

DATELE

Findable Accesible Interoperable Reusable



Eveniment Național de Diseminare

București, 27.10.2020

dr. Ella Magdalena CIUPERCĂ, ICI București

- ❑ În actualul ecosistem digital, oamenii depind din ce în ce mai mult de suportul tehnologic pentru a fi capabili să gestioneze creșterea semnificativă a volumului, complexității și vitezei de generare a datelor
- ❑ Asigurarea accesului la datele de cercetare pentru comunități științifice tot mai largi este o condiție esențială pentru creșterea transparenței cercetării și pentru accelerarea procesului de identificare a cunoștințelor
- ❑ *Inițiativa Cloud European - Construirea unei economii competitive de date și cunoștințe în Europa (2016):*
 - ❑ dezvoltarea unui mediu de încredere, deschis, care să permită comunității științifice să stocheze, să facă schimb și să reutilizeze datele și rezultatele științifice, și anume cloudul european destinat științei deschise.

□ Un set de principii în acord cu care datele de cercetare sunt

- Findable (Identificabile)
- Accessible (Accesibile)
- Interoperable (Interoperabile)
- Reusable (Reutilizabile)

□ Obiectivele FAIR

- asigurarea accesului la datele de cercetare pentru comunități științifice tot mai largi
- oferirea de sprijin pentru dezvoltarea capacității calculatoarelor de a identifica, accesa, interopera și reutiliza date cu intervenție umană minimă

- ❑ Pentru date **identificabile**: asigurați-vă că au un identificator persistent și metadata bogate, că sunt ușor de căutat și de descoperit online.
- ❑ Pentru date **accesibile**: asigurați-vă că pot fi preluate online utilizând protocoale standardizate și impunând restricții specifice dacă este necesar.
- ❑ Pentru date **interoperabile**: utilizați formatele și standardele comune, raportați-vă la vocabulare predefinite.
- ❑ Pentru date **reutilizabile**: asigurându-vă că sunt bine documentate, că prezintă informații clare despre licență, despre proveniența informațiilor și despre modul de generare a datelor.

Exemple de site-uri care conțin mai multe informații despre FAIR

<https://www.go-fair.org/resources/>

- **GO FAIR Materials**
- **GO FAIR Workshop Series**
- **FAQ**
- **Starter Kit for Research Data Management**
- **More on FAIR**
- **Glossary**

Initiative referitoare la FAIR

FAIRsFAIR (Fostering FAIR data practices) - www.fairsfair.eu

(își propune să ofere soluții practice pentru aplicarea principiilor FAIR pe tot parcursul ciclului de viață al datelor de cercetare)

GO FAIR initiative - www.go-fair.org

(vizează implementarea principiilor pentru date FAIR)

FAIRsharing - fairsharing.org

(o resursă informativă, de educație privind standardele de date și metadate, legate de bazele de date și politicile de date)

- ❑ Primul pas pentru a reutiliza datele este regăsirea acestora
- ❑ Datele și metadatele trebuie să fie ușor de identificat, de oameni și calculator
- ❑ Metadatele interpretate de calculator sunt esențiale pentru descoperirea automată a seturilor de date

- ❑ **Cerințe pentru date *Findable*:**
 - ❑ F1. (meta)datele au asignat un identificator unic la nivel global și persistent (ex. [DOI](#), [ORCID](#))
 - ❑ F2. datele sunt descrise prin intermediul metadatelor (ex. [DICOM](#))
 - ❑ F3. (meta)datele sunt înregistrate sau indexate pe suport care permite căutarea (ex. [Zenodo](#))
 - ❑ F4. Metadatele specifică clar și explicit identificatorul datelor pe care le descriu (ex. [Dublin Core](#); de consultat: [FAIRsharing](#), [RDA Metadata Directory](#), [DCC Metadata Directory](#))

- ❑ Odată ce utilizatorii găsesc datele de interes, au nevoie să știe cum pot fi acestea accesate, eventual prin autentificare și autorizare

- ❑ Cerințe pentru date *Accesibile*:
 - ❑ A1. (meta)datele pot fi regăsite pe baza identificatorului lor folosind un protocol de comunicații standardizat:
 - ❑ A1.1. protocolul este deschis, gratuit și universal implementabil (ex. HTTP, FTP)
 - ❑ A1.2. este definită, după caz, o procedură de acces
 - ❑ ex. repozitoare de date de cercetare DRYAD, Harvard Dataverse
 - ❑ registrul repozitoarelor de date de cercetare re3data.org
 - ❑ A2. (meta)datele sunt accesibile chiar și după ce datele nu mai sunt disponibile

