I mercoledì di giugno dedicati a Open Science

Una serie di incontri su cosa significa produrre "Open Science" e che cosa implica nella pratica per ricercatori e istituzioni universitarie.

Iolanda Pensa, Alessandro Pierno

Come preparare un **Data Management Plan**

Incontri online

Mercoledì 03, 10, 17 e 24 giugno 2020, dalle 16:30 alle 17:30

Su: MSTEAMS Collegio di dipartimento DACD[1]

DATI DI RICERCA

COSA SONO?

I dati della ricerca **sono tutte le informazioni** che sono state raccolte, osservate, generate o create per convalidare i risultati originali della ricerca

@alespierno

CC BY

GESTIRE I DATI DI RICERCA (RDM)

CHE SIGNIFICA?

La **gestione dei dati di ricerca** (Research Data Management) è l'efficace gestione delle informazioni che si creano nel corso della ricerca. È parte integrante del processo di ricerca

@alespierno

CC BY

Research data lifecycle diagram ©Jisc CC BY-NC-ND

GESTIRE I DATI DI RICERCA

Un'efficace gestione dei dati viene effettuata per l'intero ciclo di vita dei dati stessi

CC BY

GESTIRE I DATI DI RICERCA

MA PERCHÉ DOVREI FARLO?

Impegnarsi nel processo di RDM presso la vostra istituzione può fornire **benefici**, a voi, ai vostri Studenti, ad altri Ricercatori, ai vostri Colleghi, alla vostra Istituzione

@alespierno

CC BY

GESTIRE I DATI DI RICERCA

MA PERCHÉ DOVREI FARLO?

- Per garantire l'integrità e la riproducibilità della ricerca;
- Per garantire che i dati della ricerca siano **autentici**, **completi** e **affidabili**;
- Per ridurre al minimo il **rischio di perdita** di dati;
- Per **evitare la duplicazione** degli sforzi permettendo ad altri di riutilizzare i vostri dati;
- Per soddisfare **i requisiti dei finanziatori (Horizon 2020)**
- Per migliorare **la collaborazione** con gli altri ricercatori
- Per ricevere il **riconoscimento** per output diversi dalla pubblicazione: *dataset*, policy, protocolli, metodologie, ecc.

CC BY

IL DATA MANAGEMENT PLAN

COS'È?

a) un piano formale di gestione dei dati che accompagna il lavoro di ricerca e descrive quali dati sono raccolti, processati e generati, dove vengono archiviati, a quali licenze e limitazioni sono legati, chi ha diritto al loro utilizzo ➔ **DATA**

b) un riferimento per gestire responsabilmente i dati e renderli Findable, Accessible, Interoperable, Reusable (FAIR) ➔ **MANAGEMENT**

c) un piano per programmare e comunicare la raccolta, conservazione, riuso e divulgazione dei dati + metadati associati, sin dall'inizio dell'attività ➔ **PLAN**

Che cos'è un Data Management Plan? © Occioni, Vignocchi CC BY

IL DATA MANAGEMENT PLAN

COS'È?

Un **Data Management Plan** è il documento formale che indica nel dettaglio come saranno gestiti i dati di ricerca

@alespierno

CC BY

IL DATA MANAGEMENT PLAN

A COSA SERVE?

a illustrare tutte le **modalità** e le **azioni** che devono essere intraprese dai ricercatori per una **corretta gestione** dei dati di ricerca per la loro **valorizzazione** e alla loro **preservazione** prima, durante e dopo la realizzazione di un progetto

CC BY

IL DATA MANAGEMENT PLAN

COSA CONTIENE

1. Dataset

l'identificazione dei dati, come sono e saranno raccolti, come saranno utilizzati

2. Standard e Metadati

chi li ha creati, chi li possiede, quando sono stati creati ecc.

3. Condivisione

Come sono reperibili e utilizzabili

4. Modalità di archiviazione e conservazione

CC BY

IL DATA MANAGEMENT PLAN

PERCHÉ DOVREI FARLO?

- per soddisfare **i requisiti** posti dagli enti finanziatori (e publisher)
- per aderire al **Codice di Condotta per l'Integrità della ricerca**
- per riconoscere le **problematiche tecniche e legali** fin dall'inizio del progetto di ricerca (open source, database, licenze, diritto d'autore)

CC BY

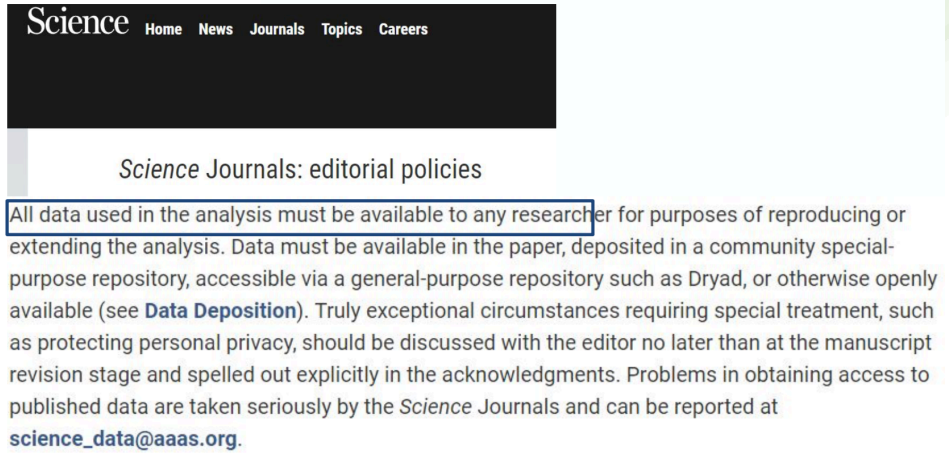
per soddisfare **i requisiti** posti dagli enti finanziatori

