

ELABORAR
UN
PLAN
DE
XESTIÓN
DE
DATOS
DE
INVESTIGACIÓN

{ **GUÍA PRÁCTICA** }
{ FEBREIRO 2020 }



Servizo de Biblioteca Universitaria
UNIVERSIDADE DA CORUÑA

Un plan de xestión de datos é un documento no que se recolle como se organizan, estruturan, gardan e conservan os datos empregados e creados no transcurso dunha investigación.

Ao comezo dunha investigación tómanse decisións relacionadas coa procura de información e datos que serán precisos para levar a bo porto a investigación. Comeza así a escolla de software, de formatos, control de copias de seguridade, etc. Se redactamos un plan de xestión de datos vainos servir para revisar as decisións iniciais tomadas sobre a xestión dos datos. Teremos un documento de referencia para garantir boas prácticas e asegurar a calidade na xestión dos datos, conservalos e ofrecelos en formatos que nos permitan o seu uso por parte dos membros do grupo de investigación, para cumprir cos mandatos dos financiadores e para a súa reutilización por terceiros se iso é preciso.

O plan debe ser un documento vivo que recolla os cambios que se produzan ao longo da investigación, como por exemplo a incorporación de novos datos, a aparición dunha patente, a integración de novos membros no grupo de investigación, etc. Polo tanto será preciso revisalo de maneira periódica e ofrecer acceso ao seu contido a todos os participantes na investigación.

CREACION DO PLAN DE DATOS DE INVESTIGACION (DATA MANAGEMENT PLAN, DMP)

Aínda que existen varias ferramentas para crear un plan de datos de investigación a mellor opción para o ámbito europeo é **DMPONLINE**, escollida polas entidades de investigación europeas.

Cales son os pasos necesarios para crear un plan de datos en  ?

PASO 1 CREAR UNHA CONTA

Usar o correo institucional e engadir un contrasinal.

Ao acceder hai que especificar a nosa institución, non imos atopar a UDC na listaxe, polo tanto temos que escoller a opción de "A miña institución non aparece" e amosará un recadro no que debemos introducir o nome oficial da nosa universidade (Universidade da Coruña).

PASO 2 ELIXIR UNHA PLANTILLA

O seguinte paso é escoller unha plantilla. Hai entidades financiadoras que teñen unha plantilla propia na que indican a información que se precisa para cumprir co seu mandato, por exemplo H2020. Outras axencias non especifican plantilla, neste caso poderemos escoller entre varias opcións.

Se o proxecto é de H2020 escolleremos a plantilla correspondente. Nela atoparemos os requisitos básicos para as tres versións que hai que entregar. Unha antes dos 6 primeiros meses, outra intermedia e unha versión final. Este tipo de proxectos esixen que se fagan tantas versións como sexan necesarias co fin de dispoñer de toda a información relativa aos principais cambios que houbo ao longo do tempo. Debemos redactar todos os elementos precisos para que os datos cumpran cos principios FAIR: Findable, Accessible, Interoperable e Reusable.

Se a investigación non está financiada ou se o financiador non especifica unha plantilla é aconsellable elixir a plantilla do European Research Council (ERC), é máis sinxela pero conta cos elementos necesarios para poder cumprir cos principios FAIR de datos de investigación.

PASO 3 REDACTAR O PLAN

O nivel de detalle da redacción pode variar dun proxecto a outro, da natureza ou complexidade dos datos, etc. Os apartados mínimos para un plan de datos son os seguintes:

► PRESENZA DE IDENTIFICADORES E DUN RESUMO SOBRE O PROXECTO.

Os identificadores son aqueles elementos que serven para coñecer a que proxecto pertencen os datos, quen son os responsables do proxecto, do control dos datos e da redacción do plan.

Teremos que indicar o ID do proxecto, o investigador principal, o responsable do plan e engadir a entidade ou entidades financiadoras, as institucións ás que pertencen os participantes do proxecto, etc.

Para identificar aos investigadores é preciso engadir o seu identificador dixital persistente ORCID.

O nome das entidades deben ir na súa forma oficial. Existe tamén un identificador dixital para entidades de investigación, o ROR (Research Organization Registry). O identificador da Universidade da Coruña é <https://ror.org/01qckj285>.

