



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o
D\

**INSTRUCCIONES PARA COMPLETAR LA MEMORIA CIENTÍFICO-TÉCNICA DE
PROYECTOS DE I+D+i “RETOS INVESTIGACIÓN” DEL PROGRAMA ESTATAL DE I+D+i
ORIENTADA A LOS RETOS DE LA SOCIEDAD**

Lea detenidamente estas instrucciones para completar correctamente la memoria-científico técnica.

1. Este modelo de memoria está restringido en su extensión máxima y por consiguiente ha de limitarse a los espacios indicados al completarla.
2. Las memorias pueden completarse en español o en inglés, a excepción del apartado 1. RESUMEN DE LA PROPUESTA, que debe completarse en ambos idiomas.
3. Se recomienda completar la memoria empleando un pc con sistema operativo Windows y usando como procesador de textos MS Word (MS Office).
4. Para completar los textos, sitúe el cursor en las zonas sombreadas. 4000 caracteres son, aproximadamente, una página.
5. Una vez terminada la memoria, guarde su archivo en formato pdf (de no más de 4Mb) y apórtelo a la solicitud telemática del proyecto en el apartado “Añadir documentos”.
6. Debido a que este formulario está diseñado para incluir únicamente texto con un tipo de letra determinado, si necesita incluir fórmulas, reacciones químicas, fórmulas matemáticas, etc., o figuras aclarativas, deberá hacerlo en los anexos I y II, respectivamente, tras citarlas en el cuerpo del texto. No deberá emplear más extensión que la indicada.
7. El formulario está adaptado para poder emplear la opción de “copiar y pegar”.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o B)

AVISO IMPORTANTE

En virtud del artículo 11 de la convocatoria **NO SE ACEPTARÁN NI SERÁN SUBSANABLES MEMORIAS CIENTÍFICO-TÉCNICAS** que no se presenten en este formato.

1. RESUMEN DE LA PROPUESTA / SUMMARY OF THE PROPOSAL

(Debe rellenarse también en inglés / It should also be completed in English)

INVESTIGADOR PRINCIPAL 1 (Nombre y apellidos):

ANTONIO HERNÁNDEZ PÉREZ

INVESTIGADOR PRINCIPAL 2 (Nombre y apellidos):

J. Tomás Nogales Flores

TÍTULO DEL PROYECTO:

CURATOR-e: custodia y gestión digital para la reutilización de datos abiertos de investigación en Humanidades y Ciencias Sociales

ACRÓNIMO:

CURATOR-e

TITLE OF THE PROJECT:

CURATOR-e: curation and management for the reuse of research open datasets in the field of Humanities and Social Sciences

ACRONYM:

CURATOR-e



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

DA

RESUMEN

Debe contener los aspectos más relevantes, los objetivos propuestos y los resultados esperados.

El resumen del proyecto debe completarse también en la aplicación de la solicitud. Su contenido podrá ser publicado a efectos de difusión si el proyecto resultara financiado en esta convocatoria.

Máximo 3500 caracteres

La ciencia enfrenta desde hace unos años un proceso acelerado de cambio de paradigma, la e-ciencia, o ciencia mejorada (enhanced science) o lo que Jim Gray llamó el cuarto paradigma. Lo que en inglés se conoce como Big Data o Data Intensive Science, un modo de hacer ciencia basado en la gran cantidad de datos existentes. La posibilidad de analizarlos, compartirlos y reutilizarlos será clave para la competencia, la productividad, la innovación y el avance del conocimiento.

Los principales organismos y agencias que financian la investigación, incluidas la Comisión Europea y la National Science Foundation (EEUU.), por citar las más relevantes, recomiendan o incluso exigen que todos los proyectos financiados incluyan un plan de gestión de los datos que se generan durante la investigación. Estos planes deben incluir, entre otros requisitos, las políticas de acceso y compartición de los datos, los tipos y formatos de datos utilizados, las políticas para su reutilización, los mecanismos de interoperabilidad y los planes de preservación.

Este proyecto parte de la hipótesis de que cada universidad o centro investigador necesitará, a corto o medio plazo, de un repositorio para el almacenamiento de grandes volúmenes de datos que se generan durante una investigación. Estos datos que ahora se producen y recopilan en grandes cantidades en formato digital favorecen el desarrollo de nuevas investigaciones, permitiendo nuevos enfoques en el análisis de datos, y pueden servir de gran ayuda en la formación de jóvenes investigadores y en la docencia. Los repositorios de datos que se proponen tienen unos objetivos y filosofía diferente a los repositorios institucionales o temáticos de resultados de la investigación que ahora existen y cuyos contenidos están fundamentalmente en forma de publicaciones.

El proyecto tiene dos objetivos fundamentales:

1) Crear un observatorio sobre gestión de datos de investigación. El observatorio tiene como fin: ayudar a bibliotecas universitarias, centros de investigación y gestores de políticas científicas a tomar decisiones bien informadas en base a lo que se está haciendo en cada momento en todo el mundo respecto a los repositorios de datos de investigación; orientarles sobre cómo diseñar planes de depósito, gestión y preservación de datos de investigación; y ayudar a concienciar a los investigadores, bibliotecarios, y demás actores de la comunicación científica de la importancia y los beneficios de compartir los datos de investigación.

2) Crear un repositorio piloto de datos de investigación para la experimentación. El objetivo del repositorio será facilitar el registro y depósito de datos, su descubrimiento, acceso y reutilización y contribuir a las políticas de depósito en acceso libre de datos de investigación. El repositorio piloto albergará datos de investigación de dos fuentes principales:

a) De los grupos de investigación españoles reconocidos por distintos organismos oficiales de las áreas de las ciencias sociales y las humanidades (con datos de encuestas, estadísticas, análisis de muestras arqueológicas, históricas o de otro tipo, corpus lingüísticos, objetos multimedia, etc.)

b) De los grupos de investigación de todas las áreas temáticas con proyectos aprobados en el Prog. Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprog de Generación de Conocimiento de esta convocatoria. Se hará en el último año, con el fin de experimentar con todo tipo de datos.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

PALABRAS CLAVE

Máximo 200 caracteres

datos de investigación; datos abiertos; repositorios; e-ciencia; metadatos; skos; biblioteca digital; humanidades; ciencias sociales; reutilización de información del sector público

SUMMARY

It should contain the most relevant topics of the project, the objectives and the expected results.

The summary should also be completed in the electronic application. It could be published for diffusion purposes if the project is financed in this call.

Maximum 3500 characters

Science faces a paradigm shift, e-science, enhanced science or what Jim Gray called the fourth paradigm. What in English is known as Big Data and Data Intensive Science, a way to make science based on the large amount of existing scientific data. The ability to analyze, share and reuse data will be key to competition, productivity, innovation and the advancement of knowledge.

Main research agencies and funding organizations, including European Commission and the National Science Foundation (USA), recommend or even require, that all funded projects include a plan for data management to archive and preserve all created data during the research project lifetime. These plans must support data access, sharing and reuse. These plans must also incorporate the use of standards to support data interoperability.

This project proposal starts from the assumption that every University or research center will need, in the short to medium term, a repository to archive valuable research data throughout the research lifecycle and for future research efforts. Research data, produced and gathered in digital format, contribute to the development of new research, fostering new approaches for data analysis and might be useful for young researchers' training and education. Data repositories which are proposed here have a different aim and philosophy from current institutional or subject-oriented repositories intended to keep research results, based upon papers or formal publications.

The project has two main objectives:

- 1) Create an observatory or research data portal. Its goals are: to guide university libraries, research centers and policy makers to take informed decisions based on what others are doing worldwide on research data repositories; to serve as a reference for designing deposit plans, managing research data and their preservation; and helping to raise awareness among researchers, librarians, and others involved actors in scientific communication about the importance and benefits of sharing research data.
- 2) Create a pilot research data repository on which to experiment with this type of data. The aim of this repository will be to facilitate data registry and deposit, as well as their discovery, access and reuse, contributing to create the deposit policies in the open access framework to research data. The pilot repository will store and treat research data from two main sources:
 - a) From research groups that are recognized by different Spanish official bodies in the Social Sciences and Humanities (which count with research data of different kinds: surveys, statistics, analysis of archaeological or historical samples, linguistic corpora, multimedia objects, etc)



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o B)

b) From research groups in all subject areas, including Social Sciences and Humanities but also hard sciences such as biology, chemistry, biomedicine or physics, with projects approved in this Call for proposals (Program of Promotion of Scientific and Technical Research Excellence, 2013-2016). The goal of including this source is to experiment with different types of data. The collection of these data will take effect in the last year of the project.

KEY WORDS

Maximum 200 characters

research data; open data; repositories; e-science; metadata; skos; digital library; humanities; social science; public sector information reuse



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

2. INTRODUCCIÓN

2.1. Antecedentes y estado actual de los conocimientos científico-técnicos de la materia específica del proyecto, incluyendo, en su caso, los **resultados previos** del equipo investigador, de otros grupos que trabajen en la misma y la relación, en su caso, entre el grupo solicitante y otros grupos de investigación nacionales y extranjeros;

- si el proyecto es continuación de otro previamente financiado, deben indicarse con claridad los objetivos y los resultados alcanzados de manera que sea posible evaluar el avance real que se propone en el proyecto solicitado.
- si el proyecto aborda un nuevo tema, deben indicarse los antecedentes y contribuciones previas del equipo investigador con el fin de justificar su capacidad para llevar a cabo el nuevo proyecto.

Máximo 16 000 caracteres

La mayor parte de las universidades y centros de investigación de todo el mundo, incluido nuestro país, disponen ya de repositorios institucionales que almacenan los resultados de investigación de sus miembros (principalmente artículos, comunicaciones a congresos y tesis doctorales). Pero además, en un futuro próximo será imprescindible, tal y como recomienda la OCDE (Organisation for Economic Cooperation and Development, 2007), la Comisión Europea (e-IRG e-Infrastructure Reflection Group, 2009 y https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/Report_2013-07-OpenResearchData-Consultation-FINAL1.pdf) o la NSF (National Science Foundation, 2011) crear nuevos repositorios de datos de investigación, que contengan las colecciones de datos (datasets) en los que se ha basado cada investigación.

Esta misma convocatoria establece en su artículo 8, apartado 2a "Proyectos que impliquen la realización de estudios mediante encuestas cuantitativas en el ámbito de las ciencias sociales con toma de datos" que "los investigadores principales de estos proyectos deberán transferir, en un plazo máximo de doce meses desde su finalización, los microdatos integrantes del estudio al Banco de Datos Específico de Estudios Sociales que, a estos efectos, radica en el Centro de Investigaciones Sociológicas, en cumplimiento de los artículos 37.1 y 37.2 de la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación, la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, y la Recomendación de la Comisión Europea, de 17 de julio de 2012, relativa al acceso a la información científica y su preservación".

Así pues, la tendencia para que los investigadores compartan sus datos de investigación y elaboren planes para su gestión y preservación irá en aumento. Y es que la capacidad de los ordenadores para capturar datos de instrumentos de medición, de simulaciones, de corpus textuales, de sensores, de redes, etc., ofrece la posibilidad de generar grandes colecciones de datos que si estuviesen bien almacenados, bien estructurados y fuesen fáciles de acceder, permitirían no sólo validar los resultados de la investigación que los han generado, sino además que otros investigadores pudieran reutilizarlos para agregarlos a otros conjuntos de datos o para realizar otros análisis, no previstos en un principio por los autores originales que recopilaron esos datos.

En España, la Red de Bibliotecas Universitarias (REBIUN) en colaboración con la FECYT (Grupo de Trabajo de "Depósito y Gestión de datos en Acceso Abierto" del proyecto RECOLECTA, 2012) elaboraron en diciembre de 2012 un informe en el que se reflejan todos los aspectos importantes que intervienen en la gestión de los datos, desde su definición, tipos de datos, actores implicados, buenas prácticas para la gestión y un panorama general de la situación en España. En nuestro país contamos aún con muy poca experiencia en este tipo de repositorios. Si bien se pueden mencionar algunos en el campo de las ciencias puras (Wulff-Barreiro, 2011; Arano, 2011), en el campo de las ciencias sociales y las humanidades apenas se puede mencionar el trabajo de la Fundación Juan March, la biblioteca de datos de CEACS, compuesta de datos primarios y secundarios de encuestas y estadísticas procedentes del CIS, INE, la



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

OCDE y otros organismos e integrada en la red Dataverse de la Universidad de Harvard. A ello se suma la colección “Recursos i dades primàries”, integrada en el Repositorio Digital de la UPF. A nivel internacional, el repositorio en ciencias sociales más importante es del Interuniversity Consortium for Political and Social Research (ICPSR) de Estados Unidos.