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

per aderire al **codice di condotta dell'integrità della ricerca**

http://www.allea.org/wp-content/uploads/2017/03/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017-1.pdf

per soddisfare **i requisiti** posti da alcuni publisher



CC BY

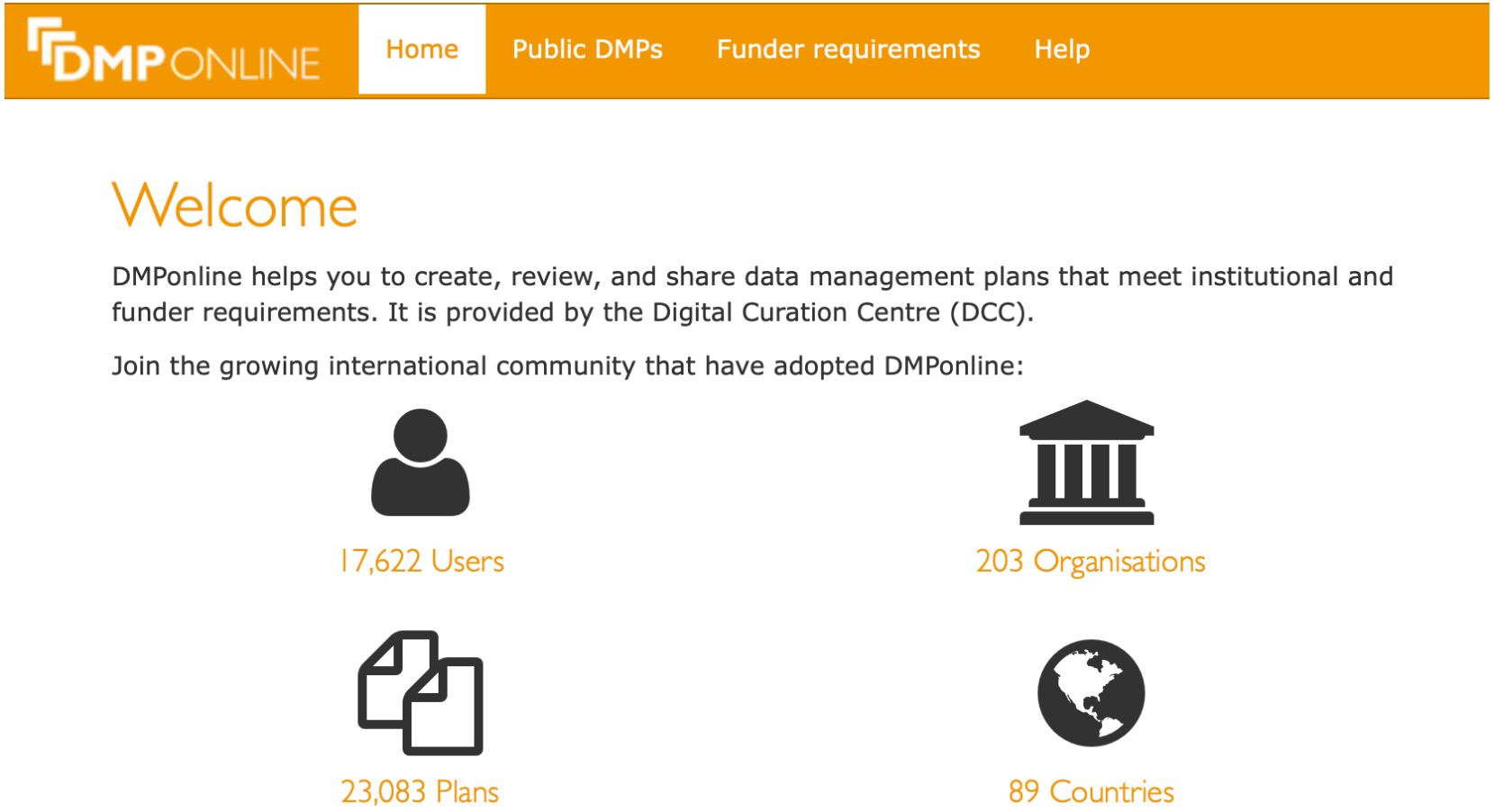
per riconoscere le **problematiche tecniche, legali**

1) quali diritti insistono sui dati?
2) un database è protetto dal diritto d'autore? E chi è il titolare?
3) Che regole ci sono nei casi di ricerca finanziata da soggetti terzi?
4) Chi è titolare dei diritti eventualmente sussistenti sulle banche dati nel caso di attività di ricerca svolta da gruppi provenienti da diversi atenei?
5) È necessario che l'ateneo regoli questi aspetti nei rapporti con i propri ricercatori?

