

RECOMANDĂRI PRIVIND SCHIMBUL DE DATE DE CERCETARE - FACILITAREA ACCESULUI LIBER LA REZULTATELE CERCETĂRILOR ÎN CONTEXTUL PANDEMIEI COVID-19



OPEN SCIENCE HUB ROMANIA - UEFISCDI
CATEDRA UNESCO PENTRU POLITICI ÎN ȘTIINȚĂ ȘI INOVARE



Acest document a fost elaborat de echipa Open Science Hub Romania împreună cu Catedra Unesco pentru Politici în Știință și Inovare din cadrul SNSPA și este pus la dispoziție sub [Licența Creative Commons Atribuire 4.0 Internațional](#) (CC BY 4.0).

În cazul în care doriți mai multe informații, ne puteți contacta la următoarele adrese:

Contact Open Science Hub Romania: open-science@uefiscdi.ro;

Contact principal din partea Catedrei Unesco pentru Politici în Știință și Inovare: Alina Irimia, open4science@gmail.com

Acest document are rolul de a oferi o sinteză de informații, referințe principale și contacte privind schimbul de date în contextul COVID-19 - ce pot fi de folos cercetătorilor români care vor să contribuie, la rândul lor, la colectarea și împărtășirea rezultatelor din știință și să beneficieze, totodată, de accesul deschis la resurse din cercetarea europeană și globală asociată pandemiei curente.

Având în vedere dinamica cercetării asociate SARS-CoV-2 și pandemiei generate de acest coronavirus, documentul poate suferi actualizări.

(Versiune 1, iunie 2020).

Informații cuprinse în document sunt colectate ca urmare a conexiunii și implicării în inițiative europene cu rol principal în sprijinirea științei deschise la nivel internațional:



<https://www.openaire.eu/advance>



<https://ni4os.eu>



<https://grants.rd-alliance.org/about-rda-eu-40>



<https://www.scienceeurope.org>

Cuprins

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|
| Context..... | 3 |
| Platforma Europeană COVID-19 Data Platform..... | 4 |
| Ghiduri și recomandări pentru schimbul liber, rapid și sigur de date de cercetare relevante pandemiei actuale - <i>Research Data Alliance (RDA) Covid-19 Working Group</i> | 5 |
| Depozitarea și schimbul datelor – pe scurt..... | 6 |
| Pregătirea datelor | 6 |
| FAIR Data..... | 6 |
| Planurile de management de date..... | 6 |
| Depozitarea datelor | 7 |
| Standarde pentru date..... | 8 |
| Alte resurse cu acces liber | 9 |

Context

Comunitatea de cercetare globală are actualmente principalul rol de a răspunde provocărilor generate de virusul SARS-CoV-2. Organizația Mondială a Sănătății a activat „[The R&D Blueprint](#)” pentru a accelera diagnosticul și descoperirea de vaccinuri și tratamente pentru acest nou virus, iar cercetătorii din întreaga lume au răspuns imediat chemării de a se implica pentru a găsi soluții. Munca acestora se poate realiza eficient dacă datele sunt distribuite cât mai rapid, liber (deschis) și organizat. Astfel, în contextul unei provocări noi pentru omenire, schimbul de date în știință (*data sharing*) își dovedește reala utilitate și trebuie asigurat cu prioritate. Există o practică și ghiduri care pot fi utilizate/adaptate pentru a realiza acest schimb rapid și în siguranță – ele se regăsesc în munca desfășurată în ultimii ani în sfera științei deschise (open science). Astfel, Comisia Europeană a lansat un apel către inițiativele open science pe care le finanțează de a-și concentra eforturile în prezent pentru a susține acest schimb eficient de date de cercetare necesare combaterii pandemiei COVID-19 și găsirii de soluții de răspuns în fața situațiilor create de acest virus.

La nivel național, Ministerul Educației și Cercetării (MEC) a prevăzut fonduri în valoare de 25 milioane lei cercetării asociate COVID-19 alocate în regim competitiv accelerat prin instrumentul Soluții - 2020 (PN III) gestionat de către UEFISCDI.

