 3.2 da FIC
PRAZO: 10 días hábiles dende a publicación na páxina web da UDC (Actualidade PDI)

**CRITERIOS E ÓRGANO DE SELECCIÓN**

CRITERIOS: Explicitados na memoria
ORGANO: Minia Manteiga e Carlos Dafonte (investigadores principais do proxecto)

**Documentación que debe presentar o solicitante:**

1. Fotocopia D.N.I.
2. Fotocopia da titulación académica.
3. Curriculum vitae.

\*Acordo sobre clasificación profesional, condicións de traballo e marco de referencia da representatividade laboral do persoal contratado con cargo a proxectos e convenios de I+D+I financiados a través do capítulo VI dos orzamentos da UDC. Asinado o 17/11/2014 entre a UDC e os sindicados CCOO e CIG

LUGAR DE PUBLICACIÓN DA RESOLUCIÓN DO CONTRATO: tablón de anuncios do Departamento de Tecnoloxías de la Información y Comunicaciones, UADI de la Facultade de Informática

A Coruña/Ferrol, 21 de 10 de 2015

O/A INVESTIGADOR/A PRINCIPAL

Asdo.: Minia Manteiga Outeiro

MANTEIGA  
OUTEIRO  
MINIA - DNI  
78781147Y

Firmado digitalmente por  
MANTEIGA OUTEIRO MINIA - DNI  
78781147Y  
Nombre de reconocimiento (DN):  
c=ES, o=UNIVERSIDAD DE A  
CORUÑA, ou=certificado electrónico  
de empleado público, ou=CIENCIAS  
DA NAVEGACION E DA TERRA,  
ou=5083, serialNumber=78781147Y,  
sn=MANTEIGA OUTEIRO,  
givenName=MINIA, cn=MANTEIGA  
OUTEIRO MINIA - DNI 78781147Y  
Fecha: 2015.10.23 12:43:40 +02'00'

**Documentación que debe presentar o solicitante:**

1. Fotocopia D.N.I.
2. Fotocopia da titulación académica.
3. Curriculum vitae.

*\*Acordo sobre clasificación profesional, condicións de traballo e marco de referencia da representatividade laboral do persoal contratado con cargo a proxectos e convenios de I+D+I financiados a través do capítulo VI dos orzamentos da UDC. Asinado o 17/11/2014 entre a UDC e os sindicatos CCOO e CIG*

## MEMORIA ANEXA A LA CONVOCATORIA DE CONTRATO

**REFERENCIA:** Proxecto MINECO ESP2014-55996-C2-2-R

**TÍTULO:** Técnicas de I.A. para a explotación do catálogo de Gaia: Análise, validación e visualización.

**CENTRO DE TRABAJO:** FACULTAD DE INFORMÁTICA. UDC.

**OBJETO:** DISEÑO Y DESENVOLVEMENTO DE APLICACIÓNS PARA A ANÁLISE DE DATOS EN ENTORNOS MULTIPLATAFORMA. TÉCNICAS DE AGRUPAMENTO (IA I ESTATÍSTICA). DESENVOLVEMENTOS EN JAVA E MATLAB NO ÁMBITO DA ASTROFÍSICA COMPUTACIONAL (PROXECTO Gaia da ESA)

**INVESTIGADORES PRINCIPAIS:** MINIA MANTEIGA OUTEIRO (ETS NÁUTICA E MÁQUINAS) E CARLOS DAFONTE VÁZQUEZ (FACULTAD DE INFORMÁTICA)

**CORREO ELECTRÓNICO DE CONTACTO:** dafonte@udc.es

### DETALLES DA CONVOCATORIA

#### **TAREFAS A REALIZAR:**

- Deseño de ferramentas de análise de datos en entornos multiplataforma.
- Desenvolvemento de algoritmos de clasificación automática mediante agrupamento (SVM, redes SOM).
- Implementación de sistemas de autenticación e cifrado para o acceso ós datos (ubicados en centros externos ESAC e CNES) para a execución dos algoritmos tanto el local como en remoto (integrado no pipeline de cómputo).
- Xeneración de toda a documentación do proxecto asociada e manuais de usuario.