Guarda, Paolo, & Galimberti, Paola. (2019, May 16). FAQ in materia di banche dati e diritti di proprietà intellettuale. Zenodo. http://doi.org/10.5281/zenodo.2875400

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

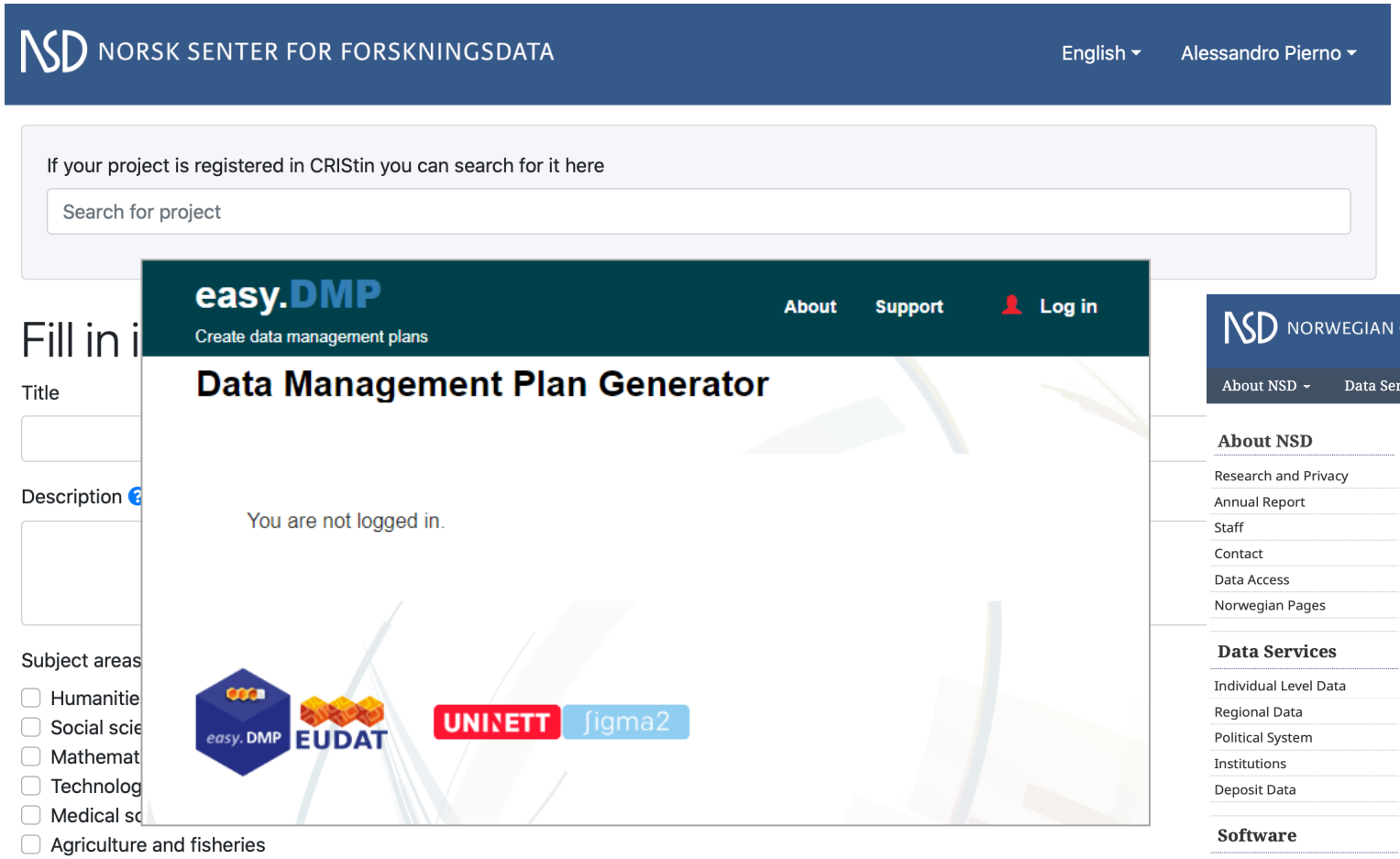


https://dmponline.dcc.ac.uk

TOOL realizzato dal **Digital Curation Centre** e messo a disposizione di tutti i ricercatori, che possono crearsi i loro *template* ed aggiungere il logo della propria istituzione