El objetivo final de CURATOR-e es doble. Por una parte, crear un observatorio en el que se recoja toda la información nacional e internacional sobre políticas y planes de gestión y preservación de datos de investigación así como estándares y repositorios que sirva de guía a bibliotecas universitarias y centros de investigación para contribuir a la toma de decisiones informadas sobre cómo afrontar la gestión de este tipo de datos. El proyecto estará centrado en el campo de las Humanidades y las Ciencias Sociales. Por otra parte, crear un repositorio piloto para que investigadores de Humanidades y Ciencias Sociales dispongan de un espacio digital en el que depositar colecciones de datos de investigación y de mecanismos para descubrir y reutilizar esos conjuntos de datos para sus labores de investigación y docencia, lo que permitirá acumular conocimiento y experiencia sobre la problemática de la puesta en marcha y gestión de estos repositorios.

En el mundo de la investigación, la conciencia sobre el valor de compartir datos para hacer avanzar más rápidamente la ciencia, la evaluación constante del trabajo del investigador así como la facilidad para acceder, difundir, procesar y preservar grandes cantidades de datos, ha propiciado el auge del movimiento de acceso abierto (Open Access - OA). Un resumen de la historia del movimiento OA y una amplia bibliografía sobre el tema se puede encontrar en Suber (2009) y Carroll (2011).

Los estados y las principales instituciones financiadoras de la investigación han fomentado el desarrollo de los repositorios institucionales, conscientes de que para avanzar e innovar, el conocimiento y los datos que se generan durante una investigación deben ser compartidos y reutilizados, con el fin de no tener que volver a repetir experimentos o para tener una base de datos de experimentación mucho más amplia.

A partir de 2002, fecha de la Budapest Open Access Initiative, muchos investigadores están obligados por mandato a depositar sus resultados de investigación, principalmente artículos y comunicaciones, en repositorios institucionales o temáticos. En el Registry of Open Access Repositories Mandatory Archiving Policies, ROARMAP (<http://roarmap.eprints.org/>), se puede acceder al más extenso listado de políticas de mandatos de depósito de archivos en repositorios de acceso abierto.

En España, el artículo 37 de la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación se refiere expresamente a la “difusión en acceso abierto”. El punto 1 dice que: “Los agentes públicos del Sistema Español de Ciencia, Tecnología e Innovación impulsarán el desarrollo de repositorios, propios o compartidos, de acceso abierto a las publicaciones de su personal de investigación, y establecerán sistemas que permitan conectarlos con iniciativas similares de ámbito nacional e internacional”. Y el punto 2 obliga a los investigadores a hacer públicas las publicaciones que realicen no más tarde de doce meses.

La existencia de estas leyes, mandatos y recomendaciones por parte de las autoridades públicas y de las agencias financiadoras de la investigación para que los investigadores publiquen los resultados de sus investigaciones y los datos obtenidos durante la misma ha provocado la creación de repositorios institucionales en prácticamente todas las universidades españolas, europeas, norteamericanas y en el resto de países más desarrollados. En noviembre de 2013, el Directorio de Repositorios de Acceso Abierto (OpenDOAR) recoge en España 109 repositorios de los más de 2.525 que existen en todo el mundo.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

De acuerdo a OpenDOAR (<http://www.opendoar.org>), los repositorios más activos son, entre otros, los de la Universidad de Salamanca, la Universidad Autónoma de Barcelona, el del Consejo Superior de Investigaciones Científicas (CSIC), la Universidad de Granada, la UNED, la Universidad Carlos III de Madrid y la Universidad Complutense. Hasta ahora, las autoridades obligaban o recomendaban la difusión y publicación de resultados de investigación en repositorios institucionales o temáticos. Ahora, conscientes de la importancia de las grandes colecciones de datos abiertos, organizaciones como la League of European Research Universities (LERU) obligan o recomiendan que además de las publicaciones se pongan a disposición de la comunidad científica también los datos recopilados durante la investigación (mediciones, observaciones, etc.) <http://www.leru.org/files/general/Open%20Access%20to%20Research%20Data-FINALdocx.pdf>

La National Science Foundation (NSF), de EEUU, obliga a que todos los proyectos que soliciten financiación presenten un plan de gestión de datos de investigación (National Science Foundation, 2011) y algo parecido hace la guía de presentación de solicitudes de proyectos de investigación para el 7º Programa Marco de la Unión Europea. Intentan con ello afrontar el reto que se ha venido a llamar de los “Big Data”.

La importancia de los Big Data quedó manifiesta en mayo de 2011, cuando los grandes medios de comunicación como The Economist (Schumpeter, 2011) o New York Times (Lohr, 2011), entre otros, se hacían eco de un informe publicado por la consultora McKinsey (Manyika, 2011) en el que se ponía de manifiesto que la cantidad de datos existentes en la actualidad y el análisis de grandes conjuntos de datos (Big Data) serán claves para la competencia, la productividad, la innovación y el conocimiento. El informe habla de la necesidad de más de 150.000 gestores de datos solo en EEUU, con capacidad de análisis estadístico, matemático y con conocimientos de tecnologías de datos.

Por tanto, previsiblemente, a medio plazo las universidades y centros de investigación, además de contar con el repositorio institucional que ya está funcionando en la mayoría de estos organismos deberán poner en marcha un repositorio de datos de investigación. El hecho de que sean repositorios independientes no significa que no exista relación entre ellos ya que los datos estarán vinculados.

Los repositorios de datos de investigación sirven, entre otros fines, para validar resultados de investigación y, por tanto, deben estar vinculados de alguna manera a las publicaciones científicas en donde se muestra para qué fueron utilizados esos datos, por lo que algunos de los problemas se podrán abordar de forma conjunta, tanto para los repositorios institucionales como para los repositorios de datos de investigación.

Para la creación de este repositorio de datos de investigación será necesario realizar un inventario y registro de dónde se pueden encontrar dichos datos, y en qué formato se encuentran. Igual que ocurre con los repositorios institucionales, es necesario definir algunas políticas de puesta en valor del repositorio, por lo que se abordarán aspectos económicos, de costes y sostenibilidad de estos repositorios; legales, de licencias de acceso y uso; culturales, de hábitos de los investigadores; institucionales, de políticas de autoarchivo de estas colecciones de datos; y técnicos, de gestión de la tecnología, de interoperabilidad de los metadatos, etc.

Actualmente, el porcentaje de investigadores que hacen el depósito de sus publicaciones en repositorios institucionales oscilan entre el 10 y el 20% (Ginsparg, 2011; Björk et al., 2010). Algunas de las causas ya han sido identificadas (Björk, 2004; Creaser et al., 2010). Uno de los objetivos de este proyecto es lograr que aumente sustancialmente el número de artículos y comunicaciones de los repositorios institucionales en nuestro país, al mismo tiempo, que también se depositan datos de investigación.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

El repositorio piloto de datos de investigación que junto con el observatorio se pretende implementar tendrá en cuenta todas las iniciativas tendentes a solucionar tanto los problemas relativos a la identificación de autores (ORCID o ResearcherID), como a la de identificación persistente de recursos (DOI, Handle, etc.). Estos problemas, así como los relativos a la interoperabilidad semántica de los metadatos (Park, 2009), tienen que ver con la falta de calidad tanto de los metadatos como de los esquemas semánticos que deben facilitar el descubrimiento y la recuperación de artículos o conjuntos de datos de investigación (Greenberg et al., 2009).

Pretendemos que la exposición de datos se realice de acuerdo a los estándares abiertos internacionales más comunes en el campo y en particular mediante datos enlazados abiertos (Linked Open Data). El proyecto pretende buscar una solución a estos problemas empleando Linked Open Data, de forma parecida a cómo se está resolviendo en el Virtual Internet Authority File (VIAF).

En España, además de la actividad llevada a cabo por el grupo de trabajo de Fecyt y Rebiun, el proyecto más relevante en esta línea probablemente sea el Datasea, iniciado a principios de este año, que trabaja en la identificación de las políticas para incrementar la competitividad de la investigación basada en la gestión de datos. El grupo está integrado por investigadores de 5 universidades españolas (UPV, UV, UB, UCV y UM) y Odisea, registro internacional de repositorios de datos de investigación, resultado del mismo grupo de investigación (Gracia-Gracia, 2012). Aunque es posible mencionar también el grupo Tecnodoc de la UC3M o el TRL de la UPF y algún otro en USAL o la UNED.

En Europa es obligado mencionar los grupos que están trabajando en los proyectos OpenAIRE - DRIVER II, el proyecto SHARE, con conjuntos de datos sobre salud, demografía y jubilaciones o los grupos del programa Managing Research Data de JISC, que están trabajando más específicamente sobre repositorios de datos de investigación. En Estados Unidos, el repositorio más importante en ciencias sociales es sin duda el ya citado Interuniversity Consortium for Political and Social Research (ICPSR). Además, nuestro grupo Tecnodoc mantiene una estrecha relación con el grupo de investigación de Metadata Research Center, en la University of North Carolina-Chapel Hill, que ha puesto en marcha el proyecto DRYAD (Greenberg et al., 2009), financiado por la NFS, para el depósito de datos en el campo de ciencias de la vida, y en el que han participado directamente dos de nuestros investigadores.

2.2. **Bibliografía** más relevante.

Máximo 8000 caracteres

Allison, J., François, S. & Lewis, S., 2008. "SWORD: Simple Web-service Offering Repository Deposit", Ariadne Issue 54. Available at: <http://www.ariadne.ac.uk/issue54/allinson-et-al/> [Acceso: 29 de de enero de de 2013].

Arano, S. et al., 2011. La comunidad «Recursos y datos primarios» de la Universitat Pompeu Fabra: los repositorios institucionales como infraestructuras científicas: estudio de caso. Revista española de Documentación Científica, 34(3), pp.385-407. Available at: <http://redc.revistas.csic.es/index.php/redc/article/view/704> [Acceso: de diciembre de 21, 2013].

Björk, B.-C., 2004. Open Access to Scientific Publications – An Analysis of the Barriers to Change? Information Research, 9, p.paper 170. Available at: <https://helda.helsinki.fi/handle/10227/647> [Acceso: 1 de febrero de 2013].

Björk, B.-C. et al., 2010. Open access to the scientific journal literature: situation 2009. E. Scalas, ed. PloS one, 5(6), p.e11273. Available at: <http://dx.plos.org/10.1371/journal.pone.0011273> [Acceso: 12 de junio de 2013]



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

DA

Carroll, M.W., 2011. Why Full Open Access Matters. PLoS Biology, 9(11), p.e1001210. Available at: <http://dx.plos.org/10.1371/journal.pbio.1001210> [Acceso: 23 de noviembre de 2013].

Council for Libraries & Information Resources. (2013). Research Data Management: Principles, Practices, and Prospects (p. 99). Available at: <http://www.clir.org/pubs/reports/pub160/pub160.pdf> [Acceso: 24 de noviembre de 2013]

Creaser, C. et al., 2010. Authors' Awareness and Attitudes Toward Open Access Repositories. New Review of Academic Librarianship, 16(sup1), pp.145-161. Available at: <http://dx.doi.org/10.1080/13614533.2010.518851> [Acceso: 15 de septiembre de 2013].

Datasea. Datos abiertos de Investigación. <http://www.datasea.es/>

Doel, W. Van den. (2013). The future of the Social Sciences and humanities in Europe: collected LERU papers on the SSH research agenda (p. 66). Available at: http://www.leru.org/files/publications/The_future_of_the_ssh_and_humanties_in_europe_final.pdf [Acceso: 15 de noviembre de 2013]

García García, A., Rodríguez-Gairín, J.-M., Saorín, T., Millán González, L., García Massó, X., Ferrer-Sapena, A., & Pese. (2012). ODiSEA : International Registry on Research Data. BiD: Textos Universitaris de Biblioteconomia i Documentació. Available at: <http://www.ub.edu/bid/29/garcia2.htm> <http://www.ub.edu/bid/29/garcia1.htm> [Acceso: 15 de noviembre de 2013]

Ginsparg, P., 2011. ArXiv at 20. Nature, 476(7359), pp.145-7. Available at: <http://dx.doi.org/10.1038/476145a> [Acceso: 13 de agosto de 2013].