Mecanismele și metodele dezvoltate până în prezent la nivel european și la care pot apela cercetătorii europeni sunt prezentate în cele ce urmează. **Parte a comunității europene, cercetătorii români sunt invitați, de asemenea, să contribuie la accelerarea cunoașterii și descoperirilor relevante pentru pandemia actuală.**

Acces liber la publicații, date și alte rezultate relevante

Comisia Europeană îndeamnă cercetătorii implicați în proiecte finanțate prin Orizont 2020 ale căror rezultate de cercetare pot fi utilizate (în orice fel) pentru a avansa cercetarea COVID-19 să ofere **acces liber imediat la publicațiile, datele și orice alt rezultat relevant**, în conformitate cu îndrumările oferite în ghidul „[Horizon 2020 projects working on the 2019 coronavirus disease \(COVID-19\), the severe acute respiratory syndrome coronavirus 2 \(SARS-CoV-2\), and related topics: Guidelines for open access to publications, data and other research outputs](#)”. Acestea pot fi proiecte care cercetează în special coronavirus, dar și alte domenii/discipline de cercetare cu relevanță pentru combaterea crizei medicale. În mod similar, Comisia Europeană îndeamnă și infrastructurile de cercetare (RIs) să furnizeze și să pună la dispoziție serviciile, instrumentele și resursele de cercetare relevante, pentru a oferi acces prioritar și personalizat la serviciile lor de cercetare dedicate COVID-19. În rândul RIs, [CERIC](#), în cadrul proiectului [ACCELERATE](#) realizează o hartă și centralizare a serviciilor puse la dispoziția cercetătorilor pentru cercetarea legată de COVID-19 (această hartă dinamică se regăsește pe website-ul proiectului [aici](#)).

Acțiunile CE se alătură unei tendințe globale de a deschide accesul la rezultate ce pot sprijini cercetarea și grăbi găsirea de soluții în fața provocărilor generate de pandemie. Amintim astfel de instituții europene importante (EUA, Inserm, UNIMED - Mediterranean Universities Union) și alți actori internaționali care s-au angajat prin declarația **Wellcome Trust** „[Statement on Data Sharing in Public Health Emergencies](#)” să facă disponibile cu acces liber (*open access* – OA) - înainte de *peer-review* - rezultate intermediare și finale relevante pentru COVID-19 pe care le partajează cu OMS. O listă completă a semnatarilor se regăsește aici: <https://wellcome.ac.uk/coronavirus-covid-19/open-data>.

Platforma Europeană COVID-19 Data Platform

Comisia Europeană a lansat în aprilie 2020 platforma [European COVID-19 Data Platform](#) în parteneriat cu [EBI-MLB](#) și [ELIXIR](#) pentru a facilita colectarea și schimbul rapid de date cu privire la Coronavirus. Platforma este una din cele zece priorități ale [Planului de acțiune ERAvsCORONA](#) lansat de către Comisia Europeană.

European COVID-19 Data Platform are două componente:

- **Hub-uri de date „SARS-CoV-2 Data Hubs”** care organizează fluxul de date privind SARS-CoV-2 și facilitează un schimb comprehensiv de date deschise pentru comunitățile europene și globale de cercetare; acestea sunt construite pe infrastructura **EMBL-EBI** ([The European Bioinformatics Institute](#)).
- **Portalul „COVID-19 Data Portal”** care reunește și actualizează continuu seturile de date și instrumentele COVID-19 relevante, găzduiește schimbul de date secvențial și facilitează accesul la alte resurse SARS-CoV-2.

Portalul **Covid19 Data Portal**, parte a European COVID-19 Data Platform colectează date din mai multe domenii specifice. Detalii despre acestea și procedura de a împărtăși noi date se regăsesc aici: <https://www.covid19dataportal.org/submit-data>.