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP



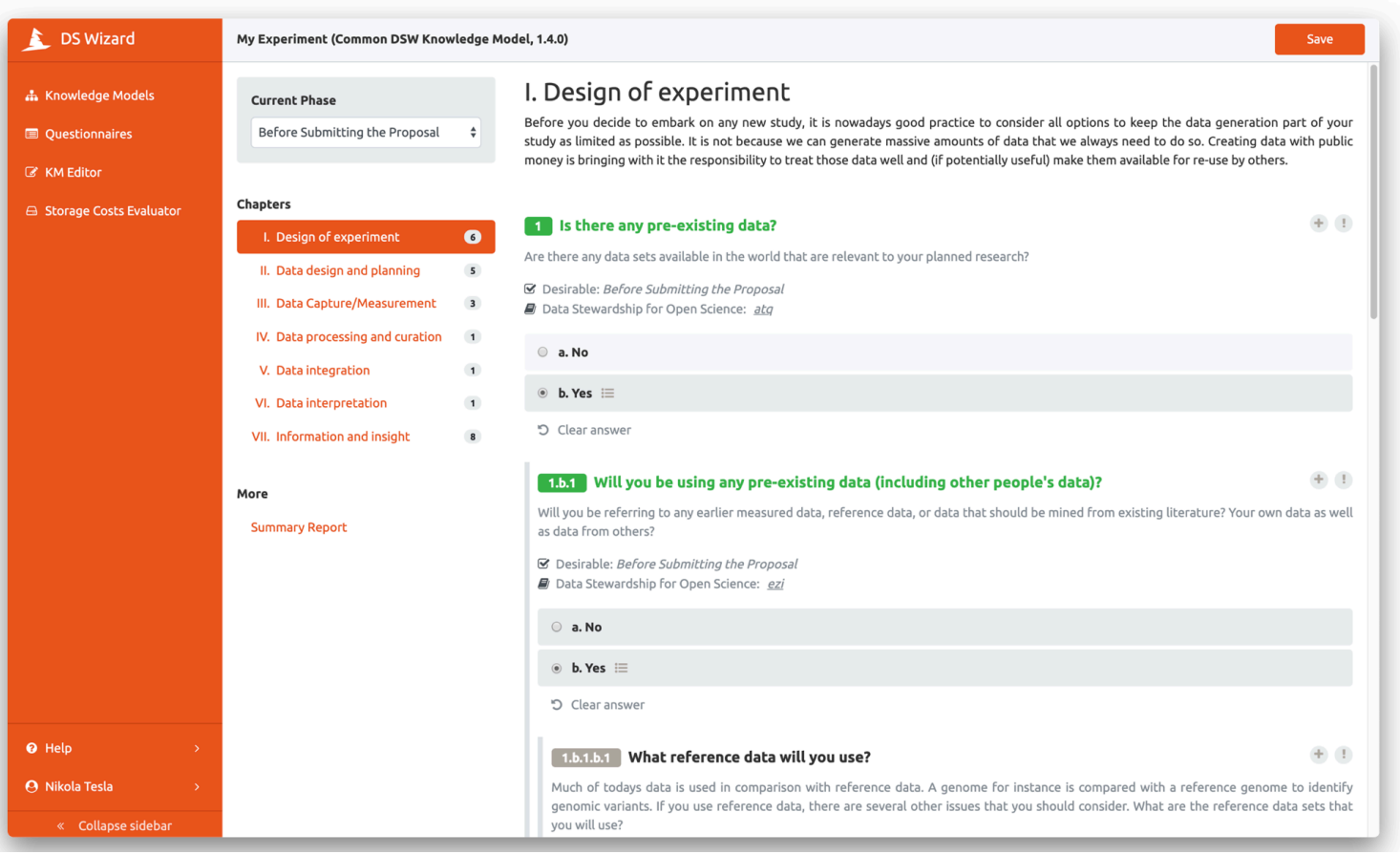
https://dmp.nsd.no

consorzio di università norvegesi
Creazione di Data Management Plan

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

DSW



https://ds-wizard.org

Data Stewardship Wizard

Creare piani di gestione intelligente dei dati per open science

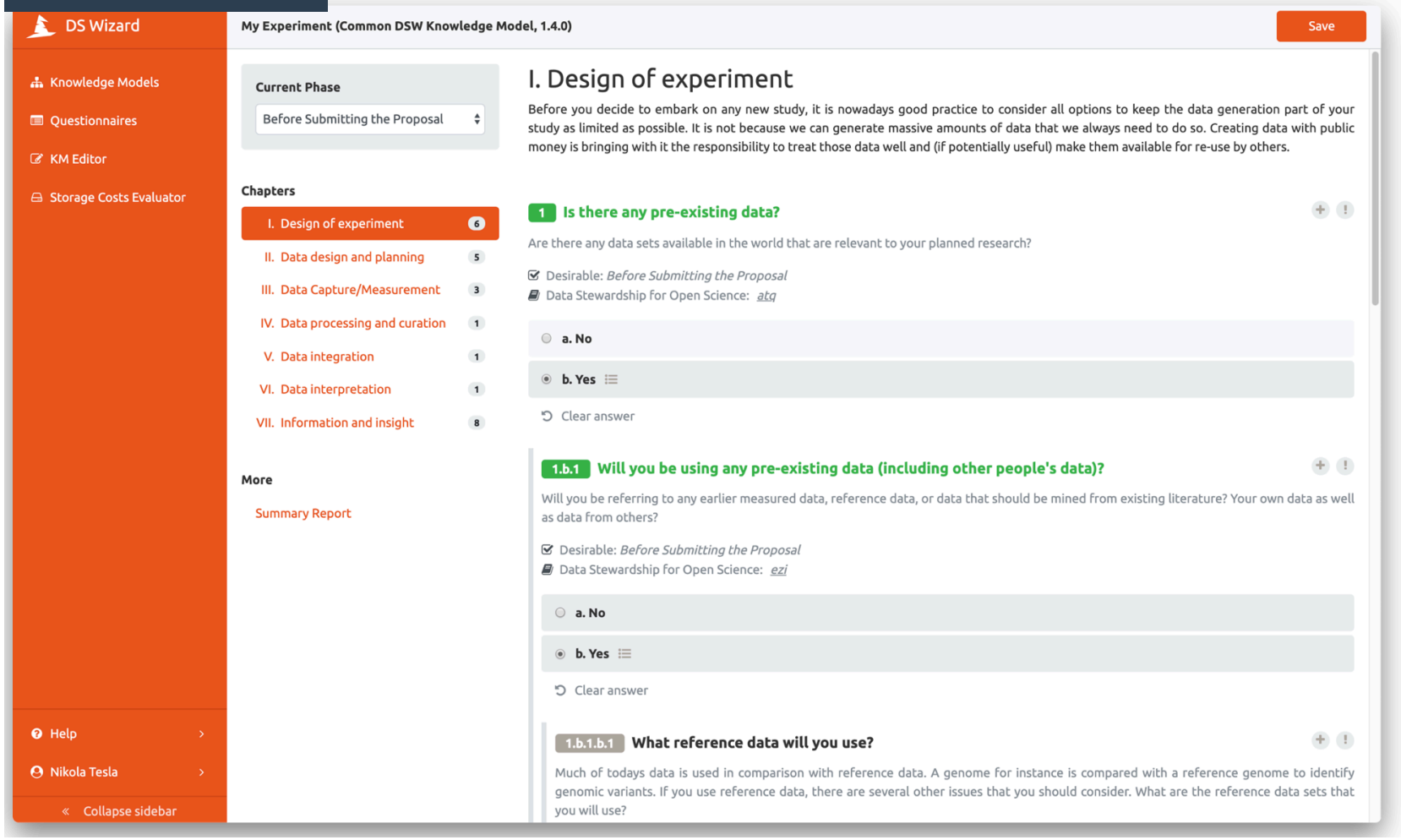
I Data Stewards acquisiscono e combinano le loro conoscenze e competenze rispetto alle esigenze specifiche di un dominio o di un'organizzazione. I ricercatori sono veramente guidati attraverso la composizione di un DMP che può essere poi esportato usando un modello e un formato selezionato, anche con l'ausilio di macchine che possono essere azionate.

