

Date de cercetare identificabile, accesibile, interoperabile și reutilizabile (FAIR)

*Identificarea cadrului general și dezvoltarea direcțiilor strategice
privind știința deschisă în România*

2022

Curprins

Accesul liber la datele de cercetare – Context	2
Datele deschise: element al politicii științei deschise.....	3
Principiile FAIR (Findable, Accessible, Interoperable, Reusable).....	5
Planul de Management al Datelor (Data Management Plan - DMP).....	11
Opțiuni strategice pentru managementul și deschiderea datelor de cercetare în România.....	14

Accesul liber la datele de cercetare – Context

Accesul liber la publicațiile științifice și la datele de cercetare reprezintă un concept bine aliniat la nivel european pentru a permite accesul la cercetarea finanțată din fonduri publice.

Intenția Comisiei Europene (CE) de a se dedica deschiderii datelor de cercetare a fost anunțată încă din iulie 2012, când, prin comunicarea intitulată „Spre un acces mai bun la informațiile științifice”¹, a anunțat că „va oferi un cadru și va încuraja accesul liber la datele de cercetare din Programul Orizont 2020”.

Această tendință a CE a fost încurajată de rapoarte de cercetare importante, cum ar fi „Riding the Wave: How Europe can gain from the rising tide of scientific data” din anul 2010 al Grupului de experți pentru date științifice a CE² urmat de raportul din anul 2014 al Research Data Alliance (RDA) intitulat „The Data Harvest: How sharing research data can yield knowledge, jobs and growth”³ care au prezentat o serie de recomandări de politici și chiar avertismente privind modul în care Europa trebuie să acționeze cât mai rapid pentru a gestiona și a împărtăși mai bine datele cercetării. Prin „Ghidul privind accesul liber la publicațiile științifice și datele de cercetare din Orizont 2020”⁴, Comisia Europeană admite că „informațiile deja plătite de către public nu ar trebui să fie plătite din nou de fiecare dată când acestea sunt accesate sau utilizate și de acestea ar trebui să beneficieze pe deplin companiile și cetățenii europeni”.

La nivel european, drumul către deschiderea datelor de cercetare rezultate din proiecte finanțate de către Comisia Europeană a fost pavat prin demararea exercițiului pilot al Programului Cadru Orizont 2020 „**Open Research Data Pilot (ORD pilot, 2017)**”⁵. În acest context, deschiderea datelor de cercetare a însemnat date care în general pot fi accesate, analizate, exploatate, reproduse și diseminate fără costuri percepute utilizatorilor. Participarea la programul pilot privind accesul liber la datele de cercetare a presupus crearea și actualizarea permanentă a unui plan de management al datelor de cercetare; încărcarea/depozitarea datelor de cercetare într-o arhivă/depozit „open access” și asigurarea accesului liber la date, acolo unde este posibil. Detaliile complete privind programul pilot pot fi regăsite în ghidul amintit mai sus.

¹ http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf

² <https://www.fosteropenscience.eu/content/riding-wave-how-europe-can-gain-rising-tide-scientific-data>

³ [https://www.rd-alliance.org/sites/default/files/attachment/The Data Harvest Final.pdf](https://www.rd-alliance.org/sites/default/files/attachment/The%20Data%20Harvest%20Final.pdf)

⁴ https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

⁵ https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Noul program de finanțare „Orizont Europa”, pune un accent deosebit și condiții clare pentru realizarea tranziției către „open science”, introducând **obligativitatea accesului liber la datele de cercetare prin intermediul unui depozit digital de încredere**, respectând principiul „cât mai deschis cu putință, dar atât de închis cât este necesar”⁶. Așadar, proiectele finanțate prin noul program au cerința obligatorie de realizare de planuri de management al datelor și să gestioneze în mod responsabil datele de cercetare respectând principiile „FAIR” (date identificabile, accesibile, interoperabile și reutilizabile). De asemenea, pentru apelurile care au prevederi referitoare la Cloud-ul European pentru Știință Deschisă (European Open Science Cloud - EOSC), obiectiv fundamental al CE în realizarea științei deschise, datele trebuie să fie depozitate în depozite de încredere care sunt integrate în EOSC în conformitate cu cerințele EOSC.

[Practical Guide to Sustainable Research Data](#) elaborat de către Science Europe vine în sprijinul organizațiilor de cercetare, organizațiilor de finanțare a cercetării și infrastructurilor de date de cercetare, în dezvoltarea strategiei lor în ceea ce privește datele de cercetare, pentru realizarea unui schimb de date sustenabil și al unor sisteme interoperabile. Ghidul este sub forma a trei matrici complementare pentru a permite colaborarea cu alte organizații. Matricile prezintă un cadru general și propun acțiuni, în șase domenii esențiale: implicare și angajament la nivel organizațional, cadrul privind politicile, aspecte financiare, training, pregătire tehnică și comunicare și creșterea gradului de conștientizare.

Datele deschise: element al politicii științei deschise

Directiva (UE) 2019/1024 a Parlamentului European și a Consiliului din 20 iunie 2019⁷ privind datele deschise și reutilizarea informațiilor din sectorul public definește datele deschise, datele provenite din cercetare și accesul liber:

Date deschise este înțeles, în general, ca desemnând datele în format deschis care pot fi liber utilizate, reutilizate și partajate de oricine în orice scop.