No resumo do proxecto indicaremos os principais obxectivos e os elementos necesarios para coñecer como se van recoller os datos, se hai que compralos, se virán de observacións, se se van a reutilizar datos de terceiros, etc. Tamén hai que indicar a tipoloxía dos datos e o seu volume, os programas necesarios para traballar con eles, procedemento para nomear ficheiros e cartafoles, etc.

▶ COMO CUMPRIR COS PRINCIPIOS FAIR (FINDABLE, ACCESSIBLE, INTEROPERABLE E REUSABLE).

FINDABLE

Debemos especificar que imos facer para que os datos se poidan encontrar. Para iso haberá que contar cun identificador persistente (doi, handle) e describir nun repositorio o conxunto ou conxuntos de datos. Na descrición engadiremos o título, responsables, financiadores, palabras clave, etc.

ACCESSIBLE

Co fin de garantir a accesibilidade é preciso escoller un repositorio fiable para depositar, preparar a documentación necesaria para poder acceder aos datos e empregar metadatos que describan o contido do conxunto de datos.

A localización e accesibilidade van depender da fiabilidade do repositorio. Os requisitos mínimos que debe cumprir un repositorio son:

- ▶ Que ofrezca un identificador persistente para que se poidan buscar, recuperar e citar os datos.
- ▶ Que os metadatos recollan información relevante para a identificación de persoas e institucións participantes na investigación (ORCID dos investigadores, nome dos financiadores, nome do proxecto, título do dataset, data, versión, relación con outros resultados de investigacións como artigos, conferencias, etc).
- ▶ Que empregue estándares de metadatos aceptados pola comunidade científica (JSON, XML, RDF, Dublin Core, etc.).
- ▶ Que asegure a autenticidade e integridade dos datos.
- ▶ Que informe sobre as licenzas e permisos para a reutilización dos datos.
- ▶ Que permita establecer diferentes niveis de acceso para que os datos sexan tan abertos como sexa posible e tan pechados como sexa necesario.

Fontes para escoller un repositorio fiable: www.re3data.org e www.fairsharing.org.

INTEROPERABLE

Temos que recoller os estándares aplicables a área de coñecemento para asegurar a interoperabilidade dos datos. Se na área non hai un estándar específico haberá que escoller un xenérico. Empregar vocabularios e metodoloxías que faciliten a comprensión e a integración dos datos no seu campo de coñecemento ou de actuación.

REUSABLE

Os datos deben poder ser reutilizados pola comunidade científica ou docente. Para iso é preciso asignarlles unha licenza de uso. A máis extendida é a Creative Commons 4.0 International. OpenAIRE recomenda as licenzas CC0 (dominio público) ou CC BY.

Haberá que especificar o uso ético dos datos respectando a lexislación sobre protección de datos e anonimizar os datos que sexan necesarios.

► INFORMACION SOBRE O ORZAMENTO.

A xestión dos datos de investigación implica unha serie de tarefas nas que se precisan soportes técnicos, programas para preparar os datos, política de conservación e reutilización, actuacións para a conservación dos datos a curto e longo prazo, etc. Todo isto terá que ser recollido nun orzamento para coñecer as necesidades económicas para garantir unha boa xestión de datos.

É interesante esta **guía para redactar un orzamento** realizada polo UK Data Service.

COMPARTIR E DIFUNDIR O PLAN

DMPONLINE permite engadir comentarios ao plan, compartilo entre os membros do grupo de investigación e asignar permisos para editar ou só de lectura. Pódese facer público dentro do propio programa.

Recomendamos descargar o plan nun editor de texto para poder editar e engadir por exemplo a filiación institucional de cada responsable. Tamén permitirá maquetar o documento cos logos dos financiadores e mellorar a súa presentación.

Podemos facer unha versión en pdf e difundilo en redes académicas (ResearchGate), redes sociais, páxinas web ou nun repositorio.

EXEMPLOS DE DMP DE H2020 ► **InteGRIDy Data Management Plan+** ► **DocksTheFuture**

GUIA MAIS COMPLETA. ► Science Europe. (2018). *Practical Guide to the International Alignment of Research Data Management* (D/2018/13.324/4). Recuperado de https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

Para calquera consulta sobre este tema lembra que podes contactar coa BUDC a través do correo electrónico **ruc@udc.es**.



Licenza CC BY 4.0 Internacional.
Elaborado por Carlota Balsa Sánchez.