@alespierno

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

FAIRDOM



https://fair-dom.org/knowledgehub/data-management-checklist/

Checklist che guida attraverso domande fondamentali che dovreste porvi durante la costruzione di un piano di gestione dei dati e del modello per la vostra proposta di sovvenzione.

La lista di controllo dei dati comprende l'integrazione dei requisiti del piano di gestione dei dati sulle sovvenzioni H2020 (PDF disponibile qui: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

Guida pratica alla gestione internazionale dei dati di ricerca

Sviluppata da esperti di Science Europe Member Organisations, questa guida mira ad allineare i requisiti della gestione dei dati di ricerca (research data management (RDM RDM) tra le varie organizzazioni di ricerca

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

GRIGLIA PER IL PIANO DI GESTIONE DEI DATI

*Gruppo di Lavoro su Dati della ricerca/Sottogruppo sul Data Management Plan
Marisol Occioni, Università di Venezia Ca' Foscari

ADMINISTRATIVE PLAN	Informazioni generali sul progetto di ricerca
Project Name	Inserire il nome del progetto
Acronimo	Inserire l'acronimo del progetto, se applicabile
Grant Reference Number	Inserire il riferimento alla call (es: call Horizon2020 ID:……) e al numero di identificazione del progetto presentato, se disponibile
Persistent Identifier	handle o DOI del DMP, ricavabile dopo l'inserimento nel repository
Funder	Inserire il nome del finanziatore/dei finanziatori Es: (H2020)
Principal Investigator/Researcher	Inserire il nome del ricercatore autore del documento
Principal Researcher ID ORCID	Inserire l'identificativo ORCID del ricercatore
DMP Creator	Nome della persona che ha generato il DMP/ ha compilato il form se diversa dal Principal Investigator/Researcher
Project Data Contact	Inserire istituzionali del ricercatore e del responsabile della gestione dei dati all'interno del progetto
Project Description	Breve descrizione del progetto, degli obiettivi e dei dati che si intendono raccogliere, memorizzare, utilizzare, produrre e distribuire
Version and Date of the DMP	Specificare la versione e la data del DMP, da aggiornare nel corso del progetto nel caso di variazioni significative (nuovi dataset, cambi di policy. Ultimo aggiornamento il giorno…

http://wikimedia.sp.unipi.it/images/Grigliapianodigestionedatiricerca.pdf

I. Sezione: Amministrativa

@alespierno

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

DATASET DESCRIPTION	Descrivere i dati esistenti o che si intendono creare, indicandone provenienza, natura e ordine di grandezza. Motivare la creazione di nuovi dataset e relativo valore aggiunto.
Provenance of data and metadata (content)	Descrivere se i dati provengono da interviste, indagini, se sono estratti da archivi disciplinari, banche dati e/o da altri progetti
Type of data	Descrivere se sono qualitativi, quantitativi
Nature and formats	Descrivere natura e formato dei dati (meglio se di formato non proprietario), ad esempio: documenti testo (DOC, ODF, PDF, TXT, etc); immagini (JPG, GIF, SVG, PNG, TIFF); video/film (MPEG, AVI, WMV, MP4); registrazioni audio (MP3, WAV, AIFF, OGG, etc); dati strutturati (HTML, JSON, TEX, XML, RDF); tabelle (CSV, ODS,TSV, XLS, SAS, Stata, SPSS portable); codici sorgente (C, CSS, JavaScript, Java, etc); configuration data (INI, CONF, etc) database (MS Access, MySql, Oracle)
Amount of data	Stimare l'ordine di grandezza dell'intero dataset (GB, MB, TB, PB)
Requirements for software e hardware	Specificare l'eventuale utilizzo di software o hardware specifici
Data re-use and integration	Nel caso di riutilizzo di dati, dichiarare la loro integrazione e riuso nel progetto, definire il contesto di riutilizzo, definire i destinatari. Pianificare se sono previste traduzioni
New Dataset value	Motivare la creazione del nuovo dataset, definendone il valore aggiunto per la comunità scientifica o altro
Audio-visual materials	In caso di video, segnalare la durata (il dato è utile per pianificare i costi di progettazione e archiviazione)
Software application	Descrivere se verranno prodotti software application per facilitare ricombinazioni di dataset di diversa origine