Greenberg, J., 2009. Theoretical Considerations of Lifecycle Modeling: An Analysis of the Dryad Repository Demonstrating Automatic Metadata Propagation, Inheritance, and Value System Adoption. Cataloging & Classification Quarterly, 47(3-4), pp.380-402. Available at: <http://dx.doi.org/10.1080/01639370902737547> [Acceso: 27 de enero de 2013].

Greenberg, J. et al., 2009. A Metadata Best Practice for a Scientific Data Repository. Journal of Library Metadata, 9(3-4), pp.194-212. Available at: <http://dx.doi.org/10.1080/19386380903405090> [Acceso: 13 de noviembre de 2013].

Grupo de Trabajo de "Depósito y Gestión de datos en Acceso Abierto" del proyecto RECOLECTA. La conservación y reutilización de los datos científicos en España(2012). FECYT. Fundación Española para la Ciencia y la Tecnología. (p. 59). Madrid. Available at: http://recolecta.fecyt.es/sites/default/files/contenido/documentos/informe_datos_cientificos_en_esp.pdf [Acceso: 23 de noviembre de 2013]

Hey, T., Tansley, S. & Tolle, K. eds., 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery 2nd ed., Redmond, Washington: Microsoft Research. Available at: http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf [Acceso: 29 de enero de 2013].

Klump, J., 2011. Criteria for the Trustworthiness of Data Centres. D-Lib Magazine, 17(1/2). Available at: <http://dlib.org/dlib/enero de 11/klump/01klump.html> [Acceso: 28 de enero de 2013].



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

Lewis, S., Castro, P. de & Jones, R., 2012. SWORD: Facilitating Deposit Scenarios. D-Lib Magazine, 18(1/2). Available at: <http://www.dlib.org/dlib/enero/12/lewis/01lewis.html> [Acceso: 29 de enero de 2013].

Lohr, S., 2011. Mining of Raw Data May Bring New Productivity, a Study Says - NYTimes.com. The New York Times. Available at: <http://www.nytimes.com/2011/05/13/technology/13data.html> [Acceso: 28 de enero de 2012].

Manyika, J. et al., 2011. Big data: The next frontier for innovation, competition, and productivity. Available at: http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation [Acceso: 28 de enero de 2013].

Meier zu Verl, C. & Horstmann, W. eds., 2011. Studies on Subject-Specific Requirements for Open Access Infrastructure, Bielefeld: Universitätsbibliothek. Available at: <https://pub.uni-bielefeld.de/luur/download?func=downloadFile&recordId=2445229&fileId=2460478> [Acceso: 30 de enero de 2013].

National Science Foundation, 2011. US NSF - CISE - Data Management Guidelines. Available at: http://www.nsf.gov/cise/cise_dmp.jsp [Acceso: 28 de enero de 2013].

Organisation for Economic Cooperation and Development, 2007. OECD Principles and Guidelines for Access to Research Data from Public Funding. Available at: <http://ec.europa.eu/research/era/docs/en/others-9.pdf> [Acceso: 28 de enero de 2013].

Park, J.-R., 2009. Metadata Quality in Digital Repositories: A Survey of the Current State of the Art. Cataloging & Classification Quarterly, 47(3-4), pp.213-228. Available at: <http://dx.doi.org/10.1080/01639370902737240> [Acceso: 13 de julio de 2013].

Schumpeter, 2011. Building with big data: . The Economist. Available at: <http://www.economist.com/node/18741392> [Acceso: 28 de enero de 2013].

Suber, P., 2009. Timeline of the open access movement. Available at: <http://dash.harvard.edu/handle/1/4724185> [Acceso: 27 de enero de 2013].

Tenopir, C., Birch, B., & Allard, S. (2012). Academic Libraries and Research Data Services: Current Practices and Plans for the Future (p. 55). Available at: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf [Acceso: 24 de noviembre de 2013]

Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2013). Academic librarians and research data services: preparation and attitudes. IFLA Journal, 39, 70–78. doi:10.1177/0340035212473089

Wulff-Barreiro, E., 2011. Approaches to Open Data for Science in Spain. Data Science Journal, 10, pp.13-23. Available at: http://www.jstage.jst.go.jp/article/dsj/10/0/10_13/_article [Acceso: 28 de enero de 2013].

IRG (e-Infrastructure Reflection Group), 2009. Report on Data Management, Espoo, Finland. Available at: http://www.e-irg.eu/images/stories/e-irg_dmf_report_final.pdf [Acceso: 28 de enero de 2013].



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

DA

2.3. Finalidad del proyecto, oportunidad de llevarlo a cabo en el contexto del reto elegido y adecuación del mismo a la Estrategia Española de Ciencia y Tecnología y de Innovación y, en su caso, a Horizonte 2020 o a cualquier otra estrategia nacional o internacional de I+D+i.

Máximo 8000 caracteres

La Estrategia Española de Ciencia y Tecnología y de Innovación marca ocho grandes retos que se enumeran en el Plan Estatal de Investigación Científica y Técnica y de Innovación para el período 2013-2016. Este proyecto se enmarca en el séptimo: Economía y sociedad digital. Específicamente en la prioridad temática y científico-técnica número tres que hace referencia a “Desarrollo, innovación y adopción de SOLUCIONES Y TECNOLOGÍAS ligadas a: (i) «cloud computing»; (ii) Open/Linked/Big Data y la reutilización información del sector público generando valor y conocimiento”.

El proyecto que se presenta se ajusta a estas prioridades y resulta pertinente porque, sin duda, los repositorios de datos comenzarán, a corto plazo, a formar parte de la e-infraestructura de universidades y centros de investigación. Los principales agentes que financian la investigación a nivel mundial recomiendan, motivan e incluso obligan, la preservación de los datasets o conjuntos de datos resultantes de los proyectos de investigación financiados con fondos públicos. La NSF de los EE.UU. obliga, desde el 18 de enero de 2011, a que cada proyecto de investigación que se presenta para obtener financiación de dicha agencia, incluya como parte de su propuesta, un plan de gestión de datos y un compromiso de depósito de los mismos en algún repositorio.

La propia Comisión Europea en su recomendación a los Estados Miembros de 17 de julio de 2012, relativa al acceso a la información científica y a su preservación (2012/417/UE) (<http://www.boe.es/doue/2012/194/L00039-00043.pdf>) sugiere a los estados de la UE que determinen “políticas claras destinadas a la difusión y el acceso abierto de los datos resultantes de la investigación financiada con fondos públicos”.

“Y que tales políticas deberán prever: unos objetivos concretos e indicadores para medir los avances, unos planes de aplicación, incluida la asignación de responsabilidades (sin olvidar la concesión de licencias) y la planificación financiera correspondiente”.

Asimismo, se deberá garantizar que, como resultado de estas políticas:

- los datos resultantes de la investigación financiada con fondos públicos sean accesibles al público y susceptibles de uso y reutilización a través de infraestructuras electrónicas digitales; determinadas preocupaciones, en particular por razones de defensa de la intimidad, confidencialidad comercial, seguridad nacional, intereses comerciales legítimos y derechos de propiedad intelectual, se tengan debidamente en cuenta; todos los datos, conocimientos técnicos e información, sea cual sea su forma o naturaleza, que obraran en poder de una empresa de colaboración pública/privada antes de la acción de investigación, y que hubieran sido identificados como tales, no tienen que cumplir esta obligación,
- los conjuntos de datos sean fácilmente identificables y puedan ligarse a otros conjuntos de datos y publicaciones a través de mecanismos apropiados, y se facilite información adicional para permitir su adecuada evaluación y utilización,
- las entidades que financian la investigación y manejan fondos públicos destinados a tal efecto y los centros académicos que reciben fondos públicos contribuyan a la ejecución de políticas nacionales mediante la puesta en marcha de mecanismos que permitan y recompensen el intercambio de datos de investigación,
- se fomenten o ejecuten programas destinados a los nuevos perfiles profesionales avanzados del ámbito de las tecnologías de tratamiento de datos”.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

Y aboga además por la preservación de la información científica, por su reutilización y por el desarrollo de infraestructuras científicas que sean interoperables con otras infraestructuras a nivel nacional e internacional. El proyecto que se presenta pretende, no sólo establecer un portal que permita el estudio y análisis comparado de políticas relacionadas con la puesta en marcha de repositorios de datos (software utilizado, políticas de licencia, de recompensa de publicación de datos, de identificación de objetos y de grupos de investigación, etc.) sino también con la puesta en marcha de un repositorio de datos pilotos con el que experimentar respecto a datos enlazados (linked data) y a la interoperabilidad con otros repositorios. El grupo dispone de los conocimientos y la experiencia necesaria para aprovechar la oportunidad tecnológica de la filosofía, tendencia o movimiento de Open Data y Linked Open Data (LOD) que aplica las tecnologías provenientes de la Web Semántica (RDF, URIs, SPARQL, etc.) para publicar conjuntos de datos de distinta naturaleza de forma abierta y enlazada, fomentando de esta forma la reutilización de dichos datos.

El proyecto tiene aspiración internacional en tanto en cuanto pretendemos contribuir al proyecto piloto de datos abiertos de investigación que la Comisión Europea está desarrollando para los proyectos de investigación financiados con fondos europeos en el marco de Horizon 2020 y en el que también tienen un fuerte interés organizaciones como COAR (Confederation of Open Access Repositories), LIBER (Ligue des Bibliothèques Européennes de Recherche - Association of European Research Libraries) u OpenAire (Open Access Infrastructure Research for Europe). Y se puede encuadrar tanto en las iniciativas de la Agenda Digital Europea, apartado 2.5 sobre investigación e innovación como en la Agenda Digital Española, objetivo número 2 "Desarrollar la economía digital para el crecimiento, la competitividad y la internacionalización de la empresa española" y objetivo número 6 "Promover la inclusión y alfabetización digital y la formación de nuevos profesionales TIC".



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

3. HIPÓTESIS Y OBJETIVOS DEL PROYECTO

Describa las razones por las cuales se considera pertinente plantear esta investigación, **la hipótesis de partida y los objetivos generales** perseguidos. Enumere brevemente, con claridad, precisión y de manera realista (acorde con la duración prevista del proyecto), los **objetivos específicos** que se persiguen con indicación expresa del investigador principal (IP) responsable en aquellos proyectos con dos investigadores principales.

Máximo 8000 caracteres

Partimos de la hipótesis de que los centros de investigación y las universidades estarán obligadas, o bien a mantener algún repositorio de datos de investigación o bien a conocer y asesorar a sus investigadores sobre dónde, cuándo y cómo depositar sus datos de investigación, puesto que será un requisito exigido por las agencias que financian la investigación. Muchas bibliotecas universitarias se verán forzadas no sólo a desarrollar este tipo de infraestructuras sino también a elaborar planes y políticas de gestión y preservación de estos datos. De la misma forma, será necesario contar con que parte del personal tenga que recibir cursos de formación o de reciclaje sobre la gestión de estos datos.

Aunque muy vinculados, estos repositorios de datos deberán ser repositorios diferentes a los repositorios institucionales y temáticos de resultados de investigación que se desarrollan desde hace años. Estas infraestructuras serán necesarias tanto en las ciencias puras como en las ciencias sociales y las humanidades. Deberán ser repositorios distintos, independientes e interoperables, puesto que las características de ambos difieren sustancialmente, tanto respecto a los objetivos para los que sirven uno y otro, como a las técnicas de gestión y mantenimiento, las políticas de acceso y depósito o la tipología de datos y cantidades de datos a preservar.

Este proyecto tiene tres grandes objetivos generales:

1. Crear un observatorio de repositorio de datos de investigación. Un espacio digital abierto en el que recopilar la información normativa (recomendaciones de la UE, mandatos de universidades o centros de investigación y financiación sobre el depósito de datos), registros de repositorios y de formatos de datos aceptables, estándares de interoperabilidad, software que se utilizan, esquemas de metadatos, vocabularios, políticas de incentivos para el depósito de datos, guías de buenas prácticas, etc. La información del observatorio no estará restringida a ningún área científica y su finalidad es ayudar a bibliotecas universitarias, centros de investigación y gestores de políticas científicas a tomar decisiones bien informadas en base a lo que se está haciendo en cada momento en todo el mundo respecto a los repositorios de datos de investigación.