Date provenite din cercetare înseamnă documente în format digital care nu sunt publicații științifice și care sunt colectate sau produse în cursul activităților de cercetare științifică și sunt utilizate ca dovezi în procesul de cercetare sau sunt acceptate în mod curent în comunitatea de cercetare drept necesare pentru validarea concluziilor și rezultatelor cercetărilor. Acestea includ statistici, rezultatele experimentelor, măsurători, observații

⁶ Prezentare a Comisiei Europene din august 2019 August 2019 privind noul program cadru „Horizon Europe”, disponibilă la adresa:

https://ec.europa.eu/info/sites/info/files/research_and_innovation/strategy_on_research_and_innovation/presentations/horizon_europe_en_investing_to_shape_our_future.pdf

⁷ <https://eur-lex.europa.eu/legal-content/RO/TXT/PDF/?uri=CELEX:32019L1024&from=RO>

rezultate din munca de teren, rezultatele anchetelor, înregistrările interviurilor și imagini. Sunt incluse, de asemenea, metadatele, specificațiile și alte obiecte digitale.

Accesul liber este înțeles ca fiind practica de a furniza acces online la rezultatele cercetării cu titlu gratuit pentru utilizatorii finali și fără restricții de utilizare și reutilizare dincolo de posibilitatea de a solicita recunoașterea calității de autor și care contribuie la îmbunătățirea calității, la reducerea duplicării inutile a cercetărilor, la accelerarea progresului științific, la combaterea fraudelor în domeniul științei și poate în general favoriza inovarea și creșterea economică. Politicile privind accesul liber vizează în special să ofere cercetătorilor și publicului larg acces la datele provenite din cercetare cât mai devreme posibil în cadrul procesului de difuzare și să faciliteze utilizarea și reutilizarea acestor date⁸.

De asemenea, Directiva stabilește obligația ca statele membre să adopte politici privind accesul liber în ceea ce privește datele provenite din cercetare care beneficiază de finanțare publică și să se asigure că aceste politici sunt aplicate de toate organizațiile care desfășoară activități de cercetare și de toate organizațiile care finanțează activități de cercetare. Conform Recomandării (UE) 790/2018 privind accesul la informațiile științifice și conservarea acestora, politicile privind accesul liber urmăresc să ofere cercetătorilor și publicului larg acces gratuit la publicațiile științifice, la datele provenite din cercetare, în mod liber și nediscriminatoriu, cât mai rapid posibil în cadrul procesului de diseminare și să permită utilizarea și reutilizarea rezultatelor cercetării științifice.

Conform Directivei, în temeiul politicilor naționale privind accesul liber, datele provenite din cercetare care beneficiază de finanțare publică ar trebui să fie puse la dispoziție sub formă de date deschise în mod implicit. Totuși, în acest context, ar trebui să se țină seama în mod corespunzător de preocupările legate de viața privată, protecția datelor cu caracter personal, confidențialitate, securitatea națională, interesele comerciale legitime, cum ar fi secretele comerciale și drepturile de proprietate intelectuală ale unor părți terțe, conform principiului „cât mai deschis cu putință, dar atât de închis cât este necesar”.

Privire de ansamblu asupra evoluțiilor și inițiativelor de referință

2001: Creative Commons

2010: Berners-Lee five-star data

2011: Data Portals

Organizația internațională Open Knowledge lansează Data Portals – inițiativă care indexează toate portalurile de date deschise din lume. (2010) [<http://dataportals.org>]

2012: Open data: data.europa.eu lansat în 2012/

portalul oficial de date publice deschise al Uniunii Europene data.europa.eu (2012), Banca Mondială OpenData Initiative 2012, Open Data Institute

⁸ <https://eur-lex.europa.eu/legal-content/RO/TXT/PDF/?uri=CELEX:32019L1024&from=RO>

2018: Recomandarea (UE) 2018/790 a Comisiei din 25 aprilie 2018 privind accesul la informațiile științifice și conservarea acestora⁹

2019: Directiva (UE) 2019/1024 a Parlamentului European și a Consiliului din 20 iunie 2019 privind datele deschise și reutilizarea informațiilor din sectorul public¹⁰

2011 – România în Parteneriatul pentru o Guvernare Deschisă,
Parteneriatul pentru o Guvernare Deschisă – Open Government Partnership OGP (2011)

2013: lansarea portalului oficial de date deschise al României, data.gov.ro,

2014: licența pentru date deschise a României OGL-ROU 1.0

2015: Coaliția pentru Date Deschise publică un Ghid pentru Publicarea Datelor Deschise

Coaliția pentru Date Deschise este o platformă a societății civile, cu scopul de a sprijini și de a monitoriza implementarea politicilor publice privind guvernarea deschisă și datele deschise în România și de a informa asupra reutilizării datelor din sectorul public în scopuri necomerciale și comerciale. [<http://datedeschise.fundatia.ro/sample-page/>].