*Gruppo di Lavoro su Dati della ricerca/Sottogruppo sul Data Management Plan
Marisol Occioni, Università di Venezia Ca' Foscari

II. Sezione: Descrizione Dataset

@alespierno

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

STANDARDS AND METADATA	Motivare la creazione di nuovi dati, in termini contenuto e qualità. Descrivere i metadati e la documentazione accompagnatoria dei dataset, con riferimenti a standard di disciplina, interoperabilità e normativa
Metadata standards and data documentation	Descrivere il tipo di metadati con riferimento a standard (es: Dublin Core o Datacite Metadata Schema)
Methodologies for data collection/generation	Descrivere le metodologie di raccolta e produzione dei dati durante il processo di ricerca. • Chi e come raccoglie i dati; • Chi e come li struttura e memorizza; • Chi e come li elabora; • Chi e come li distribuisce.
Data quality and standards	• Documentare come garantire coerenza e qualità ai dati raccolti o prodotti in termine di standard, calibrazione, validazione, review • riferirsi a idonei standard disciplinari se esistenti: nel caso in cui allo stadio attuale le tipologie di dati non rientrino in standard, preventivarne la conversione; • in caso di riuso di metadati, dichiarare se ci sono vincoli al loro riutilizzo;
Metadata quality	La riusabilità dei dati è assicurata dalla qualità dei metadati a loro associati

*Gruppo di Lavoro su Dati della ricerca/Sottogruppo sul Data Management Plan
Marisol Occioni, Università di Venezia Ca' Foscari

http://wikimedia.sp.unipi.it/images/Grigliapianodigestionedatiricerca.pdf

III. Sezione: Standard e Metadata

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

DATA MANAGEMENT, DOCUMENTATION AND CURATION	**Descrivere in termini tecnici i processi adottati per la gestione, documentazione, cura e conservazione dei dati. È un Sommario. Stimare i costi associati alla gestione e mantenimento dei dati per rendere FAIR il progetto, facendo attenzione che i costi del primo anno saranno più alti.**
Managing, storing and curating data	Predisporre un **Preservation** Plan che descriva le modalità di: 1. memorizzazione; 2. backup; 3. trasmissione; 4. cura dei dati nel breve e medio periodo, con riferimenti a pratiche, standard e regolamenti ove applicabil
Preservation strategy and standards	1. Descrivere le modalità di conservazione a lungo termine dei dati in repositories certificati (Iris o altro repository di ateneo, di comunità, di dominio o specifici del progetto); 2. se la conservazione riguarda solo una parte dei dati raccolti, 3. dichiarare la durata di conservazione dei dati nel repository e l'eventuale cancellazione dopo un determinato periodo di tempo; 4. indicare i costi annui (definire da quando) per: • I repository, gli streaming services, altri servizi o infrastrutture • l'eventuale preparazione dei dati per l'archiviazione (costi da pareri legali sui diritti o soluzioni tecniche per la conversione) • il riuso/visualizzazione dei dati.

*Gruppo di Lavoro su Dati della ricerca/Sottogruppo sul Data Management Plan
Marisol Occioni, Università di Venezia Ca' Foscari

IV. Sezione: Management Curation

@alespierno

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

DATA SECURITY AND CONFIDENTIALLY OF POTENTIALLY DISCLOSIVE INFORMATION	E' la sezione più importante per i progetti di ricerca che raccolgono dati sensibili, anche se tutti i progetti devono prevedere informazioni sulla sicurezza delle procedure adottate per conservare, trasmettere e scambiare dati. Quanto descritto in questa sezione è soggetto alla validazione da parte del Comitato Etico
Confidential and Sensitive Data	Identificare se ci sono i dati di natura confidenziale / sensibili
Data security standards	Identificare formalmente gli standard (Es: ISO 27001) l'istituto aderisce
Main risks to data security	Descrivere i rischi per i partecipanti alla raccolta dei dati, in termini di confidenzialità e sicurezza. Ovvero, per ogni rischio, identificare: • Livello del rischio (basso, medio, alto); • Probabilità di occorrenza del rischio (bassa, media, alta); • L'impatto rispetto ai diversi attori coinvolti; • Le modalità di monitoraggio; • Le contromisure adottate in caso di occorrenza.
Ethics and Legal Compliance	Con riferimento all'identificazione dei rischi legati alla gestione di dati sensibili e di tutti i metadati a loro associati, specificare se la ricerca coinvolge cellule staminali embrionali umane, persone, animali, ambiente, salute, sicurezza nazionale eccetera

*Gruppo di Lavoro su Dati della ricerca/Sottogruppo sul Data Management Plan
Marisol Occioni, Università di Venezia Ca' Foscari

V. Sezione: Confidenzialità e sicurezza

@alespierno

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

DATA SHARING AND ACCESS	Descrivere quali dati, come e in che modo saranno condivisi e resi disponibili, le politiche di accesso ai repository, di trasmissione e circolazione dei dati.
Suitable for sharing	Motivare la condivisione dei dati e di tutti i metadati associati, la modalità di condivisione. Dichiarare: • a quali licenza sono legati i dati (Creative Commons,) • la possibilità di citazione per terzi • i software e i tools per il riuso • il repository in cui verranno depositati: Zenodo o Institutional Repository: IRIS (compatibile con l'Open Access Infrastructure for Research In Europe OpenAIRE) L'eventuale non condivisione dei dati e tutti i relativi metadati va dichiarata e circostanziata.
Discovery by potential users	Identificare in modo conciso quali utenti esterni potrebbero beneficiare della condivisione dei dati e di tutti i metadati e in che modo
Restrictions or delays to sharing,	Descrivere le restrizioni di accesso ai dati per problemi di confidenzialità, agreement, privacy, proprietà intellettuale. L'eventuale ricorso ad anonimizzazioni, aggregazioni, etc. ed embargo (tempi e modalità)
Regulation of responsibilities of users	Specificare quali utenti esterni sono soggetti ad accordi relativi alla diffusione dei dati.