Domeniile din care se colectează și se împărtășesc date sunt următoarele:

Viral, non-human and cell line sequence data

Human molecular biology data

Linked viral and human molecular biology data

Viral and non-human proteomics data

Structural biology data

Viral and non-human molecular interaction data


Viral and non-human metabolomics data

Viral and other non-human molecular biology data

Compound and target data

Clinical and epidemiological data

Non-biological data

 **COVID-19 Data Portal** nu este în sine o baza de date, ci e poarta de acces pentru o structurare și vizualizare a informației relevante existente în și colectate prin diferite baze de date. În prezent agregă informații din baze de date cu acces liber de la diferite resurse EMBL-EBI, printre care Arhiva Europeană a Nucleotidelor ([European Nucleotide Archive - ENA](#)), [UniProt](#) (acces liber la informații privind secvențierea proteinelor și a genomului), [Protein Data Bank in Europe](#) (PDBe), [Electron Microscopy Data Bank](#) (EMDB), [Expression Atlas](#) (acces liber la informații privind tipare de expresie genică în condiții biologice diferite) și [Europe PubMed Central](#) (Europe PMC). Portalul va include și date genomice rezultate din epidemie. De asemenea, o activitate principală constantă a portalului este de a reuni continuu baze de date și instrumente din alte proiecte europene relevante.

În ceea ce privește vizualizarea și interpretarea informațiilor, portalul oferă funcția de căutare dedicată datelor clinice și epidemiologice și instrumente de analiză și vizualizare de date (asigurate de către EBI). Alte colaborări pentru a asigura infrastructura necesară pentru dezvoltarea portalului sunt cu Erasmus Medical Centre, Technical University of Denmark, Eotvos Lorand University, National Institute for Public Health and the Environment și Universitaetsklinikum Heidelberg.

Mai multe informații:

Informații exacte privind modalități de implicare/colaborare pot fi solicitate la adresa: covid19@ebi.ac.uk. Cei care doresc să își împărtășească datele prin portalul **COVID-19 Data Portal** sunt sfătuiți să scrie direct la adresa de email: virus-dataflow@ebi.ac.uk.

Ghiduri și recomandări pentru schimbul liber, rapid și sigur de date de cercetare relevante pandemiei actuale - *Research Data Alliance (RDA) Covid-19 Working Group*

O altă inițiativă importantă în sprijinul schimbului global de date este cea demarată de RDA, cu sprijinul CE și a altor inițiative europene (OpenAIRE, EOSC) în vederea realizării ghidurilor pentru asigurarea unui schimb de date **rapid**, în **siguranță** și cu **acces liber** și imediat. Această acțiune derulată în cadrul unui grup de lucru dedicat ([RDA Covid-19 Working Group](#)) a pornit de la recomandările și principiile deja existente în practica „open research data” ([FAIR principles](#), [TRUST Principles](#), ghidurile OpenAIRE și EOSC, ș.a.), iar cu sprijinul benevol al experților implicați în proiecte pentru „open science”, dar și în cercetarea specifică domeniilor relevante sunt dezvoltate actualmente ghiduri de „data sharing” pentru domenii specifice cu relevanță pentru SARS-CoV-2 și COVID-19.

! Ghidurile (încă în lucru) sunt disponibile persoanelor interesate și acceptă recomandări și comentarii publice. În prezent, versiunea 5 (pre-finală) disponibilă [aici](#) este supusă consultării publice până la data de 8 iunie 2020. Mai multe detalii despre activitatea grupului și resurse, aici: <https://doi.org/10.15497/rda00046>.

Pentru cine și de ce? Rezultatele grupului de lucru (RDA COVID-19 WG) contribuie la:

- 1) definirea de ghiduri clare privind schimbul de date în circumstanțele actualei pandemii astfel ca părțile interesate să urmeze cele mai bune practici pentru a maximiza eficiența muncii lor, având rol de *blueprint* pentru viitoare situații asemănătoare sau urgențe;
- 2) elaborarea de linii directoare pentru factorii de decizie pentru a maximiza schimbul calitativ de date în timp util și a oferi răspunsuri adecvate în astfel de crize;
- 3) sprijinirea cercetătorilor, factorilor de decizie, finanțatorilor, editorilor și furnizorilor de infrastructuri de partajare a datelor.