El observatorio tiene como fin ayudar a concienciar a los investigadores, bibliotecarios, personal informático y actores de la comunicación científica en general de la importancia y los beneficios de compartir los datos de investigación.

2. Crear un repositorio piloto de datos de investigación en el que experimentar con este tipo de datos. El objetivo del repositorio será facilitar el registro y depósito de datos de investigación, su descubrimiento, acceso y reutilización y contribuir a las políticas de depósito en acceso libre de datos de investigación. El repositorio piloto albergará datos de investigación de dos fuentes principales:

a) De los grupos de investigación españoles reconocidos por distintos organismos oficiales de las áreas de las Ciencias Sociales y las Humanidades (con datos de encuestas, estadísticas, análisis de muestras arqueológicas, históricas o de otro tipo, corpus lingüísticos, objetos multimedia, etc.)

b) De los grupos de investigación de todas las áreas temáticas, incluyendo las Ciencias Sociales y Humanas, pero también de las Ciencias Exactas o Experimentales, tales como Biología, Química, Biomedicina o Física, a quienes se haya aprobado proyectos en el Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma de Generación de Conocimiento de esta convocatoria. La recolección se hará en el último año de este proyecto, con el fin de experimentar con todo tipo de datos.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

DA

3. Elaborar programas de cursos y estudios sobre repositorios de datos, alineados con otras escuelas internacionales. El objetivo es identificar qué habilidades y conocimientos necesitan poseer los profesionales de la información para tratar de cubrir las nuevas necesidades que requieren el trabajar con grandes y especiales volúmenes de datos.

Objetivos específicos:

- 1.1 Identificar, localizar, describir y organizar repositorios de datos, sitios web de interés, estándares, software y fuentes de información más fiables sobre repositorios de datos de investigación. Responsable: Antonio Hernández Pérez (IP1-AHP)
- 1.2 Definir y poner en marcha sistemas de alerta sobre el tema y definir las políticas de gestión, preservación y alimentación del sitio. Responsable: Antonio Hernández Pérez (IP1)
- 1.3 Desarrollar los servicios web necesarios para el observatorio y desarrollar las políticas de difusión necesarias. Responsable: (IP1-AHP)
- 2.1 Identificar los grupos de investigación a los que se han concedido ayudas en esta convocatoria así como a los investigadores individuales y repositorios institucionales que posean, o potencialmente generen, colecciones de datos de investigación en los centros de investigación españoles en los campos de las ciencias sociales y las humanidades. Responsable: (IP1-AHP)
- 2.2 Recopilar información y datos de los grupos para evaluar técnicamente las colecciones de datos de investigación (datasets) localizadas (formatos, programas de software para su acceso, derechos, etc.) Y analizar en profundidad los aspectos éticos, legales y económicos de la apertura y/o el enlace de datos en el ámbito científico. Responsable: (IP1-AHP)
- 2.3 Elaborar y difundir una guía de buenas prácticas para la gestión de un plan de datos que tipifique, precise y describa: los formatos de datos, los estándares que se deben utilizar para los datasets y los metadatos de una investigación; los productos de la investigación (qué tipos de datos, muestras, colecciones físicas o software se generan en el desarrollo del proyecto), las políticas de acceso y reutilización de datos, y las políticas de preservación de dichos datos o colecciones de datos. Se prestará especial atención al software existente, ya que se plantea como una de las carencias iniciales de muchos investigadores del campo de las humanidades y las ciencias sociales. Responsable: (IP1-AHP)
- 2.4 Analizar y utilizar perfiles de aplicación de metadatos para la descripción de datos y para garantizar el uso de estándares de intercambio de datos, que faciliten la recuperación y la interoperabilidad con otros repositorios nacionales e internacionales. Responsable: (IP1-AHP)
- 2.5 Plantear posibles soluciones al problema de autoridades, a través de un mecanismo basado en datos abiertos enlazados (LOD) en los repositorios institucionales y de investigación de cada centro investigador analizado.
- 2.6 Desarrollar y poner en marcha un repositorio de datos de investigación conforme a las normas OpenAire y a las 16 directrices relacionadas con administración y la gestión de datos confiables que otorgan el Sello de Aprobación de Datos (Data Seal of Approval) de la Data Archiving and Networked Services (DANS), organismo internacional de acreditación de repositorios de datos de investigación. Responsable: (IP1-AHP)
- 2.7 Analizar y sistematizar la conexión de las publicaciones científicas albergadas en los repositorios institucionales y/o temáticos con los datasets que subyacen a dichas investigaciones a través de tecnologías de Linked Open Data (LOD). Responsable: (IP1-AHP)
- 3.1 Identificación y análisis de cursos y perfiles de puestos de trabajo relacionados con repositorios. Responsable: (IP1-AHP)
- 3.2 Encuestas a bibliotecarios, investigadores, gestores científicos, etc sobre habilidades y conocimientos necesarios. Responsable: (IP1-AHP)



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

4. METODOLOGÍA, PLAN DE TRABAJO Y RESULTADOS PREVISTOS

4.1. Descripción de los materiales, infraestructuras y equipamientos singulares a disposición del proyecto que permitan abordar la metodología propuesta.

Máximo 4.000 caracteres

El proyecto se encuadra dentro de las actividades del grupo de investigación Tecnodoc, reconocido oficialmente por la Comunidad de Madrid y por la Universidad Carlos III de Madrid. Sus miembros disponen de un laboratorio dotado con algunos servidores y material de oficina básico para poder operar. La práctica totalidad de los miembros del grupo son profesores que disponen de despachos y toda la infraestructura que pone a su disposición la universidad. No posee ni se estiman necesarias materiales, infraestructuras o equipamientos singulares, más allá de un servidor y su alojamiento en el centro de cálculo de la universidad, así como otro pequeño material informático y fungible que permita ejecutar adecuadamente el proyecto.

4.2. Metodología. Detalle la metodología propuesta de acuerdo con los objetivos del apartado 3.

- Deberá indicarse la viabilidad metodológica de las tareas y reseñar los hitos o entregables previstos. Si fuera necesario, se incluirá una evaluación crítica de las posibles dificultades de un objetivo específico y un plan de contingencia para resolverlas.

El personal implicado en cada una de las tareas deberá especificarse en el cronograma.

- Si solicita ayuda para la contratación de personal, justifique claramente su necesidad y las tareas que vaya a desarrollar.

Máximo 24 000 caracteres

OBJETIVO 1. CREAR UN OBSERVATORIO SOBRE GESTIÓN DE DATOS DE INVESTIGACIÓN

Actividad 1.1: Identificar, localizar, describir y organizar repositorios de datos, sitios web de interés, estándares, software y otras fuentes de información fiables sobre repositorios de datos de investigación.

Metodología:

- Búsqueda, selección y recopilación de información a nivel internacional respecto a la gestión de datos científicos, abarcando: aspectos legales (recomendaciones de la OCDE, UE, mandatos de universidades o centros de investigación, y financiación sobre el depósito de datos); técnicos (criterios de formatos de datos aceptables, estándares de interoperabilidad, software para la creación y gestión de datos, esquemas de metadatos y otros vocabularios); situación actual (registros y directorios de repositorios) y nuevos proyectos; y otros aspectos políticos y organizativos (programas de incentivos para el depósito de datos, guías de buenas prácticas), etc.
- Implementación de un sitio web para la difusión de esta información, con capacidades para la alimentación continua de información, la creación de alertas, la participación, el establecimiento de comunidades de usuarios, etc.
- Arquitectura, organización y volcado en el Observatorio de la información recopilada.
- Creación de una breve guía y de un folleto explicativo sobre la importancia del desarrollo de planes para la gestión de datos de investigación, dirigido a gestores de políticas científicas bibliotecas y a los órganos de gobierno de universidades y centros de investigación.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

Entregables:

E1. Publicación de guía y folleto sobre la importancia de desarrollar planes para la gestión de datos de investigación.

Hitos

H1. Sitio web del Observatorio: definición, organización de las secciones iniciales y alimentación del portal con la información recopilada en esta fase inicial.

Actividad 1.2. Definición de las políticas de gestión, preservación y mantenimiento del Observatorio.

Metodología:

- Diseño de flujos de trabajo con relación a los contenidos, actualización y mantenimiento del observatorio.
- Elaboración de políticas y estrategias de selección, tratamiento, gestión y mantenimiento de los contenidos del observatorio.

Hitos

H2. Diseño de la arquitectura de la información del observatorio.

H3. Guía de responsabilidades y políticas del Observatorio, con descripción de roles, flujos de trabajo, estrategia de contenidos y políticas explícitas de gestión, preservación y mantenimiento del Observatorio.

Actividad 1.3.: Desarrollo de servicios y sistemas de difusión y alerta del Observatorio

Metodología:

- Diseño Web del sitio del Observatorio.
- Creación del wiki del proyecto para comunicación interna y externa y gestión de la documentación del proyecto.
- Fase inicial de difusión con comunicaciones a listas de distribución especializadas, nacionales e internacionales, y a otros espacios de comunicación (facebook, twitter).

Entregables

E2. Lanzamiento del wiki del Observatorio en Línea

E3. Lanzamiento de la primera versión del sitio Web del Observatorio en línea.

Hitos

H4. Dotar al Observatorio de visibilidad y reconocimiento en la comunidad local, nacional e internacional.

H5. Concienciar y crear de hábitos de actualización continua en cuanto a la práctica de gestión de datos de investigación.

H6. Constitución de comunidades de discusión y práctica en torno a la gestión de datos de investigación.

OBJETIVO 2. CREAR UN REPOSITORIO PILOTO DE DATOS DE INVESTIGACIÓN PARA LA EXPERIMENTACIÓN

Actividad 2.1: Elaborar los criterios de selección de los grupos de investigación que se incluirán como participantes y suministradores de datos en el proyecto piloto.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

2.1.a) Identificar grupos de investigación españoles que trabajan con colecciones de datos de investigación en los campos de las ciencias sociales y humanidades. Se pretende crear una base de datos con investigadores y grupos de investigación activos en centros españoles durante los últimos 5 años, y de entre el universo de investigadores tratar de encontrar a los que ya han trabajado con colecciones de datos. La limitación de 5 años se debe a que no es objetivo prioritario del trabajo recuperar todos los datasets utilizados años atrás, sino desarrollar la metodología para que a partir de la puesta en marcha de un repositorio los investigadores puedan hacer uso de él, si quieren depositando colecciones de datos antiguas.

2.1.b) Identificar los grupos de investigación a los que se han concedido ayudas en esta convocatoria en áreas distintas a las mencionadas (por ejemplo, en Biología, Física, Biomedicina). Sus datos, sin embargo, sólo se incluirían en una fase posterior, una vez obtenidas conclusiones y experiencias con los grupos en ciencias sociales y humanidades.

Metodología:

- Búsqueda sistemática en sitios web y bases de datos y a través de contactos con instituciones, para la identificación de grupos de investigación reconocidos por distintos organismos oficiales de las áreas de las ciencias sociales y las humanidades que cumplan con los criterios mencionados.
- Revisión de la resolución de concesión de ayudas del Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma Estatal de Generación de Conocimiento 2013-2016, e identificación de los grupos de investigación con proyectos aceptados de acuerdo a los criterios establecidos.
- Recopilación de información sobre estos grupos, y creación de un catálogo de grupos de investigación.

Hitos:

H1. Detectar y/o desarrollar una interfaz que permita la localización y visualización en un mapa por áreas temáticas de los grupos e investigadores; por localización geográfica, relaciones y grados de colaboración, etc.

H2. Detección y/o construcción de un esquema y vocabulario controlado sobre temáticas de trabajo de grupos de investigación.

Entregables:

E1. Base de datos con nombres de investigadores y grupos de investigación, afiliación, temática de líneas de investigación de acuerdo a un vocabulario controlado y campos de observaciones sobre la probabilidad potencial de que sean grupos de investigación que trabajen con colecciones de datos (datasets)

Actividad 2.2: Recopilar información sobre colecciones de datos de los grupos de investigación seleccionados en la fase anterior.

A partir de los investigadores y grupos identificados que trabajan con colecciones de datos, el trabajo a realizar es lograr recopilar colecciones de datos para su descripción y establecer acuerdos con estos grupos para su participación en el proyecto piloto de repositorio para la gestión integral de datos de investigación.

