Firmante: Universidade da Coruña - Q6550005J URL de validación: https://sede.udc.gal/services/documents_validation CSV: U8P8VGGGK4UBA9GSQTT9EVA1
Este documento es copia auténtica de un documento custodiado por la Universidade da Coruña. Fecha de creación: 13/02/2017 14:17:57 Página 1 de un total de 5 página(s).

UNIVERSIDADE DA CORUÑA

GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

ANEXO III

REF.: 2017/CP/050

CONVOCATORIA DE CONTRATACIÓN

AXUDA

Referencia: Proxecto MINECO ESP2016-80079-C2-2-R Título: Minería de Datos de Gaia para estudiar la Vía Láctea			
CENTRO: Facultade de Informática, Campus de Elviña			
OBXECTO DO CONTRATO: DESENVOLVEMENTO DE APLICACIÓNS JAVA EN ENTORNOS DE CÓMPUTO INTENSIVO NO CAMPO DA ASTROFÍSICA COMPUTACIONAL (PROXECTO Gaia da ESA) [I] Colaborador/a nas tarefas do (proxecto/convenio) de investigación que se indica.			
LUGAR E HORARIO: Laboratorio 3.2. Grupo de Investigación LIA[2]. De luns a venres de 9h-14h e 16h-18:30h			
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 ☐ * Investigador en formación ☒ * Require matrícula programa doutoramento SUG Técnico de apoio á investigación ☐ Axudante de apoio á investigación ☐ Técnico administrativo ☐ Axudante administrativo ☐			
XORNADA	Tempo Completo: ☒	Tempo Parcial: ☐	
DURACIÓN	18 meses prorrogables	DATA APROX. DE INICIO	01/4/2017
RETRIBUCIÓNS BRUTAS : 1368,5 (incluído prorrateo de pagas extras)			
PARTIDA ORZAMENTARIA: 2017-6140210573-541A64902-00 (Ref: ESP2016-80079-C2-2-R)			

* 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 memoria da convocatoria
REQUISITOS: O candidato debe estar 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 desde a publicación no taboleiro de anuncios da sede electrónica da UDC. Os requisitos e méritos a valorar deberán terse cumprido e acreditado cando remate o prazo de presentación de instancias para participar na convocatoria.

ÓRGANO DE SELECCIÓN E PUBLICACIÓN RESOLUCIÓN

ORGANO: Carlos Dafonte e Minia Manteiga (investigadores principais do proxecto)
LUGAR DE PUBLICACIÓN DA RESOLUCIÓN: Taboleiro de anuncios do Departamento de Tecnoloxías da Información e as Comunicacións (UADI, Facultade de Informática)

A Coruña, 13 de febreiro de 2017

O/A INVESTIGADOR/A PRINCIPAL Asdo.: Carlos Dafonte

Documentación que debe presentar o solicitante:

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

NOMBRE DAFONTE VAZQUEZ JOSE CARLOS - NIF 76620329S

Firmado digitalmente por NOMBRE DAFONTE VAZQUEZ JOSE CARLOS - NIF 76620329S
Nombre de reconocimiento (DN): c=ES, o=FNMT, ou=FNMT Clase 2 CA, ou=500052865, cn=NOMBRE DAFONTE VAZQUEZ JOSE CARLOS - NIF 76620329S
Fecha: 2017.02.13 11:00:47 +01'00'

**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*

MEMORIA ANEXA A LA CONVOCATORIA DE CONTRATO

REFERENCIA: Proxecto ESP2016-80079-C2-2-R

TÍTULO: Minería de Datos de Gaia para estudiar la Vía Láctea

CENTRO DE TRABALLO: FACULTADE DE INFORMÁTICA. UDC.

OBJETO: DESENVOLVEMENTO DE APLICACIÓNS JAVA EN ENTORNOS DE CÓMPUTO INTENSIVO NO CAMPO DA ASTROFÍSICA COMPUTACIONAL (PROXECTO Gaia da ESA) [I]

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

CORREO ELECTRÓNICO DE CONTACTO: dafonte@udc.es

DETALLES DA CONVOCATORIA

TAREFAS A REALIZAR:

- Optimización de algoritmos de clasificación non supervisada en entornos de computación distribuida mediante Apache-Spark.
- Optimización de algoritmos de clasificación non supervisada en entornos de computación distribuida mediante cálculo con GPUs utilizando a tecnoloxía NVidia CUDA.
- Integración de algoritmos de agrupamento no entorno de cálculo proporcionado por ESA-ESAC para o proxecto ESA-Gaia (GSP-Spec)
- Desenvolvemento de aplicacións web para a visualización e análise dos resultados de clasificación.
- Deseño da integración (entrada e saída de datos) para o entorno de clasificación mediante tecnoloxías Virtual Observatory (VO).
- 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 ESP2016-80079-C2-2-R durante o tempo de vixencia (probablemente se poida prorrogar o contrato algúns meses máis)

REQUISITOS DOS CANDIDATOS:

Titulación: Enxeñeiro en Informática

Experiencia: Experiencia en optimización de cómputo intensivo e visualización.