Cercetătorii români care doresc ca datele rezultate din cercetările lor să fie parte în schimbul acesta european și global de date pot consulta ghidul RDA și pot adresa întrebări specifice privind schimbul de informații pe domenii specifice contactând subgrupurile dedicate:

1. [Clinical](#)
2. [Community Participation](#)
3. [Epidemiology](#)
4. [Indigenous Data](#)
5. [Legal and Ethical](#)
6. [Omics](#)
7. [Social Sciences](#)
8. [Software](#)

De asemenea, comunicarea cu reprezentanți ai grupului *RDA Covid-19 Working Group* poate fi facilitată de [Open Science Hub Romania](#) care este, totodată, reprezentant național al RDA ([RDA Node Romania](#)).

Depozitarea și schimbul datelor – pe scurt

Pregătirea datelor

FAIR Data – date identificabile, accesibile, interoperabile, reutilizabile (**F**indable, **A**ccesible, **I**nteroperable, **R**eusable).

Recomandările europene privind managementul datelor de cercetare, arhivarea pe termen lung și deschiderea datelor au în centru principiile FAIR. Gestionarea datelor pe modelul și după criteriile FAIR nu presupune automat și deschiderea tuturor datelor, care nu întotdeauna pot fi făcute publice/deschise tuturor, ci este importantă în primul rând pentru o depozitare adecvată și sigură a informației electronice. Principiile descriu capacitatea sistemelor computaționale de a găsi, accesa, asigura interoperabilitatea și reutiliza datele fără sau cu o foarte mică intervenție a omului. În tabelul de mai jos (tabel 1) sunt foarte pe scurt menționate caracteristicile principale. Detalii complete despre FAIR se regăsesc în [website-ul GO-FAIR](#). De asemenea, o auto-evaluare privind gradul de “FAIRness” poate fi realizată cu ajutorul instrumentului *FAIR self-assessment tool* (deținut de [Australian Research Data Commons](#)): <https://www.ands-nectar-rds.org.au/fair-tool>.

| Principiu | Caracteristici |
|-----------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Identificabile <i>Findable</i> | Metadatele și datele trebuie să fie ușor de accesat, atât pentru oameni, cât și pentru computer; pentru aceasta este necesară utilizarea de identificatori unici persistenti (precum DOI). |
| Accesibile <i>Accessible</i> | Principiul detaliază modul în care datele și metadatele pot fi preluate automat de la identificatorii lor. Accesibilitatea este susținută prin utilizarea de protocoale de comunicare standardizate. Metadatele sunt accesibile chiar și atunci când datele nu mai sunt disponibile. Accesibil în spiritul FAIR nu înseamnă neapărat deschis sau fără costuri. Mai multe informații: <i>FAIR Accessor</i> https://peerj.com/preprints/2522v2/#supp-1 . |
| Interoperabile <i>Interoperable</i> | Principiul detaliază integrarea datelor în așa fel încât să relaționeze cu aplicații sau alte fluxuri de date din surse diferite și să permită analize, stocare și procesare. |
| Reutilizabile <i>Reusable</i> | Principiul oferă ghidarea pentru optimizarea reutilizării datelor; metadatele și datele trebuie să fie bine descrise pentru a putea fi replicate și combinate în diferite cercetări viitoare |

Tabel 1. Principiile FAIR - pe scurt

Planurile de management de date (*Data management Plans* – DMPs)

Există instrumente suport care oferă diferite modele (*templates*) de planuri de management al datelor, precum:

- [DMPonline](#) (al *Digital Curation Centre*) ține cont de criteriile a diferiți finanțatori (conține și model cu cerințele Orizont 2020). Prin intermediul platformei DMPonline pot fi create planuri personalizate DMPs;
- [ARGOS](#) automatizează crearea, managementul, schimbul și conectarea DMPs cu rezultate/proiecte de care sunt legate (elaborat de OpenAIRE și EUDAT CDI ținând cont de principiile FAIR și practicile *open science*);
- Comisia Europeană oferă [template](#) proiectelor finanțate prin H2020 pentru elaborarea DMPs;

- [The Data Stewardship Wizard](#) (realizat de [ELIXIR Cehia](#) și [ELIXIR Olanda](#)) oferă de asemenea sprijin în elaborarea DMPs ținând cont de cerințele FAIR.