2016: Ordonanța de Urgență nr. 41/2016: Art. 11 din OUG nr. 41/2016: Instituțiile publice și organele de specialitate ale administrației publice centrale au obligația de a publica seturi de date de interes public pe portalul www.data.gov.ro. Procedura de publicare și de actualizare a acestor seturi de date se stabilește prin Hotărâre a Guvernului.

2022: Proiectul de Lege privind datele deschise și reutilizarea informațiilor din sectorul public este aprobat în luna ianuarie 2022 de către Guvernul României. Această lege transpune Directiva 2019/1024/UE a Parlamentului European și a Consiliului din 20 iunie 2019 privind datele deschise și reutilizarea informațiilor din sectorul public, publicată în Jurnalul Oficial al Uniunii Europene Seria L, nr. 172/56 din data de 26.06.2019. Directiva a actualizat cadrul legislativ european, astfel încât acesta să reflecte progresele din domeniul tehnologiilor digitale, stimulând totodată inovarea digitală, în special în ceea ce privește inteligența artificială.

Principiile FAIR (Findable, Accessible, Interoperable, Reusable)

În condițiile producerii unui volum tot mai mare de date în cercetare și a științei bazate tot mai mult pe tehnologie, utilizarea optimă a datelor și metodelor de cercetare este o provocare complexă¹¹, care trebuie abordată din diverse perspective: a cercetătorilor care doresc să-și facă cunoscute realizările, a revistelor și regândirii serviciilor oferite, a furnizorilor de software și instrumente de analiză și procesare a datelor, a organizațiilor de finanțare a cercetării preocupate de o mai bună administrare a datelor și de impactul cercetării, a

⁹ <https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A32018H0790>

¹⁰ <https://eur-lex.europa.eu/legal-content/RO/TXT/PDF/?uri=CELEX:32019L1024&from=RO>

¹¹ <https://www.force11.org/fairprinciples>

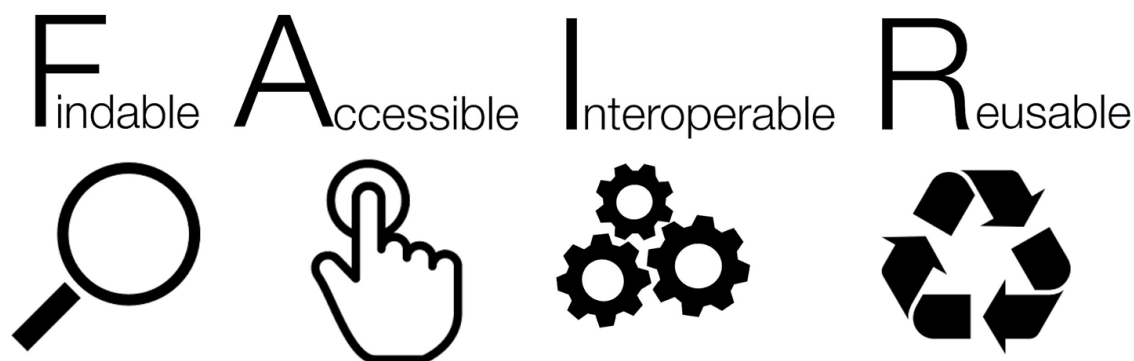
comunității științifice interesate de integrarea și analiza rezultatelor care pot conduce la descoperiri avansate și inovare.

Încercarea de a realiza analize și a stabili conexiuni, de a descoperi anumite tipare/ modele în volume mari de date interconectate a devenit uzuală în activitatea de cercetare. În acest context, furnizarea datelor de cercetare în format citibil de mașină („machine-readable”), precum și elaborarea unui set minim de principii în vederea facilitării accesului, interoperabilității și reutilizării datelor, agreate ca standard de comunitatea științifică, apare ca o necesitate și o provocare, în același timp.

În acest context a luat naștere conceptul FAIR încă din anul 2008, cu ocazia evenimentului „Jointly Designing a Data FAIRport” organizat de Barend Mons¹² împreună cu Lorentz Center, Dutch Techcenter for the Life Sciences și Netherlands eScience Center.

În anul 2016 „Principiile FAIR pentru managementul și administrarea datelor științifice” au fost publicate în Digital Science¹³, iar, începând de atunci, conceptul a fost adoptat și promovat intens la nivel european în vederea arhivării pe termen lung și reutilizării datelor de cercetare și a realizării Cloud-ului European pentru Știință Deschisă (EOSC).

Principiile descriu capacitatea sistemelor computaționale de a găsi, accesa, asigura interoperabilitatea și reutiliza datele fără sau cu o foarte mică intervenție a omului, în contextul creșterii volumului, complexității și vitezei cu care sunt produse datele”¹⁴.



Sursa: https://en.wikipedia.org/wiki/FAIR_data#/media/File:FAIR_data_principles.jpg

Cele 4 principii pot fi descrise succint astfel:

¹² Barend Mons este considerat creatorul conceptului FAIR, este în prezent director GoFAIR și presedinte CODATA (2018-2022); o prezentare recentă privind FAIR este disponibilă public la: <https://media.uc3m.es/video/5d25b90d8f4208bbb58b4567>

¹³ M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg et al., The FAIR guiding principles for scientific data management and stewardship, Scientific Data 3 (2016), 160018. doi:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).