Metodología:

- Búsqueda de las colecciones de datos de grupos e investigadores



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

-Comunicación formal con los investigadores y los equipos de investigación, a los que se les suministrará la guía, el folleto elaborado en la actividad 1.1. y se les explicará los objetivos del proyecto. A estos grupos se les ofrecerá la realización de un plan integral de gestión de datos para su almacenamiento, descripción y preservación. Contamos ya con la colaboración de varias colecciones de datos de algunos de los investigadores de nuestra universidad.

-Tipificación y descripción de las colecciones de datos localizadas, en cuanto a su temática, formatos empleados, políticas de acceso, uso y reutilización, etc.

-Utilizando la metodología de mapas conceptuales, construir un vocabulario controlado de temáticas de colecciones de datos, relacionadas con los grupos de investigación.

Hitos:

H3. Acuerdos formales con los grupos de investigación que participarán en el proyecto de repositorio piloto.

Entregables

E2. Base de datos con inventario y tipología de colecciones de datos de investigación.

E3. Mapa conceptual / visualización gráfica de las temáticas y ubicación de las colecciones de datos.

Actividad 2.3: Evaluación técnica de las colecciones de datos identificadas y/o recopiladas

Una vez realizados los análisis de los aspectos técnicos, organizativos, legales y políticos de los distintos repositorios y colecciones de datos, se procederá a una evaluación técnica de estas colecciones.

Metodología:

-Establecimiento y desarrollo de los criterios de evaluación.

-Aplicación de los criterios de evaluación técnicos de colecciones de datos.

Hitos:

H4. Informe de evaluación de las colecciones seleccionadas aplicando el conjunto de criterios propuesto.

H5. Guía para evaluación de los requisitos técnicos de los datasets para poder formar parte de un repositorio de datos de investigación.

H6. Documento guía para la "curación", la preparación de los datos de las colecciones.

Actividad 2.4: Análisis comparativo de otros repositorios de datos

Esta actividad consiste en un estudio comparado de los aspectos técnicos, tecnológicos, económicos y sociales de otras experiencias internacionales de repositorios de datos de investigación, especialmente aquellos dedicados a las humanidades y las ciencias sociales.

Metodología:

-Recopilación bibliográfica y análisis comparativo de las tecnologías y estándares que están utilizando otros repositorios de datos de investigación. Incluye análisis de software de gestión del repositorio, estándares para el almacenamiento, la descripción y la interoperabilidad de estos objetos digitales.

-Recopilación bibliográfica y análisis comparativo de las políticas de depósito y acceso de los distintos repositorios (quién y qué se puede depositar, formatos admitidos, definición de niveles de acceso, etc.)



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

- Recopilación bibliográfica y análisis comparativo de los costes económicos de los repositorios (costes de instalación, de mantenimiento, de formación, etc.)
- Recopilación bibliográfica y análisis comparativo de los problemas legales relacionados con el acceso a las colecciones de datos y su reutilización.
- Recopilación bibliográfica y análisis comparativo de estrategias de incentivos para el depósito de estos datasets en repositorios.

Entregables:

E4. Publicación en la wiki del proyecto de informes sobre cada uno de los análisis comparativos antes mencionados.

Actividad 2.5. Análisis de los estándares y modelos de metadatos aplicables a la descripción de datos de investigación, que sustenten la recuperación en el repositorio piloto y aseguren la interoperabilidad con otros repositorios nacionales e internacionales.

Metodología:

- Comparación de metadatos de investigador con schemas formales utilizados en el ámbito de los repositorios de datos. Encuesta y validación al proyecto DRYAD y a través de la Comunidad internacional del metadatos del Dublin Core (DCMI-Science).
- Identificación de datasets y los metadatos que usan los investigadores.
- Conversión a elementos canónicos descriptivos en los dataset analizados. Utilizando técnicas de crowdsourcing con los investigadores del área (Ciencias Sociales y Humanidades) para determinar los valores de metadatos de usuario que utilizan y mejorar su valor (provenance, subject, etc.)

Hitos:

- H7. Perfil de Aplicación de Metadatos CURATOR-e.
- H8. Descripción de la semántica de los elementos de metadatos.
- H9. Definición del schema en RDF/XML.

Actividad 2.6. Identificación de investigadores individuales participantes en el proyecto piloto y definición de mecanismos de control de autoridades en base a de datos abiertos enlazados

Metodología:

- Obtener un listado de autores/investigadores en España, a través de los centros de investigación que proporcionen datos para el prototipo CURATOR-e, así como a través de ORCID, ResearchID, GoogleCitations, ScimaGo, SSCI, Mendeley, Latindex, ISNI, etc.
- Normalizar los datos de los autores/investigadores a partir de vocabularios controlados de autoridad, como VIAF o ISNI, codificar estos datos en RDF, publicarlos y enlazarlos con otras fuentes de datos LOD.

Hitos:

- H10. Registro de autoridad de investigadores en Ciencias Sociales y Humanidades publicado como LOD.

Actividad 2.7. Analizar las listas de autoridad incluidas en LOD-LAM Data Hub (CKAN: <http://thedatahub.org/>) y definir parámetros de compatibilidad

Metodología:



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

-Aplicar la metodología empleada en proyectos anteriores, como HIVE (Helping Interdisciplinary Vocabulary Engeneering), para vocabularios de materia y listados de autoridad de investigadores

Entregables:

E5. Creación de un vocabulario de listas de autoridades en SKOS.

Actividad 2.8. Publicar el vocabulario de listas de autoridad resultante como Linked Open Data

Metodología:

-Convertir el listado/los listados de autoridad a SKOS (Simple Knowled Organization Systems) y analizar su publicación en CKAN.

Entregables:

E6. Publicación del listado como Linked Open Data.

Actividad 2.9. Elaboración y difusión de una guía de buenas prácticas

Con los resultados de la actividad 2.2 y los estudios realizados en las actividades 2.3 a 2.8, se procederá a la elaboración de una guía de buenas prácticas para la elaboración y gestión de planes de datos que tipifique, precise y describa: los formatos de datos, los estándares que se deben utilizar para los datasets y los metadatos de una investigación; los productos de la investigación (qué tipos de datos, muestras, colecciones físicas o software se generan en el desarrollo del proyecto), las políticas de acceso y reutilización de datos, y las políticas de preservación de dichos datos o colecciones de datos. Se prestará especial atención al software existente, ya que se plantea como una de las carencias iniciales de muchos investigadores del campo de las humanidades y las ciencias sociales.

Metodología:

- Edición y publicación de una guía de buenas prácticas en la gestión de datos científicos.

- Elaboración de planes de difusión específicos para distintos actores: a) grupos de investigación e investigadores individuales, b) bibliotecas de universidades y centros de investigación; c) órganos de decisión en materia científica en las instituciones; y d) gestores de políticas científicas a nivel regional y nacional.

-Difusión, a través del Observatorio, de la información recopilada.

Hitos

H11. Planes de difusión adaptados a los cuatro tipos de actores identificados.

H12. Evaluación y validación piloto de la guía con una muestra seleccionada de grupos.

Entregables

E7. Guía de buenas prácticas para la gestión de datos científicos.

E8. Publicación de un artículo en una revista académica sobre las conclusiones de los estudios y la evaluación de la guía.

Actividad 2.10. Puesta en marcha del repositorio piloto CURATOR-e

Se pretende desarrollar y poner en marcha un repositorio de datos de investigación (con los grupos de investigación seleccionados en la actividad 2.1.a) conforme a las normas OpenAire y a las 16 directrices relacionadas con



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

administración y la gestión de datos confiables que otorgan el Sello de Aprobación de Datos (Data Seal of Approval) de la Data Archiving and Networked Services (DANS), organismo internacional de acreditación de repositorios de datos de investigación.

Metodología:

- Elaboración de un plan para la puesta en marcha del repositorio piloto.
- Puesta en marcha del repositorio piloto con primeras pruebas de descripción de colecciones de datos de los grupos seleccionados en la actividad 2.1.a.
- Pruebas de interoperabilidad con proveedores de servicios para el intercambio de datos.
- Recopilación de la actividad y de los problemas detectados por los investigadores en las actividades anteriores.
- Revisión y corrección de los procedimientos y problemas detectados
- Puesta en marcha del repositorio en fase de producción de colecciones de datos.
- Promoción del repositorio para el autodepósito de datos.
- Evaluación del repositorio de acuerdo a las directrices DANS.
- Solicitud de acreditación del repositorio por la DANS para la obtención del Data Seal of Approval (DSA)

Hitos

- H13. Primeras pruebas para el autodepósito por parte de los investigadores.
- H14. Validación del repositorio con base en los estándares internacionales (sello DSA)

Entregables:

- E9. Puesta en producción del repositorio piloto de datos de investigación.

Actividad 2.11. Analizar y sistematizar la conexión de las publicaciones científicas albergadas en los repositorios institucionales y/o temáticos con los datasets que subyacen a dichas investigaciones a través de tecnologías Linked Open Data, aplicadas a los grupos seleccionados en la actividad 2.1.a..

Metodología:

- Búsquedas temáticas en el ámbito de los datasets en los repositorios abiertos de publicaciones y encuestas a los investigadores colaboradores con la aportación de sus datos, para definir sus publicaciones resultantes.
- Partimos de la hipótesis que ninguno de los repositorios institucionales/temáticos españoles codifican sus metadatos en LOD. No obstante, se utilizará el análisis realizado en la actividad 2.1, para corroborar esta hipótesis:
 - a) En el caso de que exista un repositorio de publicaciones con metadatos ya definidos en LOD, o fácilmente exportables a RDF en un SPARQL Endpoint, se intentará la interconexión de dicho/s repositorio/s de publicaciones con CURATOR-e.
 - b) En el caso de que no exista ningún repositorio de publicaciones definido en LOD, se utilizará OAI2LOD para convertir los metadatos OAI-PMH de uno de los repositorios de pruebas de TECNODOC de publicaciones a un SPARQL Endpoint para amplificar las pruebas de concepto en relación a los repositorios de publicaciones y de datasets y su interoperabilidad.
- Inclusión en la Guía de Buenas Prácticas (Actividad 2.9), reflejando la necesidad de publicación de los metadatos de repositorios en triples de URIs en RDF y almacenados en SPARQL Endpoint, para facilitar su enlazado.

Hitos:

- H15. Obtención de porcentajes de correlación de datos-publicaciones en ciencias sociales y humanidades según la muestra definida en la actividad 2.1.a.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o B)

Entregables:

E10. Ampliación del prototipo de repositorio de datos CURATOR-e, a través de datos enlazados con repositorios de publicaciones.

Actividad 2.12. Incorporación al repositorio piloto de los datos correspondientes a los grupos de investigación seleccionados en la actividad 2.1.b) y experimentación con los mismos.

Metodología:

- Evaluación técnica de las colecciones de datos identificadas y/o recopiladas de este conjunto de grupos
- Estudio sobre similitudes y diferencias de los hallazgos generados con el primer conjunto de grupos de investigación (ciencias sociales y humanidades) con respecto a esta muestra (otras áreas temáticas).

Hitos:

H16. Actualización del observatorio con la información proporcionada con los resultados de este estudio comparativo
H17. Actualización o ampliación de la guía de buenas prácticas creada en la actividad 2.9 con los aspectos particulares de cada una de las áreas que afecten de manera específica la gestión de sus datos de investigación.

OBJETIVO 3. ELABORAR PROGRAMAS DE CURSOS Y ESTUDIOS SOBRE REPOSITORIOS DE DATOS, ALINEADOS CON OTRAS ESCUELAS INTERNACIONALES

Actividad 3.1. Identificación y análisis de cursos y perfiles de puestos de trabajo relacionados con repositorios.

Metodología:

- Revisión de las principales fuentes (listas de distribución, asociaciones profesionales, organizaciones académicas e instituciones científicas nacionales e internacionales) en las que aparezcan periódicamente cursos sobre la creación, gestión y mantenimiento de repositorios, así como ofertas laborales relacionadas con el tema.
- Extracción de los resultados obtenidos y categorización de aquellos que aparezcan con mayor frecuencia, tanto en la categoría de cursos como de perfiles de trabajo.
- Cotejo de los perfiles laborales requeridos con los cursos ofrecidos.