Depozitarea datelor

Datele necesare pentru o colectare, arhivare adecvată și încurajate spre a fi făcute publice sunt cele care stau la baza unei publicații științifice, dar și date de sine stătătoare.

Acestea trebuie să fie depuse într-un depozit de încredere care asigură următoarele:

- identificatori unici și persistenți (PID);
- găsirea și identificarea datelor, respectiv integritatea și autenticitatea datelor;
- arhivare pe termen lung;
- licențe;
- infrastructura organizațională¹;
- securitate;
- tehnologie (infrastructura tehnică);
- metadate standardizate;
- curățarea datelor și asigurarea calității lor;
- acces (eventual liber, facil pentru reutilizare);
- etică și confidențialitate;
- format comun;
- proveniență și eventual istoric al intervențiilor asupra seturilor de date.

În acest context, certificarea depozitelor este fundamentală pentru garantarea încrederii și sustenabilității depozitelor de date. În acest fel, părțile interesate pot fi încrezătoare că depozitele asigură integritatea, autenticitatea, exactitatea, siguranța datelor, precum și accesul la date pe termen lung. În timp ce principiul „FAIR” definește proprietățile datelor și metadatelor, principiul „TRUST - Transparency, Responsibility, User focus, Sustainability and Technology” descrie caracteristicile depozitelor de date care sunt responsabile de managementul și diseminarea datelor pe o perioadă lungă de timp.

| Principiu | Îndrumări pentru depozite |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Transparency</i> | Să fie transparente cu privire la serviciile de depozitare specifice și datele păstrate care pot fi verificate prin mijloace accesibile în mod public. |
| <i>Responsibility</i> | Să fie responsabile pentru asigurarea autenticității și integrității datelor și încrederea și persistența serviciilor. |
| <i>User Focus</i> | Să se asigure că normele privind managementul datelor și așteptările comunităților de utilizatori cărora li se adresează sunt îndeplinite. |
| <i>Sustainability</i> | Să asigure sustenabilitatea serviciilor și să păstreze stocurile de date pe termen lung. |
| <i>Technology</i> | Să furnizeze infrastructura și alte resurse pentru a sprijini servicii sigure, persistente și de încredere. |

Tabel 2: Principiile „TRUST” – pe scurt. Sursa: Lin et al., 2020. *The TRUST Principles for Digital Repositories. Scientific Data* (<https://doi.org/10.1038/s41597-020-0486-7>)

¹ Conform [CoreTRUSTSeal](#), ar trebui ca depozitul să aibă o finanțare adecvată și un număr suficient de personal calificat cu expertiză în arhivarea datelor printr-un sistem clar de guvernare pentru îndeplinirea eficientă a misiunii.

Depozitele sigure și de încredere pentru date de cercetare pot fi găsite prin intermediul registrelor [R3data – Registry of Research Data Repositories](#), [Core Trust Seal certified repositories](#) sau [ROAR \(Repository of Open Access Repositories\)](#). De asemenea, [Zenodo](#) sprijină depozitarea datelor în diverse formate.

În ceea ce privește strict cercetarea COVID-19, în completarea celor enumerate mai sus ca parte a portalului european COVID-19, informații și o colecție draft de baze de date se găsesc centralizat și aici: <https://fairsharing.org/collection/COVID19Resources>.