¹⁴ <https://www.go-fair.org/fair-principles/>

1. F (Findable): Identificabile (ușor de găsit)

Primul pas în procesul de reutilizare a datelor este găsirea acestora. Metadatele și datele trebuie să fie ușor de accesat, atât pentru oameni, cât și pentru computer. Metadatele în format ușor de citit de mașină sunt foarte importante pentru găsirea automată a datelor și serviciilor¹⁵.

F1. Metadatele au alocări identificatori unici la nivel global și persistenți.

Identificatorii unici și persistenți înlătură ambiguitatea cu privire la datele publicate, alocând un identificator unic fiecărui element din metadate și fiecărui concept din setul de date. Multe arhive de date generează automat identificatori unici și persistenți la nivel global pentru seturile de date depozitate. Acești identificatori permit înțelegerea și interpretarea datelor și sunt esențiali pentru interoperarea om - mașina/computer, reprezentând un element absolut necesar implementării științei deschise. În plus, aceștia facilitează citarea sursei în procesul de reutilizare a datelor.

Identificatorii trebuie să îndeplinească două condiții:

- să fie unici la nivel global (nu pot fi reutilizați/relocați). Există diferite platforme/servicii care asigură generarea de identificatori și care folosesc algoritmi ce garantează unicitatea;
- să fie persistenți; este nevoie de timp și bani pentru a menține legăturile (link-urile) active, de aceea ele au tendința să devină inactive la un moment dat, însă platformele/serviciile specifice pentru generarea de identificatori asigură, într-o oarecare măsură, rezolvarea acestei probleme.

Exemple de platforme/servicii care furnizează identificatori unici și persistenți la nivel global:

- Identifiers.org – identificatori sub formă de URIs și CURIEs: <http://identifiers.org>
- Identificatori unici universali: https://en.wikipedia.org/wiki/Universally_unique_identifier
- Identificatori URLs persistenți: <http://www.purlz.org>
- Digital Object Identifier (DOI): <http://www.doi.org>
- Archival Resource Key: <https://escholarship.org/uc/item/9p9863nc>
- Research Resource Identifiers: <https://scicrunch.org/resources>
- Identificatori pentru organizații de finanțare a cercetării: <https://www.crossref.org/services/funder-registry/>
- Identificatori pentru organizații de cercetare: <https://www.grid.ac>

¹⁵Procesul de FAIR-ificare are 7 pași, mai multe informații pot fi găsite aici: <https://www.go-fair.org/fair-principles/fairification-process/>

F2. Datele sunt descrise prin intermediul metadatelor.

Metadatele ar trebui să fie cuprinzătoare și extensive, incluzând informații descriptive despre context, calitate și condiții sau caracteristici ale datelor. Argumentul din spatele acestui principiu este că datele ar trebui să poată fi găsite pe baza informațiilor din metadate, chiar în absența identificatoarelor.

F3. Metadatele includ, în mod clar și explicit, identificatorul datelor pe care le descriu.

Asocierea dintre fișierul cuprinzând metadatele și setul de date corespunzător trebuie să se facă explicit prin menționarea identificatorului în metadate.

F4. Metadatele sunt înregistrate și indexate în cadrul unei resurse care poate fi găsită la căutare.

Resursele digitale pot fi făcute cunoscute/găsite la căutare inclusiv prin indexare. De exemplu Google indexează pagini web astfel încât să poată fi găsite prin intermediul opțiunii de căutare Google. Pentru datele de cercetare această indexare trebuie însă să fie specifică¹⁶.

2. A (Accesible): Accesibile

După ce utilizatorul găsește datele care-i sunt necesare, trebuie să afle cum le poate accesa, în termeni de autentificare și autorizare, dacă este cazul.

A1. Metadatele pot fi regăsite utilizând un protocol de comunicare standardizat.

Acest principiu se concentrează asupra modului în care datele și metadatele pot fi preluate de la identificatorii lor în mod automat. Este necesar să se stabilească clar cine poate accesa datele și să se specifice cum se face acest lucru – de aceea trebuie gândite protocoale clare care să limiteze pe cât posibil intervențiile umane. În același timp, trebuie avut în vedere faptul că un acces securizat nu poate fi asigurat prin intermediul unui protocol în totalitate automatizat; de exemplu, pentru datele foarte sensibile se recomandă menționarea în metadate a informațiilor de contact ale persoanei cu care să se poată discuta accesul la date.

Mai multe informații: FAIR Accessor <https://peerj.com/preprints/2522v2/#supp-1>.

A1.1 Protocolul este deschis, fără costuri și interoperabil universal.

¹⁶ Mai multe informații la: <https://github.com/FAIRDataTeam/FAIRDataPoint/wiki>

Pentru maximizarea reutilizării datelor, protocolul trebuie să nu implice costuri și să poată fi folosit în mod liber și, prin urmare, să poată fi implementat global pentru a facilita găsirea datelor (orice persoană cu acces la internet să poată accesa cel puțin metadatele).