Hitos:

H1. Lista de principales contenidos que aparecen en los cursos analizados.
H2. Lista de perfiles laborales relacionados con los repositorios más demandados.
H3. Planteamiento de posibles materias, demandadas en los perfiles laborales, que no sean impartidas en los cursos analizados

Actividad 3.2. Encuestas a bibliotecarios, investigadores, gestores científicos, etc. sobre habilidades y conocimientos necesarios para la gestión de repositorios.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

Metodología:

- Establecimiento de un cuestionario semiabierto, con preguntas cerradas basadas en los resultados de la actividad 3.1., así como con algunas cuestiones abiertas. que permitan añadir reflexiones no incluidas en las anteriores.
- Envío a gestores, investigadores y bibliotecarios involucrados en proyectos relacionados con repositorios de datos, a partir de los datos obtenidos en las actividades 1.1 y 2.6.
- Recopilación y análisis de los resultados obtenidos
- Comparación de esos resultados con los obtenidos en la actividad 3.1

Hitos:

- H4. Elaboración de una lista definitiva de materias susceptibles de ser transmitidas mediante cursos, ya sean presenciales o en línea
- H5. Agrupación de esas materiales en subconjuntos que puedan ser impartidos tanto de forma independiente, en cursos de corta duración, como en cursos conjuntos de mayor duración.

Actividad 3.3. Diseño e impartición de cursos de formación dirigidos a bibliotecas y a investigadores para la gestión de datos de investigación.

Metodología:

- Diseño de cursos modulares que puedan ser organizados tanto para grupos específicos de profesionales como para audiencias más amplias con varios tipos de usuarios, ya sean presenciales o en línea.
- Determinación de colectivos que puedan recibir estos cursos: bibliotecas pertenecientes a REBIUN, investigadores que trabajen con datos de investigación (datasets), gestores de investigación.
- Propuesta y organización de seminarios o cursos a esos colectivos.
- Evaluación de la viabilidad del diseño de un curso práctico para investigadores (estudiantes de doctorado) sobre el uso de los datos específicos incluidos en Curator-e, siguiendo el ejemplo de la "Utrecht Data School" (<http://newmediastudies.nl/magazine/digging-deep-data-utrecht-data-school>).

Entregables:

E1.Realización de cursos y seminarios, tanto presenciales como a distancia, sobre repositorios de datos

Hitos:

- H6. Difusión de los conocimientos necesarios para la creación y gestión de repositorios de datos.
- H7. Concienciación sobre la importancia de los repositorios de datos entre la comunidad académica y científica, especialmente en España.
- H8. Incremento en el número de repositorios de datos existentes, y mejora en el funcionamiento de los ya existentes.



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

4.3.1. Cronograma. Para cada objetivo deben indicarse: el investigador responsable del mismo, los participantes involucrados, el período de ejecución (expresado en trimestres) y los hitos y entregables esperados con indicación del trimestre previsto (Tx) para su consecución.

Máximo 24 000 caracteres

Ejemplo:

O1: Breve descripción del objetivo 1

Responsable: Nombre y apellidos

Participantes: Nombre y apellidos P1; Nombre y apellidos P2; Nombre y apellidos P3; Nombre y apellidos P4; Nombre y apellidos Pn

Período de ejecución (en trimestres); p. ej., T1 – T2

H1 – Breve descripción del hito 1 y trimestre previsto, p. ej., T1

H2 – Breve descripción del hito 2 y trimestre previsto, p. ej., T2

Hn – Breve descripción del hito n ...

E1 – Breve descripción del entregable 1 y trimestre previsto, p. ej., T2

En – Breve descripción del entregable n...

O1: CREAR UN OBSERVATORIO SOBRE GESTIÓN DE DATOS DE INVESTIGACIÓN

Consistirá en un espacio digital abierto en el que recopilar la información normativa (recomendaciones de la UE, mandatos de universidades o centros de investigación y financiación sobre el depósito de datos), registros de repositorios y de formatos de datos aceptables, estándares de interoperabilidad, software que se utilizan, esquemas de metadatos, vocabularios, políticas de incentivos para el depósito de datos, guías de buenas prácticas, etc.

La información del observatorio no estará restringida a ningún área científica y su finalidad es ayudar a bibliotecas universitarias, centros de investigación y gestores de políticas científicas a tomar decisiones bien informadas en base a lo que se está haciendo en cada momento en todo el mundo respecto a los repositorios de datos de investigación.

El observatorio tiene como fin ayudar a concienciar a los investigadores, bibliotecarios, personal informático y actores de la comunicación científica en general de la importancia y los beneficios de compartir los datos de investigación.

Responsable: Antonio Hernández Pérez (IP1-AHP)

Participantes: Ivette Aportela Rodríguez, Gema Bueno de la Fuente, Antonio Hernández Pérez, M^a Eugenia Iglesias Moreno, Liliana Melgar Estrada, Tomás Nogales Flores, Virginia Ortiz-Repiso Jiménez, David Rodríguez Mateos, Técnico

Período de ejecución: T1-T6

Hitos y Ejecutables:

H1 - Definición del sitio web del Observatorio, organización de las secciones iniciales y alimentación del portal con la información recopilada en esta fase inicial. T1



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

H2 - Diseño de la arquitectura de la información del observatorio. T2

H3 - Guía de responsabilidades y políticas del Observatorio, con descripción de roles, flujos de trabajo, estrategia de contenidos y políticas explícitas de gestión, preservación y mantenimiento del Observatorio. T2

E1 - Publicación de guía y folleto sobre la importancia de desarrollar planes para la gestión de datos de investigación. T3

E2 - Lanzamiento del wiki del Observatorio en Línea. T3

E3 - Lanzamiento de la primera versión del sitio Web del Observatorio en línea. T4

H4 - Dotar al Observatorio de visibilidad y reconocimiento en la comunidad local, nacional e internacional. T5

H5 - Concienciar y crear de hábitos de actualización continua en cuanto a la práctica de gestión de datos de investigación. T5

H6 - Constitución de comunidades de discusión y práctica en torno a la gestión de datos de investigación. T6

O2: CREAR UN REPOSITORIO PILOTO DE DATOS DE INVESTIGACIÓN PARA LA EXPERIMENTACIÓN

El objetivo del repositorio será facilitar el registro y depósito de datos de investigación, su descubrimiento, acceso y reutilización y contribuir a las políticas de depósito en acceso libre de datos de investigación.

El repositorio piloto albergará datos de investigación de dos fuentes principales:

a) De los grupos de investigación de todas las áreas temáticas con proyectos aprobados en esta convocatoria de ayudas correspondientes al Prog. Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma Estatal de Generación de Conocimiento 2013-2016.

b) De los grupos de investigación reconocidos por distintos organismos oficiales de las áreas de las ciencias sociales y las humanidades (con datos de encuestas, estadísticas, análisis de muestras arqueológicas, históricas o de otro tipo, corpus lingüísticos, objetos multimedia, etc.)

En principio, el repositorio no aspira a convertirse en un servicio, sino en plataforma de experimentación para la concienciación de los investigadores respecto a la importancia de los datos de investigación, la problemática del registro y depósito de datos, la interoperabilidad con otros repositorios internacionales, la descripción de los datos que se depositen, la utilización de Linked Open Data, la usabilidad y accesibilidad de los mismos, la variedad y problemática de formatos, etc. No obstante, el servicio desarrollado podrá ser cedido a cualquier servicio de apoyo a la investigación que pueda estar interesado en asumir la gestión del repositorio a largo plazo.

Participantes: Ivette Aportela Rodríguez, Gema Bueno de la Fuente, Antonio Hernández Pérez, M^a Eugenia Iglesias Moreno, Liliana Melgar Estrada, Eva Méndez Rodríguez, José Antonio Moreira González, Nancy Diana Gómez, Tomás Nogales Flores, Virginia Ortiz-Repiso Jiménez, David Rodríguez Mateos, Sonia Sánchez Cuadrado, Técnico

Periodo de ejecución: T3-T12



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

D\

Hitos y ejecutables:

H1 - Detectar y/o desarrollar una interfaz que permita la localización y visualización en un mapa por áreas temáticas de los grupos e investigadores; por localización geográfica, relaciones y grados de colaboración, etc. T3

H2 - Detección y/o construcción de un esquema y vocabulario controlado sobre temáticas de trabajo de grupos de investigación. T3

E1 - Base de datos con nombres de investigadores y grupos de investigación, afiliación, temática de líneas de investigación de acuerdo a un vocabulario controlado y campos de observaciones sobre la probabilidad potencial de que sean grupos de investigación que trabajen con colecciones de datos (datasets). T4

H3 - Acuerdos formales con los grupos de investigación que participarán en el proyecto de repositorio piloto. T4

E2 - Base de datos con inventario y tipología de colecciones de datos de investigación. T5

E3 - Mapa conceptual / visualización gráfica de las temáticas y ubicación de las colecciones de datos. T5

H4 - Informe de evaluación de las colecciones seleccionadas aplicando el conjunto de criterios propuesto. T5

H5 - Guía para evaluación de los requisitos técnicos de los datasets para poder formar parte de un repositorio de datos de investigación. T6

H6 - Documento guía para la "curación", la preparación de los datos de las colecciones. T6

E4 - Elaboración y publicación en la wiki de los siguientes informes T6

Recopilación bibliográfica y análisis comparativo de las tecnologías y estándares que están utilizando otros repositorios de datos de investigación. Incluye análisis de software de gestión del repositorio, estándares para el almacenamiento, la descripción y la interoperabilidad de estos objetos digitales.

Recopilación bibliográfica y análisis comparativo de las políticas de depósito y acceso de los distintos repositorios (quién y qué se puede depositar, formatos admitidos, definición de niveles de acceso, etc.)

Recopilación bibliográfica y análisis comparativo de los costes económicos de los repositorios (costes de instalación, de mantenimiento, de formación, etc.)

Recopilación bibliográfica y análisis comparativo de los problemas legales relacionados con el acceso a las colecciones de datos y su reutilización.

Recopilación bibliográfica y análisis comparativo de estrategias de incentivos para el depósito de estos datasets en repositorios.

H7 - Perfil de Aplicación de Metadatos CURATOR-e. T7

H8 - Descripción de la semántica de los elementos de metadatos aplicables a la descripción de datos de investigación. T7

H9 - Definición del schema resultante de H6 y H7 en RDF/XML. T7

H10 - Obtención de registros de autoridad de investigadores en Ciencias Sociales y Humanidades publicado como LOD, mediante centros de investigación que proporcionen datos para el prototipo CURATOR-e, así como a través de ORCID, ResearchID, GoogleCitations, ScimaGo, SSCI, Mendeley, Latindex, ISNI, etc. T8

H11 - Planes de difusión adaptados a los cuatro tipos de actores identificados. T8



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

- E5 - Creación de un vocabulario de listas de autoridades en SKOS, basado en H10. T8
- E6 - Publicación del listado como Linked Open Data. T9.
- E7 - Guía de buenas prácticas en la gestión de datos científicos. T9
- H12 - Evaluación y validación piloto de la guía con una muestra seleccionada de grupos. T9
- H13 - Primeras pruebas para el autodepósito por parte de los investigadores. T9
- E8 - Publicación de un artículo en una revista académica sobre las conclusiones de los estudios y la evaluación de la guía. T10
- H14 - Validación del repositorio con base en los estándares internacionales (sello DSA). T10
- H15 - Obtención de porcentajes de correlación de datos-publicaciones en ciencias sociales y humanidades según la muestra definida en el actividad 2.1.a). T10
- E9 - Puesta en producción del repositorio piloto de datos de investigación T11.
- E10 - Ampliación del prototipo de repositorio de datos CURATOR-e, a través de datos enlazados con repositorios de publicaciones. T11
- H16 - Actualización del observatorio con la información proporcionada con los resultados de este estudio comparativo. T12.
- H17 - Actualización o ampliación de la guía de buenas prácticas creada en la actividad 2.9 con los aspectos particulares de cada una de las áreas que afecten de manera específica la gestión de sus datos de investigación. T12

O3: ELABORAR PROGRAMAS DE CURSOS Y ESTUDIOS SOBRE REPOSITORIOS DE DATOS, ALINEADOS CON OTRAS ESCUELAS INTERNACIONALES Elaborar programas de cursos y estudios sobre repositorios de datos, alineados con otras escuelas internacionales

El objetivo es identificar qué habilidades y conocimientos necesitan poseer los profesionales de la información para tratar de cubrir las nuevas necesidades que requieren el trabajar con grandes y especiales volúmenes de datos, así como incidir en la comunidad profesional y académica, motivando a la investigación sobre estos temas.