**XUSTIFICACIÓN DA DURACIÓN DEL CONTRATO:** Apoio ás tarefas de cómputo intensivo de datos reflexadas no proxecto ESP2014-55996-C2-2-R durante o tempo de vixencia (probablemente se poida prorrogar o contrato outros 10 meses; estamos á espera dunha modificación do orzamento solicitada ó MINECO)

#### **REQUISITOS DOS CANDIDATOS:**

**Titulación:** Enxeñeiro en Informática

**Experiencia:** Experiencia en técnicas de I.A. para minería de datos e visualización.

Como investigador en formación, o candidato debe estra matriculado nun programa de doutoramento do SUG.

#### **CRITERIOS DE VALORACIÓN DESGLOSADOS**

- Expediente académico: 20%
- Experiencia laboral en investigación relacionada 30%
- Coñecementos en Java e deseño web 15%
- Experiencia en sistemas de autenticación non convencionáis 20%
- Coñecementos de inglés 10%
- Outros méritos: 5%

## DESCRIPCIÓN DEL PROYECTO EN EL QUE SE ENMARCA EL TRABAJO A REALIZAR:

### ESP2014-55996-C2-2-R

Nuestro grupo de investigación participa desde el año 2006 en la preparación científica de la misión GAIA de la Agencia Espacial Europea, proyecto espacial clave que ha sido puesto en órbita a finales de 2013 y cuyo objetivo fundamental es obtener un “censo estereoscópico de la Galaxia” hasta magnitud 20. Para ello se medirán de forma precisa posiciones y distancias a las estrellas, velocidades radiales (en el caso de los objetos más brillantes) y propiedades físico-químicas, de un total de aproximadamente mil millones de objetos. Formamos parte del equipo científico de GAIA a través de la unidad de coordinación CU8 “Clasificación y Parámetros Astrofísicos” en el marco del Gaia *Data Processing and Analysis Consortium* (DPAC), que es el consorcio científico encargado de diseñar, implementar, gestionar y ejecutar la reducción de datos de la misión hasta la producción del catálogo final.

CU8 tiene como misión la preparación de herramientas de análisis para la clasificación de todas las fuentes que observará el satélite en clases discretas (estrellas, binarias, galaxias, cuásares, objetos del sistema solar, y fuentes compuestas) junto con una estimación de los principales parámetros astrofísicos que las caracterizan. En estos momentos el satélite está operativo, tomando datos de las fuentes de la galaxia, y en los próximos meses tendrá lugar el procesado de los primeros conjuntos de datos de la misión.

Participamos en tres paquetes de trabajo de CU8: GWP-S-823 cuyo objetivo es la parametrización físico-química de estrellas a partir de datos del instrumento RVS; GWP-S-836 (que lideramos) encargado del análisis general de los *outliers* de la misión; y GWP-S-824 que realizará una clasificación, no supervisada, de todas las fuentes observadas.

Cada seis meses aproximadamente se realizan workshops internos del CU8, en los que se presentan los avances y resultados obtenidos tanto de tipo científico (funcionamiento de los algoritmos de clasificación) como técnico (implementación del software y documentación asociada). Por ello, además de artículos en revistas y presentaciones en congresos, se elaboran también documentos técnicos, que son informes internos de CU8 que reflejan el trabajo realizado.

A partir del año 2013 nuestro grupo también se ha involucrado de forma significativa en la unidad de coordinación 9, CU9, encargada de elaborar una propuesta del archivo final de Gaia. En concreto participamos en los paquetes de trabajo que se encargan de la validación de los datos del catálogo (WP 945), Agrupamiento y selección avanzada de datos para visualización multidimensional (WP 985), Herramientas de acceso avanzado y Data Mining (WP 973) y actividades de divulgación (WP963).

#### **Paquete de trabajo CU8-GWP-S-823 “Generalized Stellar Parametrizer-Spectroscopy.**

Este paquete de trabajo tiene como misión elaborar herramientas de análisis automático para la extracción de los principales parámetros físico-químicos de las estrellas observadas por el espectrógrafo RVS, de resolución intermedia y rango espectral de alrededor del triplete IR del Call. Como continuación del trabajo que se ha venido realizando en los últimos años, hemos elaborado una aplicación informática de análisis basada en redes de neuronas artificiales y en el procesado de los espectros sintéticos de RVS en los dominios transformados de Fourier y Wavelet, que es capaz de extraer la temperatura y gravedad efectiva, la metalicidad y una estimación de la abundancia de los elementos  $\alpha$ , para espectros con baja relación señal a ruido de estrellas de tipos F, G y K, las más abundantes en la Galaxia.