A1.2 Protocolul permite o procedură de autentificare și autorizare, acolo unde este necesar.

Accesibil în spiritul FAIR nu înseamnă neaparat deschis sau fără costuri. Mai degrabă, se referă la faptul că este necesar să se menționeze condițiile exacte în care pot fi accesate datele. Prin urmare, chiar și datele foarte bine protejate și private pot fi gestionate cu respectarea principiilor FAIR. Ideal este ca accesibilitatea să fie specificată în așa fel încât mașina să poată automat să înțeleagă cerințele și, fie să le execute în mod automat, fie să avertizeze utilizatorul cu privire la aceste cerințe. Este de preferat să se solicite utilizatorilor să-și creeze cont de acces pentru a utiliza o arhivă. În acest fel este posibilă identificarea posesorului sau contributorului pentru fiecare set de date și pot fi setate drepturi specifice de utilizare.

A2. Metadatele sunt accesibile chiar și atunci când datele nu mai sunt disponibile.

Seturile de date au tendința de a se degrada sau dispărea în timp deoarece există costuri pentru ca sursele de date să fie menținute *online*. Când se întâmplă acest lucru, *link*-urile nu mai sunt valide, iar utilizatorii pierd timp căutând date care nu mai sunt disponibile. Stocarea metadatelor este mai simplă și mai ieftină.

Metadatele au propria valoare mai ales în stadiul de planificare a cercetării, în special în cazul repetării studiului. Chiar dacă datele inițiale nu mai există sau nu mai sunt disponibile, devine posibilă găsirea persoanelor, instituțiilor sau publicațiilor asociate cu cercetarea originală.

3. I (Interoperable): Interoperabile

De obicei, datele trebuie să fie integrate împreună cu altele și să relaționeze cu aplicații sau fluxuri de date în vederea realizării de analize, stocare și procesare.

I1. Metadatele folosesc un limbaj formal, accesibil și aplicabil pe scară largă pentru reprezentarea cunoștințelor.

I2. Metadatele folosesc vocabulare care respectă principiile FAIR.

I3. Metadatele includ referințe la alte metadate.

Mai multe: https://en.wikipedia.org/wiki/Programming_language.

4. R (Reusable): Reutilizabile

Scopul principiilor FAIR este să optimizeze reutilizarea datelor. Pentru arhivarea acestora, metadatele și datele trebuie să fie bine descrise pentru a putea fi replicate și combinate în diferite cercetări viitoare. De aceea, managementul datelor de cercetare trebuie să înceapă din stadiul de planificare a proiectului prin realizarea de planuri de management al datelor (*Data Management Plans* – DMPs) cu respectarea principiilor FAIR.

[RDA FAIR Data Maturity Model Working Group](#) oferă o listă detaliată adnotată de indicatori de abordat în vederea creșterii “FAIRness of data”.

Metadatele pot fi publicate folosind formate sau standarde cunoscute cum ar fi DCAT, DCAT-AP, ADMS, attribute CKAN, schema INSPIRE, DCIP sau VoID.

Portalul data.gov.ro folosește standardul DCAT-AP. Acesta este un standard mai ușor de folosit, bazat pe DCAT („Data Catalogue”) și creat în mod specific pentru seturile de date publicate de sectorul public din Europa. DCAT-AP ajută la expunerea seturilor de date deschise, prin posibilitatea de a fi găsite din orice portal de date european care folosește același protocol.

Licențe Creative Commons

Publicarea în „Open Access” se face de regulă prin adoptarea unui regim de licențiere de tip Creative Commons BY (CC BY), care este cea mai răspândită formă, urmată de CC BY-NC și CC BY NC-ND¹⁷. Semnificația acestor licențe este următoarea:

- CC BY: Oricine este liber să copieze și să distribuie lucrarea în orice mediu cu atribuirea corectă a sursei. De asemenea, este liber să adapteze materialul publicat în orice formă și pentru orice scop, inclusiv comercial¹⁸.
- CC BY-NC: Licența este identică cu cea anterioară, dar nu conține și posibilitatea folosirii în scopuri comerciale¹⁹.
- CC BY NC-ND: Această licență păstrează din CC BY doar posibilitatea de copiere și distribuție²⁰.

Identificatori personali și de date:

- Localizator uniform de resurse (URI).
- Digital Object Identifier (DOI).

¹⁷ Katie Shamash “Article processing charges (APCs) and subscriptions - Monitoring open access costs” Jisc, May 2016

¹⁸ <https://creativecommons.org/licenses/by/2.0/>

¹⁹ <https://creativecommons.org/licenses/by-nc/2.0/>

²⁰ <https://creativecommons.org/licenses/by-nc-nd/2.0/>

Planul de Management al Datelor (Data Management Plan - DMP)

Planul de management al datelor (DMP) reprezintă o piatră de temelie pentru gestionarea responsabilă a rezultatelor cercetării, în special a datelor și sunt obligatorii în programul Orizont Europa pentru proiectele care generează și/sau reutilizează date. DMP - ul abordează aspecte importante privind managementul datelor de cercetare, jucând un rol cheie în a ajuta cercetătorii să gestioneze în mod adecvat datele de cercetare în conformitate cu principiile FAIR.