*Gruppo di Lavoro su Dati della ricerca/Sottogruppo sul Data Management Plan
Marisol Occioni, Università di Venezia Ca' Foscari

VI. Sezione: Condivisione Accesso

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

*Gruppo di Lavoro su Dati della ricerca/Sottogruppo sul Data Management Plan
Marisol Occioni, Università di Venezia Ca' Foscari

RESPONSIBILITIES AND RESOURCES	Specificare le responsabilità a livello di progetto/Istituto. Se possibile individuare anche il nominativo del referente . Segnalare e giustificare inoltre eventuali risorse aggiuntive
Responsibility	specificare chi ha la responsabilità: • nella raccolta dei dati e creazione di metadati • nel garantire la qualità dei dati • nel garantire la sicurezza dei dati • nel garantire l'archiviazione e condivisione dei dati • nel garantire la conformità con la policy dei dati dell'istituto
Additional resources required to deliver the plan	specificare le richieste, possibilmente corredate da una stima dei costi per: - Specialist/Technical expertise, consulenza legale, consulenza tecnica (o costi per formazione dello staff) per la gestione dei dati e l'archiviazione a lungo termine, Hardware o Software, Costi relativi ai Data Repositoriy
Compliance with ethical and legal requirements	Dichiarare se il progetto è conforme con i requisiti legali e etici di istituto

VII. Sezione: Responsabilità e Risorse

CC BY

STRUMENTI UTILI PER LA REDAZIONE DI UN DMP

*Gruppo di Lavoro su Dati della ricerca/Sottogruppo sul Data Management Plan
Marisol Occioni, Università di Venezia Ca' Foscari

RELEVANT INSTITUTIONAL POLICIES ON DATA SHARING AND DATA SECURITY	Riferimenti alle politiche di Istituto
Research Data Policy	URL alla Policy
Open Access Repository and Scientific Publications	URL alla Policy sul repository di Ateneo
Where the DMP will be deposited	Indicare che il DMP verrà caricato in Iris o altro repository
How DMP will be accessible	Indicare come viene assicurato l'accesso al DMP
DMP preservation Responsibility	Indicare chi è responsabile per la conservazione del DMP

VIII. Sezione: Policy

CC BY

ALTRE RISORSE UTILI

Esempi di DMP

Guida agli esempi di DMPs del DCC
http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples

Catalogo DMP di LIBER (European Research Library Association):
https://libereurope.eu/dmpcatalogue/

DMPs in Zenodo:
https://zenodo.org/search?page=1&size=20&q=Data%20Management%20Plan

DMP Template Supsi
Template-DataMP_Suspi_da compilare.docx

CC BY

CALCOLARE I COSTI DI GESTIONE DEI DATI

I costi per la gestione e l'archiviazione dei dati variano e dipendono dal progetto, dal volume, dal dominio, dal livello di documentazione e dalla conservazione dei vostri dati.

Calcolo dei costi
Esistono check-list per determinare i costi della gestione dei dati che in genere sono coperti dal budget di progetto:

OpenAIRE Costing tool https://www.openaire.eu/how-to-comply-to-h2020-mandates-rdm-costs