Responsable: Antonio Hernández Pérez (IP1-AHP)

Participantes: Ivette Aportela Rodríguez, Gema Bueno de la Fuente, Antonio Hernández Pérez, M^a Eugenia Iglesias Moreno, Liliana Melgar Estrada, Eva Méndez Rodríguez, José Antonio Moreiro González, Nancy Diana Gómez, Tomás Nogales Flores, Virginia Ortiz-Repiso Jiménez, David Rodríguez Mateos, Sonia Sánchez Cuadrado, Técnico

Periodo de ejecución: T11-T12

Hitos y entregables:

- H1 - Obtención de listas de principales contenidos en cursos relacionados con la gestión de repositorios de datos. T10
- H2 - Obtención y categorización de listas de requisitos aparecidos en ofertas laborales relacionadas con la gestión de repositorios de datos. T10



Convocatoria de ayudas a Proyectos de I+D+i
«RETOS INVESTIGACIÓN»
PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
SOCIEDAD
2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

H3 - Planteamiento de posibles materias, demandadas en los perfiles laborales, que no sean impartidas en los cursos analizados. T10

H4 - Elaboración de programas de cursos que incorporen materias no impartidas en otras actividades formativas ya en marcha. T11

H5 - Agrupación de esas materiales en subconjuntos que puedan ser impartidos tanto de forma independiente, en cursos de corta duración, como en cursos conjuntos de mayor duración. T11

E1 - Realización de cursos y seminarios, tanto presenciales como a distancia, sobre repositorios de datos. T11

H6 - Difusión de los conocimientos necesarios para la creación y gestión de repositorios de datos. T12

H7 - Concienciación sobre la importancia de los repositorios de datos entre la comunidad académica y científica, especialmente en España. T12

H8 - Incremento en el número de repositorios de datos existentes, y mejora en el funcionamiento de los ya existentes. T12



Convocatoria de ayudas a Proyectos de I+D+i
 «RETOS INVESTIGACIÓN»
 PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
 SOCIEDAD
 2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

o

4.3.2. Cronograma (gráfico). En el cronograma, marque la duración del OBJETIVO (X) y señale con H1...Hx los HITOS y con E1...Ex los ENTREGABLES, en su caso, de cada objetivo

Objetivo	Año 1 (trimestres)				Año 2 (trimestres)			
	1	2	3	4	5	6	7	8
1 (p. ej.)	X	X	H1; H2	E1		H1; H2	E1	
2 (p. ej.)			X	X	H1	H2; E1	E2	H3

AÑOS 1 y 2

Objetivo	Año 1 (trimestres)				Año 2 (trimestres)			
	1	2	3	4	5	6	7	8
1	H1	H2; H3	E1;E2	E3	H4; H5	H6		
2			H1;H2	E1;H3	E2;E3;H4	H5;H6;E4	H7;H8;H9	H10;H11;E5

Subdirección General de Proyectos de Investigación



Convocatoria de ayudas a Proyectos de I+D+i
 «RETOS INVESTIGACIÓN»
 PROGRAMA DE I+D+i ORIENTADA A LOS RETOS DE LA
 SOCIEDAD
 2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o B)

AÑOS 3 y 4

Objetivo	Año 3 (trimestres)				Año 4 (trimestres)			
	9	10	11	12	13	14	15	16
2	E6-7H12-13	E8;H14;H15	E9-10	H16-17				
3		H1;H2;H3	H4;H5;E1	H6-7-8				



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

5. IMPACTO ESPERADO DE LOS RESULTADOS

Describe brevemente el **impacto** científico-técnico, internacional, social o económico que se espera de los resultados del proyecto en el reto elegido.

El impacto esperado de los resultados debe completarse también en la aplicación de la solicitud. Su contenido podrá ser publicado a efectos de difusión si el proyecto resultara financiado en esta convocatoria¹.

Máximo 3500 caracteres

El impacto que se pretende conseguir, con los resultados obtenidos, se dirige en dos direcciones que son y deben ser necesariamente complementarias. En primer lugar, dotar a nuestro país de información precisa para crear un marco normalizado de actuación en cuestión de la gestión de los datos de investigación que se generan en los organismos públicos. España es uno de los países pioneros y más activos en el acceso abierto a la información, debe dar un paso más allá para regularizar y establecer los mecanismos necesarios para el acceso abierto a los datos de investigación, tal y como recomienda la Comisión Europea para el año 2020.

Este proyecto contribuirá a dotar a las diferentes administraciones públicas de la información necesaria para emprender con éxito la gestión de archivos/repositorios abiertos de datos de investigación. Realidad que se extenderá.

La segunda dirección se enfoca a concienciar a los investigadores de la necesidad de poner sus datos en abierto para que puedan generar nuevas investigaciones. Para ello, es necesario dotar a la comunidad investigadora de una herramienta, de uso sencillo, para depositar los datos de sus investigaciones. Pero también es imprescindible llevar a cabo una labor de alfabetización en este sentido. El repositorio piloto servirá para dar a conocer el potencial investigador de este "movimiento" y, al mismo tiempo, aumentar el número de publicaciones en acceso abierto de España, aumento la visibilidad de la actividad investigadora de nuestro país.

¹ Si el proyecto resultara financiado en esta convocatoria, el órgano concedente podrá solicitar las conclusiones y los resultados del mismo, que podrán ser publicados a efectos de difusión.



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

6. DIFUSIÓN Y TRANSFERENCIA DE LOS RESULTADOS

6.1. PLAN DE DIFUSIÓN.

Máximo 4000 caracteres

El proyecto CURATOR-e permitirá establecer mecanismos para buscar y conocer dónde se encuentran los grandes depósitos de datos de investigación en las universidades y centros de investigación españoles en el área de las humanidades y las ciencias sociales. Asimismo, permitirá transferir dicha información a las bibliotecas o unidades encargadas de la gestión de repositorios, la experiencia metodológica sobre auditorías para la identificación, localización, descripción y evaluación de estas colecciones de datos.

Como canal de transferencia a la sociedad se contactará con el grupo de trabajo en torno a RECOLECTA, Recolector de Ciencia Abierta. RECOLECTA, como proveedor de servicios basado en el protocolo OAI-PMH, para recopilar literatura científica proveniente de los repositorios de las Universidades españolas (fundamentalmente, de publicaciones), tendrá que plantearse también, a corto o medio plazo, el tratamiento de datasets, la implementación de OAI-ORE, y la interoperabilidad entre resultados-datos y resultados-publicaciones de investigación.

REBIUN (Red de Bibliotecas Universitarias) incluye entre los objetivos de la línea 2 de su nuevo Plan Estratégico 2020: promover mandatos y políticas institucionales de acceso abierto a la producción científica de cada universidad para incrementar la visibilidad y el impacto de la universidad. En este sentido se utilizará la Universidad Carlos III de Madrid, a través del servicio de Biblioteca, para adoptar CURATOR-e como prototipo para presentar la necesidad y motivar políticas institucionales en relación a la apertura e interoperabilidad no sólo de las publicaciones o literatura científica, sino también de los datos.

La experiencia adquirida en la creación y gestión de este repositorio permitirá a los centros y las bibliotecas que decidan acometer el proceso de implantación de estos repositorios conocer con antelación los problemas técnicos, económicos y sociales que habrán de tener en consideración, tanto técnicos (formatos de datos, metadatos, gestión de grandes volúmenes de datos, interoperabilidad, software, etc.), como económicos (costes de implantación, costes de almacenamiento, costes de transacción, costes de tiempo/persona, etc.) y sociales (cultura de los investigadores, promoción y uso del repositorio, etc.)

Plan de difusión. Contribuciones científico - técnicas

La difusión final de los resultados obtenidos por la investigación se desarrollará mediante dos vías fundamentales: el servidor web del proyecto, que será una herramienta de trabajo tipo Wiki pero de acceso público, en el que se irán publicando los informes de progreso y que albergará también el repositorio piloto para los investigadores o grupos de investigación que deseen depositar y compartir sus colecciones de datos de investigación en abierto (open datasets). Asimismo, para potenciar la difusión de resultados se utilizarán canales sociales como Twitter, o bien listservs relacionados con el tema (DC-Science, OS-Repositorios, etc.)

En cuanto a la difusión mediante los canales habituales de comunicación científica, se prevé que el proyecto dé lugar, al menos, a las siguientes iniciativas:

- Publicación de comunicaciones en congresos y reuniones científicas de relevancia nacional e internacional. Tendrán carácter prioritario aquellos específicos sobre repositorios, e-ciencia, metadatos y acceso abierto de carácter internacional.



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

- Publicación de artículos científicos en las principales revistas de la especialidad. Se priorizará la difusión en revistas de carácter internacional. Se publicará en revistas del ámbito de la Documentación pero también en otras de distintas ramas de las Humanidades y las Ciencias Sociales.
- Elaboración de informes de situación, de alcance nacional e internacional, segmentado por áreas temáticas, sobre la gestión de datos de investigación en repositorios, revistas y otros agentes de distribución.
- Realización de seminarios y asesoramiento a instituciones.

6.2. PLAN DE TRANSFERENCIA Y EXPLOTACIÓN, en su caso, de los resultados del proyecto, incluyendo aquellas entidades interesadas en los resultados del proyecto, concretando su participación y/o aportaciones al desarrollo del mismo.

Máximo 4000 caracteres

El plan previsto es transferir la experiencia y el repositorio piloto al grupo de trabajo de gestión de datos de la Red de Bibliotecas Universitarias (Rebiun), algunos de cuyos responsables han mostrado ya su interés en el proyecto. Idealmente, esperamos que los resultados del proyecto contribuyan a que se materialicen políticas nacionales o autonómicas consensuadas que permitan llevar a cabo el tratamiento y gestión de los datos de investigación. Estos datos conforman un activo tangible para el emprendimiento y desarrollo de nuevas investigaciones y multiplican las inversiones realizadas con presupuestos públicos ya que revierten o pueden revertir en nuevas investigaciones que acrecenten el acervo investigador de España. Las entidades interesadas serán fundamentalmente todos los organismos públicos nacionales, europeos e internacionales ya que por una parte podrán servir para regular las políticas nacionales en este ámbito y, por otro, servirán para equiparar la política española en tratamiento de datos de investigación con la de otros países. Y, en definitiva, contribuirá, igual que ocurrió con el movimiento Open Access a la internacionalización y la libre circulación de los datos científicos generados con los recursos y esfuerzos de todos



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

7. FINANCIACIÓN PÚBLICA Y PRIVADA (PROYECTOS Y CONTRATOS DE I+D+i) DEL EQUIPO DE INVESTIGACIÓN

Indique únicamente la financiación en los últimos 5 años (2009-2013), ya sea de ámbito autonómico, nacional o internacional. Incluya también las solicitudes pendientes de resolución.

Debe indicar: referencia, título, investigador principal, entidad financiadora, duración y cuantía de la subvención, además de las siguientes claves: 0 = es el mismo tema; 1 = está muy relacionado; 2 = está algo relacionado; 3 = sin relación; C o una S según se trate de una concesión o de una solicitud.

Máximo 20 proyectos o contratos.

Máximo 12 000 caracteres

Ejemplo:

1. Referencia XXX2009-nnnnn. "Título", "Investigador principal", MICINN, 01/2010-12/2012. xx.xxx €. 1-C
2. Referencia XXX2012-nnnnn. "Título", "Investigador principal", MINECO, 01/2013-12/2015. xxx.xxx €. 3-C
3. FP7-NMP-20xx-SMALL-x. "Título", "Investigador principal", UE, 01/2013-12/2015. xxx.xxx €. 2-C
4. FP7-PEOPLE-2012-ITN. Título", "Investigador principal", UE, 01/2013-12/2015. xxx.xxx €. 1-S
- ...
- 20.