În cadrul programului Orizont Europa, DMP - ul este definit ca un document care subliniază încă de la începutul proiectului principalele aspecte ale ciclului de viață al rezultatelor cercetării, incluzând în special datele. Acesta include informații privind proveniența, organizarea și pregătirea („curation”) acestora, precum și informații în ceea ce privește accesul, conservarea/păstrarea, partajarea și eventuala ștergere a acestora, atât în timpul cât și după finalizarea proiectului.

În plus, redactarea unui DMP este considerată o activitate direct legată de metodologia cercetării, adică un bun management al datelor va eficientiza munca/economisi timp, va contribui la protejarea informațiilor și la creșterea impactului și a valorii datelor în rândul beneficiarilor și al altor actori interesați, în timpul și după finalizarea cercetării. Astfel, DMP - ul este un mijloc cheie de sprijin atunci când se planifica și se implementează un proiect de cercetare și, în mod ideal, iar completarea acestuia ar trebui inițiată înainte de începerea proiectului.

Un plan de management al datelor ar trebui să fie un document „viu”, livrabil, parte a unui proces care începe în momentul planificării unui proiect de cercetare și este actualizat atunci când este necesar, pe măsură ce proiectul progresează. Un DMP cuprinde, în general, descrierea modului în care toate tipurile de date (brute și procesate) sunt colectate, utilizate și stocate într-un proiect, cine sunt deținătorii datelor și cine are, în final, responsabilitatea pentru colectarea, prelucrarea și arhivarea lor.

O bună practică menționată în cadrul programului Orizont Europa în ceea ce privește DMP-urile este de a le face livrabile publice, nerestricționate pentru a fi accesibile în mod liber, cu excepția cazului în care există motive legitime pentru a le păstra confidențiale. O altă bună practică este publicarea DMP-urilor în reviste specializate sau platforme de publicare, cum ar fi [RIO](#) etc., sau depozitarea lor în depozite publice specifice DMP, cum ar fi [DMPOnline](#) și altele.

Există în prezent resurse (informații și instrumente cu acces liber) pentru realizarea acestor planuri și pentru arhivarea pe termen lung a datelor. Instrumente precum **DMPonline**²¹ al Digital Curation Centre ajută la scrierea de astfel de planuri, punând la dispoziție diferite modele (templates) de planuri de management al datelor realizate după criteriile a diferite agenții de finanțare (printre care și model cu cerințele programului Orizont 2020 și Orizont Europa). De asemenea, prin intermediul platformei DMPonline pot fi create planuri personalizate și salvate în diferite formate. Un alt instrument online este [ARGOS](#) dezvoltat de OpenAIRE. [Data Stewardship Wizard](#), un instrument comun al infrastructurilor ELIXIR CZ și ELIXIR NL, îi ajută pe cercetători să înțeleagă ce este necesar pentru administrarea datelor astfel încât acestea să fie FAIR și să-și construiască propriile planuri de management al datelor.

Template - ul general utilizat în proiectele finanțate din programul Orizont 2020 conține²²:

- Sumarul datelor;
- Detalii privind organizarea datelor respectând principiile FAIR;
- Informații privind costurile și resursele;
- Informații privind securitatea datelor;
- Aspecte etice.
- Versiunea îmbunătățită și actualizată a template – ului menționat mai sus, utilizată în proiectele finanțate din programul Orizont Europa conține și o secțiune care se referă la²³: alte rezultate ale cercetării - pe lângă managementul datelor, beneficiarii ar trebui, de asemenea, să ia în considerare și să planifice managementul altor rezultate ale cercetării care pot fi generate sau reutilizate pe parcursul proiectelor lor. Astfel de rezultate pot fi fie digitale (de exemplu, software, fluxuri de lucru, protocoale, modele etc.) sau fizice (de exemplu, materiale noi, anticorpi (“antibodies”), reactivi („reagents”), probe etc.);

De asemenea, [Practical Guide to the International Alignment of Research Data Management - Extended Edition](#) elaborat de către Science Europe prezintă cerințele cele mai importante care trebuie avute în vedere la elaborarea și evaluarea DMP.

Depozitele digitale (“repository”) de încredere joacă un rol central în ceea ce privește depozitarea datelor de cercetare pe termen lung și asigurarea accesului liber la aceste date cât mai curând posibil și în termenele stabilite în DMP, prin aplicarea licențelor deschise.

Accesul liber ar trebui să fie implicit pentru datele de cercetare în conformitate cu principiul „cât mai deschis cu putință, dar atât de închis cât este necesar”. Aceasta înseamnă că, prin

²¹ <https://dmponline.dcc.ac.uk/>

²² http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

²³ <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/how-to-participate/reference-documents:programCode=HORIZON>

excepție, beneficiarii pot sau trebuie să păstreze anumite date închise din motive justificate; beneficiarii trebuie să explice în DMP excepția (excepțiile) în temeiul căreia aleg sau trebuie să restricționeze accesul la unele sau la toate datele de cercetare.