#### **Paquete de Trabajo CU8-GWP-S-836, dedicado al “Análisis de Outliers”.**

Somos responsables ante la ESA del paquete de trabajo de análisis de “ouliers”. El objetivo de este paquete es analizar todas aquellas observaciones atípicas o dispersas que los programas de clasificación supervisada de CU8-Gaia no sea capaz de clasificar con un grado de probabilidad aceptable. Nuestra misión consiste en analizar la información sobre los mismos, intentar agruparlos en clases mediante el uso de técnicas estadísticas (análisis cluster) y de Inteligencia Artificial (redes de neuronas, sistemas basados en el

conocimiento), comparar sus propiedades con la información de otros *surveys*, y detectar posibles nuevos tipos de objetos.

El algoritmo propuesto combina métodos de análisis cluster (Fuzzy C-means, FCM) con modelos de redes de neuronas artificiales competitivas (redes SOM). Así, el mapa resultado a partir de las redes SOM se segmenta mediante análisis cluster (FCM) obteniendo subconjuntos de objetos con propiedades similares cuya naturaleza física se debe estudiar por cruce con otros catálogos e introduciendo conocimiento de expertos.

### Paquetes de trabajo de CU9

Una de las tareas primordiales antes de la publicación de cada versión del catálogo es la validación de los datos que contiene. Nuestro grupo participa en varios aspectos clave en CU9. Actualmente, estamos desarrollando un conjunto de herramientas estadísticas que permiten la caracterización y validación del contenido del catálogo (objetivo 042); estas herramientas deberán permitir un análisis automático del archivo completo para producir informes estadísticos. Estos informes serán usados como parte de la documentación oficial del catálogo (objetivo 041) y constituirán el primer paso en la validación de su contenido. También participaremos en varios aspectos específicos de la validación del catálogo: el análisis de sus datos astrométricos y fotométricos (objetivo 043) y el uso de clusters abiertos para validaciones a pequeña escala (objetivo 044). Estas tareas precisan ser acompañadas del desarrollo de herramientas software que realicen el análisis de los datos del archivo y la aplicación de nuestra experiencia y conocimiento en la interpretación de los resultados.

Minería de datos: El Sistema de archivo en ESAC proporcionará soporte hardware y software para actividades de minería de datos. La implementación de esta infraestructura no está planificada hasta la tercera versión del catálogo (previsto para 2017) pero requiere de varias actividades preparatorias para poder estar en marcha a tiempo. Nuestro equipo es el responsable de estas actividades preparatorias a través de principalmente dos tareas:

- La implementación de un *testbed* para el desarrollo de procedimiento de minería de datos (objetivo 050). Se trata de un sistema prototipo en el que las aplicaciones de minería serán instaladas y ejecutadas. Esto se realizará utilizando las facilidades disponibles en el CESCA (Centro de Supercomputación de Cataluña), que da soporte al proyecto GENIUS.
- Se desarrollará un *framework* para minería de datos que deberá estar disponible en el entorno de producción en ESAC (objetivo 049). Este *framework* será la base para el desarrollo de aplicaciones científicas sobre el archivo de Gaia.

A Coruña, 21/10/2015

Asdo.: Minia Manteiga Outeiro

**MANTEIGA  
OUTEIRO  
MINIA - DNI  
78781147Y**

Firmado digitalmente por MANTEIGA  
OUTEIRO MINIA - DNI 78781147Y  
Nombre de reconocimiento (DN):  
c=ES, o=UNIVERSIDAD DE A  
CORUÑA, ou=certificado electrónico  
de empleado público, ou=CIENCIAS  
DA NAVEGACION E DA TERRA,  
ou=5083, serialNumber=78781147Y,  
sn=MANTEIGA OUTEIRO,  
givenName=MINIA, cn=MANTEIGA  
OUTEIRO MINIA - DNI 78781147Y  
Fecha: 2015.10.23 12:35:38 +02'00'