- 2013/304PN004. "Adaptación de contenidos de una Plataforma de formación a distancia en CIE10MC (Clasificación internacional de Enfermedades, 10ª revisión-modificación Clínica)/PCS (Procedure Codification System)", "Antonio Hernández Pérez", Ministerio de Sanidad, Política Social e Igualdad, 10/2013-12/2013. 53.167,40€, 3-C.
- 2013/07CM013. "Curso online para el desarrollo de proyectos de mejora de la seguridad del paciente e implantación de buenas prácticas en centros asistenciales", "David Rodríguez Mateos", Ministerio de Sanidad, Política Social e Igualdad, 10/2013-12/2013. 11.718 €, 3-C.
- Referencia, "Workshop on Digital Libraries Education", "Jeffrey Pomerantz y Eva Mª Méndez Rodríguez" National Science Foundation Estados Unidos, Fecha de Inicio – Fin del Proyecto, 19.018,33 €, 1-C.
- 2008/00482/001, "Estudio, catalogación, digitalización y publicación en internet del fondo fotográfico de la emigración española a través de la revista Carta de España", "Antonio Hernández Pérez", Ministerio de Trabajo e Inmigración, 01/2009-12/2009, 40.172 €, 3-C.
- SEJ2007-67677. "Situación de los Centros de Documentación: rentabilidad y uso, condiciones laborales y formación de los documentalistas en los medios de comunicación autonómicos, provinciales y locales en las Comunidades Autónomas de Madrid, Galicia y Cataluña", "Eulàlia Fuentes i Pujol", MEC, 10/2007-09/2010. 76.860€, 3-C.
- PHB2007-0099-PC. "Renovación del proyecto Sistema de definición de perfiles profesionales y de competencias de formación para estudiantes de información y documentación: estudio comparativo para Brasil y España basado en la gestión semántica del conocimiento", "José Antonio Moreira González", MEC, 01/2010-12/2011. 7200€, 3-C.
- PCI A1/049617/11. "Consolidación de la capacidad tecnológica y formativa para la enseñanza a distancia y el autoaprendizaje de los docentes y profesionales de la información en los centros de la Universidad Eduardo Mondlane (Mozambique)", "José A. Moreira González y Manuel Valente Mangué". Secretaría de Estado de Cooperación Internacional. 12/2011-06/2013, 34.055 + 3000€, 3-C.



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

- CCG10-UC3M/HUM-5440. "Impacto del e-Learning en la Enseñanza y actividades de los docentes en Educación superior de la Universidad Carlos III de Madrid", "David Rodríguez Mateos", CAM y UC3M, 01/2011-12/2011. 1.250€, 3-C
- 2011/00555/001. "Realización de recursos informativos on line y asesoramiento para la implantación de buenas prácticas en seguridad del paciente", "David Rodríguez Mateos", Ministerio de Sanidad, Política Social e Igualdad, 10/2011-12/2011. 128.430€, 3-C.



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

8. RELACIÓN DE LAS PERSONAS QUE COMPONEN EL EQUIPO DE TRABAJO

Relacione las personas del equipo de trabajo que participarán en la ejecución del proyecto de investigación (de acuerdo con el artículo 18.9 de la convocatoria). Recuerde que el currículum de los doctores aquí incluidos deberá aportarse en la aplicación de solicitud del proyecto.

Máximo 8000 caracteres

Indique NOMBRE Y APELLIDOS y las siguientes claves según proceda:

TITULACIÓN: Doctor (D); Licenciado o ingeniero (L); Graduado (G); Máster (M); Formación profesional (FP); Otros (O)

TIPO DE CONTRATO: En formación (F); Contratado (C); Técnico (PT); Entidad extranjera (EE); Otros (OC)

DURACIÓN DEL CONTRATO: Indefinido (I); Temporal (T)

1. Nombre y apellidos. G-F-T
2. Nombre y apellidos. FP-PT-I
3. Nombre y apellidos. EE

Equipo de investigación

Antonio Hernández Pérez. D-C-I

Virginia Ortiz-Repiso Jimenez. D-C-I

Eva Méndez Rodríguez. D-C-I

Tomas Nogales Flores. D-C-I

José Antonio Moreira González. D-C-I

Equipo de Trabajo

Gema Bueno. D-C-T

David Rodríguez Mateos. D-C-T

Sonia Sánchez Cuadrado. D-OC

Ivett María Aportela Rodríguez. M-F-T

Liliana Melgar Estrada. M-F-T

María Eugenia Iglesias Moreno. M-F-T

Nancy Diana Gomez. M-OC

Virginia Bazán Gil D-OC



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

9. CAPACIDAD FORMATIVA DEL EQUIPO SOLICITANTE

Complete únicamente en caso de que se solicite la inclusión del proyecto en la convocatoria de “*Contratos predoctorales para la formación de doctores*”².

Especifique:

- plan de formación previsto en el contexto del proyecto solicitado;
- relación de tesis realizadas o en curso con indicación del nombre del doctorando, el título de tesis y la fecha de obtención del grado de doctor o de la fecha prevista de lectura de tesis (últimos 10 años).

Máximo 8000 caracteres

Los miembros del equipo pertenecientes a esta universidad forman parte del grupo Tecnodoc, reconocido por la Comunidad de Madrid, grupo consolidado y con capacidad para acoger a becarios de Formación de Personal Investigador, e integrarlos directamente en el desarrollo del proyecto. De hecho, tres de los miembros del equipo investigador de CURATOR-e, están realizando sus tesis doctorales en el seno de este grupo formal de investigación, y sus directores son asimismo miembros más senior de Tecnodoc.

De igual forma, en el desarrollo de este proyecto, el grupo tiene la capacidad de integrar más investigadores en esta misma circunstancia, provenientes de nuestro programa de Doctorado en Documentación: archivos y bibliotecas en el entorno digital, másteres universitarios, u otra procedencia, que tengan interés en esta línea de trabajo de Open Access/Open Science. De hecho, dos de los integrantes de este equipo, comenzaron de esta forma su actividad en la universidad y uno de ellos (Gema Bueno) ha completado específicamente su tesis doctoral sobre un modelo de repositorio institucional para la gestión, intercambio, reutilización y preservación de materiales digitales de docencia y aprendizaje en una biblioteca universitaria.

Desde el año 1996, las personas que forman el equipo de investigación y el de trabajo han participado, dependiendo de su edad y trayectoria, como docentes en másteres de investigación dirigidos tanto a la formación de doctores como a la especialización profesional. Esta docencia ha estado enfocada fundamentalmente a las tecnologías de la información y en diferentes aspectos, modalidades y enfoques. Los miembros del equipo de investigación han dirigido más de una docena de tesis doctorales, participado en numerosos tribunales de tesis y tutorizado muchos trabajos de Fin de Máster y de Fin de Grado. Todos ellos han versado sobre las tecnologías de la información y su aplicación en el contexto actual de desarrollo.

El departamento al que está adscrito la mayoría del equipo investigador recibe habitualmente tanto a investigadores postdoctorales como a becarios en formación, especialmente a través del programa de doctorado. El IP es director del Programa de Doctorado en Documentación, programa verificado por la ANECA.

En la actualidad, los miembros imparten docencia de grado y másteres y doctorado relacionada con el uso y gestión de información electrónica y con la aplicación de las tecnologías de la información en diversas titulaciones de todo tipo de ramas del conocimiento: la Licenciatura en Periodismo y los Grados en Información y Documentación, Periodismo, Derecho, Administración de Empresas, Ciencias Políticas, Finanzas y Contabilidad, Ingeniería Mecánica, Ingeniería Informática, Ingeniería Telemática, Ingeniería en Sistemas de Comunicaciones, Ingeniería en Sistemas Audiovisuales e

² La inclusión en la convocatoria de “*Contratos predoctorales para la formación de doctores*” solo será posible en un número limitado de los proyectos aprobados



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

Ingeniería Informática, así como los grados conjuntos en Informática y Administración de Empresas, y Derecho y Administración de Empresas.

La mayoría de los integrantes del proyecto imparten docencia en un máster oficial, el Máster en Bibliotecas y Servicios de Información Digital; uno de ellos es el director del Máster, y otro forma parte del comité de dirección del mismo. Todos los miembros del proyecto participan además como docentes en otros másters propios, todos ellos impartidos por la Universidad Carlos III de Madrid: especialmente, el Máster en Gestión y Producción de E-Learning (a distancia), desde 2007, en el que varios integrantes del proyecto imparten sendos módulos sobre tecnologías de la información, y, específicamente, sobre tecnologías y estándares en elearning, áreas claves para el desarrollo del repositorio de recursos educativos abiertos. Además, los miembros del proyecto participan en otros másters propios sobre Documentación Audiovisual, Comunicación Científica, Comunicación en Moda y Belleza, Guión de Cine y Televisión, y Museología y otros cursos especializados en organismos de relevada importancia en el área de bibliotecas digitales.

Por último, uno de los miembros del equipo es la actual vicerrectora adjunta de Infraestructuras en la Universidad Carlos III de Madrid, mientras que otro de los miembros es la directora del Departamento de Biblioteconomía y Documentación. La implicación de ambos cargos universitarios es fruto de la toma de conciencia sobre la necesidad de aplicar los resultados obtenidos tanto en la formación de grado como de posgrado impartida por el Departamento, además de en la propia actividad investigadora de la universidad a la que pertenecen.

Tesis realizadas

Modelo de repositorio institucional de contenido educativo: la gestión de materiales digitales de docencia y aprendizaje en la biblioteca universitaria. Realizada por Gema Bueno de la Fuente (18/6/2010). Director: Antonio Hernández Pérez

Sistema de recuperación conceptual mediante niveles semánticos en la representación de esquemas de metadatos. Realizada por Vicente Palacios Madrid (20/09/2010). Director: José A. Moreiro González

Análisis de los criterios de relevancia documental mediante consultas de información en el entorno web. Realizada por Valentín Moreno Pelayo (8/10/2010). Directores: José A. Moreiro González y Sonia Sánchez Cuadrado.

La colaboración y la visibilidad en las disciplinas de física en science citation index y arxiv (2.000 – 2.005). Realizada por Ana Isabel Bonilla Calero (26/11/2009). Directores: Isabel Gómez Caridad y Antonio Hernández Pérez

Tesis en curso:

Expert tagging behavior: a case study in the film domain. Liliana M. Melgar Estrada. Dirigida por Eva Méndez Rodríguez (fecha prevista de lectura - 2014)

Datos abiertos: evaluación y análisis de nivel de apertura. Nancy Diana Gómez. Dirigida por Eva Méndez Rodríguez y Antonio Hernández Pérez (fecha prevista de lectura - 2016)



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R\

**10. IMPLICACIONES ÉTICAS Y/O DE BIOSEGURIDAD
DE LA INVESTIGACIÓN PROPUESTA**

Si en la aplicación electrónica de solicitud ha contestado **afirmativamente** en alguno de los aspectos relacionados con implicaciones éticas o de bioseguridad allí recogidos, explique los aspectos éticos referidos a la investigación que se propone, las consideraciones, procedimientos o protocolos a aplicar en cumplimiento de la normativa vigente, así como las instalaciones y las preceptivas autorizaciones de las que dispone para la ejecución del proyecto.

Máximo 8000 caracteres



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R)

ANEXO I

Incluir las fórmulas, reacciones químicas, etc que por el tipo de letra del texto del formulario no hayan podido ser insertadas en el mismo. El número de ecuación debe coincidir con su llamada en el texto.

Máximo 2 páginas

Respete la extensión máxima indicada. Recuerde que en virtud del artículo 11 de la convocatoria **NO SE ACEPTARÁN NI SERÁN SUBSANABLES MEMORIAS CIENTÍFICO-TÉCNICAS** que no se presenten en este formato.



Convocatoria de ayudas « RETOS INVESTIGACIÓN »
PROGRAMA ESTATAL DE I+D+I ORIENTADA A LOS
RETOS DE LA SOCIEDAD

2013

MEMORIA CIENTÍFICO-TÉCNICA DE PROYECTOS INDIVIDUALES (TIPO A o

R)

ANEXO II

Únicamente en el caso en el que se considere necesario para aclarar ciertos aspectos del proyecto, incorporar las imágenes o figuras (formato TIFF, JPEG o GIF), hasta **un máximo de 6**, a las que se ha hecho referencia en el texto.

Respete el número máximo de figuras indicado. Recuerde que en virtud del artículo 11 de la convocatoria NO SE ACEPTARÁN NI SERÁN SUBSANABLES MEMORIAS CIENTÍFICO-TÉCNICAS que no se presenten en este formato.