Standarde pentru date

Pentru SARS-CoV-2 și COVID-19

EMBL-EBI a pregătit un document preliminar dedicat standardelor de date SARS-CoV-2 pentru a fi publicate în Arhiva Europeană a Nucleotidelor (ENA), disponibil în Anexa 3 a ghidului Comisiei Europene.

De asemenea, Grupul de lucru al RDA pentru Covid-19 (**RDA Covid-19 WG**) oferă indicații compacte în ghidul [RDA COVID-19; recommendations and guidelines, 5th release, 28 May 2020](#) (pag. 17, 21).

Standarde generale pentru date din domenii tematice

[FAIRsharing](#) îndeplinește rolul unui portal pentru standarde de date, baze de date și politici privind datele în domenii diverse (inclusiv al științelor vieții).

[Digital Curation Centre](#) și Research Data Alliance prin [Metadata Standards Directory](#) oferă informații privind standarde de metadata disciplinare.

Alte resurse cu acces liber

Portalul OpenAIRE

[OpenAIRE COVID-19 Gateway](#) agregă toate tipurile de informații referitoare la COVID-19 (publicații, date, software și alte rezultate de cercetare) pe care le conectează între ele pentru a înlesni accesarea lor, navigarea și descoperirea într-un portal unic.

Cataloage de publicații și date cu acces liber puse la dispoziția comunității pentru un schimb rapid de informații privind COVID-19 ², pot fi accesate direct la link-urile de mai jos:

[Baza de date a OMS](#)

[The EU Clinical Trials Register and the European Union Drug Regulating Authorities Clinical Trials database](#) - EudraCT database

[The ClinicalTrials.gov database](#)

[European Network of Centres for Pharmacoepidemiology and Pharmacovigilance](#), ENCePP

Reviste care oferă acces liber la resurse relevante pandemiei:

[BMJ](#)

[Cambridge University Press](#)

[Centers for Disease Control and Prevention](#)

[Chinese Medical Association](#)

[Cochrane](#)

[Elsevier](#)

[European Centre for Disease Prevention and Control \(ECDC\)](#)

[JAMA Network](#)

[The Lancet](#)

[LITCOVID: US National Library of Medicine](#)

[New England Journal of Medicine](#)

[Oxford University Press](#)

[PLOS](#)

[Public Health England](#)

[Science](#)

[Springer Nature](#)

[SSRN \(Preprints\)](#)

[Wiley](#)

² Sursa: World Health Organisation website: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov>

Referințe:

European Commission (2020), *Horizon 2020 projects working on the 2019 coronavirus disease (COVID-19), the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and related topics: Guidelines for open access to publications, data and other research outputs*, versiune 1.1, 24 aprilie 2020, disponibil la: https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-guidelines-oa-covid-19_en.pdf (12 mai 2020);

Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. *Sci Data* **7**, 144 (2020). <https://doi.org/10.1038/s41597-020-0486-7>

Proffitt, Allison, *Data Sharing During A Pandemic: EMBL-EBI's COVID-19 Portal, GA4GH Take On Data Sharing*, Bio-IT World, 23 aprilie 2020, disponibil la: <http://www.bio-itworld.com/2020/04/23/data-sharing-during-pandemic-embl-ebi-covid-19-ga4gh.aspx> (28 mai 2020);

UEFISCDI (website). *Open Science Hub Romania*, disponibil la: <https://uefiscdi.gov.ro/open-science-hub> (28 mai 2020);

RDA COVID-19; recommendations and guidelines, 5th release 28 May 2020 [DOI: 10.15497/RDA00046](https://doi.org/10.15497/RDA00046);

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>;

Science Europe (2018). *Practical Guide to the International Alignment of Research Data Management: D/2018/13.324/4*, Science Europe, noiembrie 2018, disponibil la: https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf;

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; et al. (15 March 2016). *The FAIR Guiding Principles for scientific data management and stewardship*. Scientific Data. **3**: 160018. [doi:10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18);

World Health Organization (website). *Global research on coronavirus disease (COVID-19)*, disponibil la: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov> (2 mai 2020).