- ❑ Din punct de vedere la utilizatorului, datele trebuie să poată fi:
 - ❑ integrate cu alte date
 - ❑ interoperabile cu aplicații sau fluxuri de lucru pentru a fi analizate, stocate și procesate

- ❑ Cerințe pentru date *Interoperabile*:
 - ❑ I1. (meta)datele utilizează un limbaj formal, accesibil, partajat și aplicabil pe scară largă pentru reprezentarea cunoștințelor
 - ❑ I2. (meta)datele folosesc vocabulare, tezaure, ontologii care respectă principiile FAIR
 - ❑ I3. (meta)datele includ referințe calificate la alte (meta)date

- ❑ Pentru optimizarea reutilizării datelor, metadatele și datele trebuie să fie bine descrise, astfel încât să poată fi replicate și/sau combinate în diferite configurații

- ❑ Cerințe pentru date *Reusable*:
 - ❑ R1. (meta)datele au o multitudine de attribute precise și relevante (ex. fișiere .README):
 - ❑ R1.1. (meta)datele sunt eliberate cu o licență clară și accesibilă de utilizare (ex. Creative Common)
 - ❑ R1.2. (meta)datele sunt asociate cu proveniența lor
 - ❑ R1.3. (meta)datele respectă standardele comunității relevante din domeniu

- Pentru a verifica nivelul de respectare a principiilor FAIR, cercetătorii au la dispoziție documetații specifice:
 - “How FAIR are your data” checklist (<http://doi.org/10.5281/zenodo.1065991>)
 - “A design framework and exemplar metrics for FAIRness” (<https://www.biorxiv.org/content/10.1101/225490v3>)
 - SATIFYD - Self-Assessment Tool to Improve the FAIRness of Your Dataset (<https://satisfyd.dans.knaw.nl/>)

- ❑ Principiile FAIR se aplică datelor deschise, dar și datelor care pot fi partajate cu respectarea anumitor restricții
- ❑ Principiile FAIR se aplică nu numai datelor, ci și altor obiecte digitale specifice activității de cercetare: algoritmi, instrumente, fluxuri de lucru etc.
- ❑ FAIR și Open sunt concepte complementare în definirea politicilor europene sau naționale:
 - ❑ Directiva UE actualizată privind datele deschise și reutilizarea informațiilor din sectorul public (UE 2019/1024 din 5.04.2019) solicită adoptarea de politici și măsuri pentru asigurarea accesului liber la datele provenite din cercetare care au beneficiat de finanțare publică, inclusiv prin aplicarea principiilor FAIR.

- ❑ Facilitează maximizarea potențialului datelor de cercetare și a impactului cercetării prin creșterea vizibilității și a reutilizării
- ❑ Îmbunătățesc reproductibilitatea și fiabilitatea cercetării
- ❑ Facilitează alinierea la abordările și standardele internaționale
- ❑ Permit integrarea în noi parteneriate de cercetare sau de valorificare a rezultatelor cercetării
- ❑ Promovează instrumente și paradigme de cercetare inovative

- ❑ EOSC se va implica în asigurarea posibilității cercetătorilor europeni de a beneficia de oferta științei bazate pe date, punând la dispoziție un mediu cu servicii gratuite și deschise de stocare, management, analiză și reutilizare a datelor din discipline diferite
- ❑ *Strategia europeană a datelor (2020)*: are ca obiectiv crearea unei piețe unice pentru date, ceea ce va permite circulația liberă în interiorul UE și între domenii de activitate, în beneficiul întreprinderilor, cercetătorilor și al administrației publice
 - ❑ A European strategy for data. COM(2020) 66 final, Brussels, 19.2.2020.
https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf

- ❑ Wilkinson, M., et al. (2016). “*The FAIR Guiding Principles for scientific data management and stewardship*”, Scientific Data 3.
<https://doi.org/10.1038/sdata.2016.18>
- ❑ <https://www.openaire.eu>
- ❑ <https://www.eosc-portal.eu/>
- ❑ <https://www.fosteropenscience.eu/>
- ❑ <https://www.go-fair.org/>