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

Firmante: Universidade da Coruña - Q6550005J URL de validación: https://sede.udc.gal/services/documents_validation CSV: U8P8VGGGK4UBA9GSQTT9EVA1
Este documento es copia auténtica de un documento custodiado por la Universidade da Coruña. Fecha de creación: 13/02/2017 14:17:57 Página 2 de un total de 5 página(s).

CRITERIOS DE VALORACIÓN DESGLOSADOS (sobre 100 puntos)

- Expediente académico **(10 puntos)**
    - Utilizarase a media do Expediente en Base 10 (RD 1125/2003)

- Experiencia laboral en investigación relacionada **(40 puntos)**
    - Participación en proxectos de investigación (Autonómico / Nacional / Europeo). (5 puntos cada 6 meses).
    - Publicacións indexadas JCR no ámbito do traballo (2-5 puntos segundo sexa Q1-Q4).
    - Outras publicacións non indexadas (máximo 1 punto)

- Coñecemento de Java e desenvolvemento de aplicacións web **(10 puntos)**
    - Cursos dun mínimo de 15 horas (2 puntos)
    - Bolsas relacionadas (1 punto/mes)

- Coñecementos en deseño de aplicacións móbiles **(10 puntos)**
    - Cursos dun mínimo de 15 horas (2 puntos)
    - Bolsas relacionadas (1 punto/mes)

- Coñecementos en optimización de cómputo con CUDA **(10 puntos)**
    - Cursos dun mínimo de 15 horas (2 puntos)
    - Bolsas relacionadas (1 punto/mes)

- Experiencia en técnicas de IA aplicada **(15 puntos)**
    - Participación en contratos (5 puntos)
    - Patentes e rexistros de software (5 puntos)
    - Bolsas relacionadas (1 punto/mes)

- Coñecementos de inglés **(5 puntos)**
    - Certificaciones oficiales y equivalentes A2 -> C2 (puntuación 1-5 puntos segundo o nivel)

Os proxectos fin de carreira ou de mestrado tamén se valorarán nos diferentes apartados (hasta 10 puntos) segundo a calificación obtida.

Firmante: Universidade da Coruña - Q6550005J URL de validación: https://sede.udc.gal/services/documents_validation CSV: U8P8VGGGK4UBA9GSQTT9EVA1
Este documento es copia auténtica de un documento custodiado por la Universidade da Coruña. Fecha de creación: 13/02/2017 14:17:57 Página 3 de un total de 5 página(s).

Firmante: Universidade da Coruña - Q6550005J URL de validación: https://sede.udc.gal/services/documents_validation CSV: U8P8VGGGK4UBA9GSQTT9EVA1
Este documento es copia auténtica de un documento custodiado por la Universidade da Coruña. Fecha de creación: 13/02/2017 14:17:57 Página 4 de un total de 5 página(s).

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

ESP2016-80079-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 18. 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 fisico-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). También colaboramos desde 2015 en el paquete de documentación de la *Data Release* (también en CU9) y en un paquete de trabajo en CU4, relacionado con el estudio de las nebulosas planetarias.

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 CaII. 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 α, 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 diversas técnicas pero se basa principalmente en modelos de redes de neuronas artificiales competitivas (redes SOM). Así, el mapa resultado a partir de las redes SOM permite separar subconjuntos de objetos con propiedades similares cuya naturaleza física se debe estudiar por cruce con otros catálogos y facilita la incorporación de conocimiento de experto. Uno de los retos más importantes en estos momentos es conseguir el que el cómputo sea lo suficientemente rápido para ejecutarse en SAGA-Hadoop.

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 (objetivos O17, O112, O113, O21 y O22); 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 (objetivos O111, O114 y O21) 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 (O213) 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: Tanto el Centro de Supercomputación de Cataluña como el GAC (Gaia Archive System) instalado en ESAC (que da soporte a todas las *Data Releases* de Gaia*)* proporcionará soporte hardware y software para actividades de minería de datos. La implementación de esta infraestructura está parcialmente implantada y todavía requiere infraestructura software y hardware para posibilitar la explotación científica de los datos. Nuestro equipo es el responsable de estas actividades preparatorias a través de principalmente tres 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 europeo 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.
- El desarrollo de una aplicación compatible V.O. que facilite en análisis de los mapas entrenados en la reléase del catálogo y que permita también la integración con herramientas disponibles en astrofísica computacional (por ejemplo, Topcat y Vaex).

NOMBRE DAFONTE VAZQUEZ JOSE CARLOS - NIF 76620329S

Firmado digitalmente por NOMBRE DAFONTE VAZQUEZ JOSE CARLOS - NIF 76620329S
Nombre de reconocimiento (DN): c=ES, o=FNMT, ou=FNMT Clase 2 CA, ou=500052865, cn=NOMBRE DAFONTE VAZQUEZ JOSE CARLOS - NIF 76620329S
Fecha: 2017.02.13 10:58:58 +01'00'

A Coruña, 13/2/2017

Asdo.: José Carlos Dafonte Vázquez

Firmante: Universidade da Coruña - Q6550005J URL de validación: https://sede.udc.gal/services/documents_validation CSV: U8P8VGGGK4UBA9GSQTT9EVA1
Este documento es copia auténtica de un documento custodiado por la Universidade da Coruña. Fecha de creación: 13/02/2017 14:17:57 Página 5 de un total de 5 página(s).