În ceea ce privește selectarea unui depozit digital de încredere pentru arhivarea datelor, în general, ar trebui urmărită îndeplinirea următoarelor criterii minime²⁴:

1. Furnizarea de identificatori persistenți și unici („Persistent and Unique Identifiers-PID”):
 - Permite identificarea datelor (chiar și dacă datele în sine dispar, referința la ele se păstrează cu ajutorul metadatelor);
 - Permite căutarea, citarea și găsirea cu ușurință a datelor;
2. Metadata utilizând protocoalele acceptate pentru „open access” și „FAIR Data”:
 - Asigură recuperarea metadatelor în mod automat;
 - Permite găsirea datelor și referințe la informațiile relevante, cum ar fi date și publicații conexe;
 - Furnizează informațiile care sunt disponibile și stocate public, chiar și pentru datele nepublicate, protejate, retrase sau șterse;
 - Utilizează metadata standardizate care sunt acceptate de către comunitatea științifică.
3. Accesul la date și licențe de utilizare
 - Asigură confidențialitatea și respectarea drepturilor de proprietate intelectuală ale autorilor;
 - Permite accesul la date în condiții bine specificate: informații despre licențe și condițiile de utilizare;
 - Asigură autenticitatea și integritatea datelor.
4. Stocarea datelor
 - Asigură păstrarea pe termen lung a metadatelor și a datelor;
 - Asigură transparența în ceea ce privește misiunea și politicile de păstrare a datelor (inclusiv sustenabilitatea financiară și tehnică, perioada de stocare a datelor, ș.a.).

Un registru cunoscut și de încredere al depozitelor dedicate datelor de cercetare cu acces liber este registrul [re3data](#) care oferă posibilitatea găsirii unui depozit adecvat, general sau specific unei discipline. Un alt registru al depozitelor dedicate datelor de cercetare cu acces liber este [OpenDOAR](#).

Exemple de arhive/depozite cu scop general pentru rezultatele cercetării multidisciplinare, inclusiv date, software și publicații:

- www.zenodo.org - depozit de uz general pentru date, software și publicații;

²⁴ http://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

- <https://figshare.com/> – depozit pentru orice rezultate de cercetare (pentru toate formatele de fișiere).

Exemple de arhive/depozite specifice unei discipline:

- [ELIXIR Deposition Databases](#) și [ELIXIR Core Data Resources](#) (depozite pentru arhivarea datelor experimentale din științele vieții).

Opțiuni strategice pentru managementul și deschiderea datelor de cercetare în România

- Elaborarea de politici privind conservarea rezultatelor științifice de către instituțiile academice și de cercetare care beneficiază de finanțare din fonduri publice;
- Stabilirea unor politici instituționale pentru gestionarea datelor cercetării, precum și elaborarea unor planuri de punere în aplicare;
- Asigurarea accesului liber la datele provenite din cercetare conform principiului „cât mai deschis cu putință, dar atât de închis cât este necesar” pentru proiectele de cercetare care produc/generează și reutilizează date de cercetare și în conformitate cu principiul gestionării responsabile a datelor de cercetare, începând cu noul ciclul de finanțare:
 - Date FAIR - ușor de identificat, accesibile, interoperabile și reutilizabile conform ghidurilor și protocoalelor europene (utilizând identificatori persistenți și unici²⁵ precum Digital Object Identifier; identificatorii persistenți asigură că seturile de date sunt ușor de identificat și pot fi legate de alte seturi de date și publicații prin mecanisme adecvate);
 - Date de cercetare disponibile sub protecția licențelor deschise (de preferat Creative Commons Attribution International Public License – CC BY SA 4.0 sau Creative Commons Public Domain Dedication (CC0) sau o licență cu drepturi echivalente) care să îndeplinească cerințele definite prin Declarația de la Berlin²⁶;
 - Metadatele aferente datelor depozitate - deschise în temeiul licenței Creative Commons Public Domain Dedication (CC 0) sau echivalentă (în măsura în care interesele sau constrângerile legitime sunt protejate), în conformitate cu principiile FAIR (în special care pot fi acționate automat – „machine-actionable”) și care să ofere informații cel puțin despre următoarele: seturi de date (descriere, data depozitării, autor(i), locul și embargo (dacă există); finanțarea; denumirea, acronimul și numărul proiectului; licența; identificatori

²⁵ http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?action=display&doc_id=4607

²⁶ Declarația de la Berlin <https://openaccess.mpg.de/Berlin-Declaration>

persistenți pentru setul de date, autorii implicați în acțiune și, dacă este posibil, pentru organizațiile lor și pentru grantul obținut.