UK Data Service Costing Data Management

https://zingtree.com/host.php?tree_id=511095771 - 1

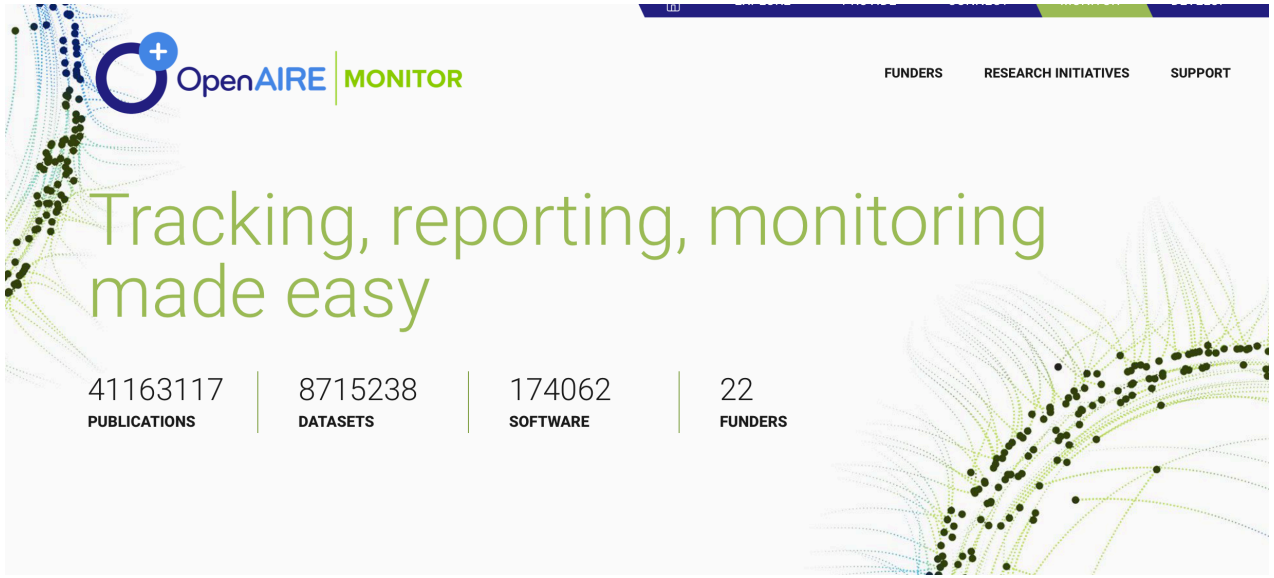
CC BY

OpenAIRE è la grande infrastruttura per il monitoraggio e la verifica delle politiche sull'Open Access finanziata dalla Commissione Europea.

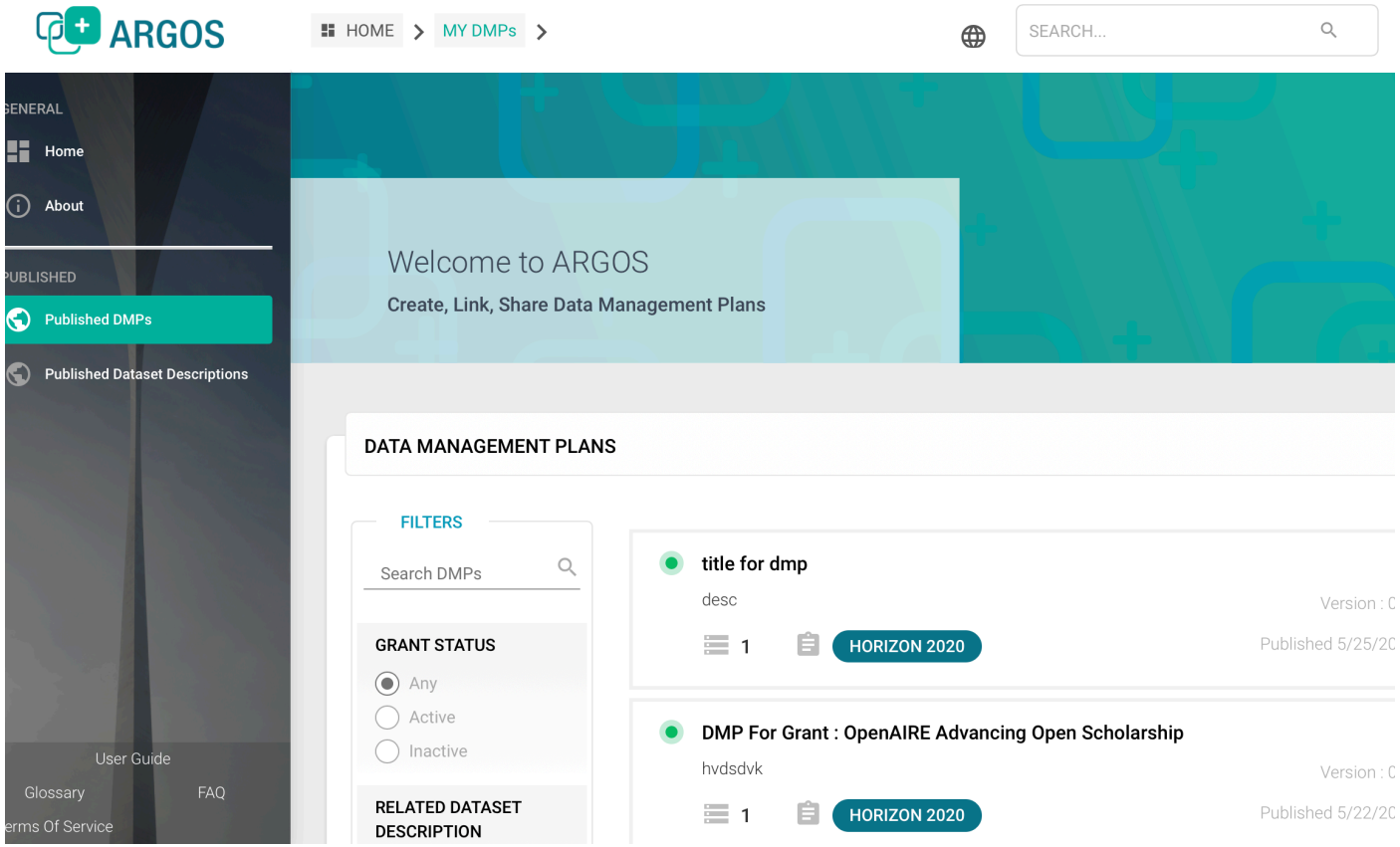
OpenAIRE collega i progetti ai risultati della ricerca (pubblicazioni, dati, software ecc) finanziata dalla Commissione Europea

https://www.openaire.eu

https://monitor.openaire.eu



CC BY



ARGOS è un servizio aperto che semplifica la gestione, la validazione, il monitoraggio e la manutenzione dei Data Managment Plan.

Permette di creare DMP azionabili e interoperabili

https://devel.opendmp.eu/explore-plans

CC BY