- Obligatorietatea elaborării și actualizării în mod regulat a planurilor de management al datelor (DMP) care să abordeze aspecte importante privind managementul datelor de cercetare, cu respectarea principiilor “FAIR” și ale accesului liber la date („open data”), în cadrul proiectelor de cercetare, dezvoltare și inovare finanțate din fonduri publice. De asemenea, susținerea infrastructurilor de cercetare pentru elaborarea și implementarea planurilor de management al datelor asociate experimentelor;
- Publicarea DMP, cu excepția cazului în care există motive (interese legitime sau alte constrângeri) pentru a nu face acest lucru. În cazul în care acestea sunt făcute publice, se recomandă, de asemenea, ca accesul liber să fie oferit prin aplicarea licenței CC BY pentru a permite o reutilizare largă;
- Costurile asociate managementului datelor provenite din cercetarea finanțată din fonduri publice considerate eligibile prin contractele de finanțare;
- Arhivarea/depozitarea datelor într-un depozit digital („repository”) de date de încredere „open access” și asigurarea accesului liber, cât mai curând posibil și în termenele stabilite în DMP. Accesul liber ar trebui să fie implicit pentru datele de cercetare în conformitate cu principiul „cât mai deschis cu putință, dar atât de închis cât este necesar”. Aceasta înseamnă că, prin excepție, beneficiarii pot sau trebuie să păstreze anumite date închise din motive justificate; beneficiarii trebuie să explice în DMP excepția (excepțiile) în temeiul căreia aleg sau trebuie să restricționeze accesul la unele sau la toate datele de cercetare. Aceste excepții pot fi: dacă acordarea accesului liber este împotriva intereselor legitime ale beneficiarului, inclusiv în ceea ce privește exploatarea comercială; dacă este contrar oricăror altor constrângeri, cum ar fi normele de protecție a datelor, confidențialitatea, secretele comerciale, regulile de securitate, drepturile de proprietate intelectuală;
- Promovarea și sprijinirea inițiativelor existente și a altora noi care susțin știința deschisă, în special pentru depozitarea pe termen lung a datelor provenite din cercetare, în vederea integrării în baze de date disciplinare și/sau în EOSC;
- Furnizarea de îndrumări cercetătorilor cu privire la modul în care trebuie să se conformeze politicilor de management al datelor științifice și sprijinirea lor în acest sens, în special în ceea ce privește dezvoltarea unor competențe solide privind planificarea managementului datelor și a infrastructurilor digitale care sprijină accesul la datele cercetării și păstrarea acestora.

Informații suplimentare:

- 'Implementing Research Data Management Policies Across Europe – Experiences From Science Europe Member Organisations', Science Europe, ianuarie 2020: https://www.scienceeurope.org/media/jikjlb2g/se_rdm_best_practices.pdf
- 'Practical Guide to the International Alignment of Research Data Management': D/2018/13.324/4, Science Europe, noiembrie 2018: https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf
- What is a Data Management Plan (DMP) and how do I create one? <https://www.openaire.eu/what-is-a-data-management-plan-and-how-do-i-create-one>
- Recomandări RDA & CODATA Legal Interoperability Of Research Data: Principles And Implementation Guidelines: <https://www.rd-alliance.org/group/rdacodata-legal-interoperability-ig/outcomes/rda-codata-legal-interoperability-research-data>
- FAIR Data in Trustworthy Data Repositories Webinar (DANS/EUDAT/OpenAIRE Webinar - Dec. 2016) <https://www.openaire.eu/fair-data-in-trustworthy-data-repositories-webinar-dans-eudat-openaire-webinar-dec-2016>
- Open Research Data in H2020 and Zenodo repository (OAWEEK 2016 webinar - Oct. 2016): <https://www.openaire.eu/open-research-data-in-h2020-and-zenodo-repository-oaweek-2016-webinar-oct-2016>
- Open research data and data management for Horizon 2020 projects (16 June 2016): <https://www.openaire.eu/open-research-data-and-data-management-for-horizon-2020-projects>
- Webinars on How to write a Data Management Plan - Joint EUDAT-OpenAIRE webinar (July 2016): <https://www.openaire.eu/webinars-on-how-to-write-a-data-management-plan-joint-eudat-openaire-webinar-july-2016>
- Research Data Management: An introductory Webinar - Joint OpenAIRE and EUDAT (May 2016): <https://www.openaire.eu/research-data-management-an-introductory-webinar-from-openaire-and-eudat>
- DMPonline pentru diferite țări: <https://github.com/DMPRoadmap/roadmap/wiki/Installations-inventory>
- Ghidul și template-ul DMP furnizat de Comisia Europeană pentru datele proiectelor H2020: http://ec.europa.eu/research/participants/data/ref/h2020/gm/reporting/h2020-tpl-oa-data-mgt-plan_en.docx
- Recomandările și protocoalele OpenAIRE (<https://guidelines.openaire.eu/en/latest/>) și OAI-PMH (protocol internațional pentru metadata: https://guidelines.openaire.eu/en/latest/data/use_of_oai_pmh.html, <https://www.openarchives.org/pmh/>)
- Metodologie pentru publicarea datelor deschise: http://ogp.gov.ro/nou/wp-content/uploads/2019/02/Metodologie-date-deschise_feb2019.pdf
- Inventar al datelor colectate/produse/gestionate de structurile/departamentele instituției (modelul în format editabil se regăsește la adresa: <http://ogp.gov.ro/wpcontent/uploads/2018/06/Anexa-model-inventar.xlsx>)
- Planul de publicare a seturilor de date deschise - modelul în format editabil se regăsește la adresa: <http://ogp.gov.ro/wp-content/uploads/2018/06/Anexa-modelplan-de-publicare.xlsx>