

Tot ce trebuie să știi despre
datele FAIR

CE SUNT DATELE FAIR ?

Creșterea continuă a producției și disponibilității resurselor digitale impune crearea datelor cu luarea în considerare a longevității lor. Asigurarea accesului la datele de cercetare pentru comunități științifice tot mai largi a ușurat procesul de identificare a cunoștințelor și a crescut transparența cercetării. În lumina acestui fapt, în 2016 a fost lansat 'The FAIR Guiding Principles for scientific data management and stewardship' (Principiile directe FAIR pentru managementul și administrarea datelor științifice). FAIR nu este un standard, ci un set de principii în acord cu care datele de cercetare sunt Findable (Identificabile), Accessible (Accesibile), Interoperable (Interoperabile) și Reusable (Reutilizabile), așa cum se prezintă pe pagina următoare. În afară de a facilita reutilizarea datelor de către cercetători individuali, principiile FAIR pun un accent important pe dezvoltarea capacității calculatoarelor de a permite identificarea și utilizarea datelor în mod automat. Aceste principii subliniază faptul că 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. În acest context, cheia o reprezintă capacitatea de acțiune a calculatorului (capacitatea sistemelor de calcul de a identifica, accesa, interopera și reutiliza date cu intervenție umană minimă sau absentă) propusă de principiile FAIR.



O concepție greșită, întâlnită frecventă în rândul cercetătorilor, o constituie echivalarea datelor FAIR cu datele deschise. Nu este corect pentru ca datele FAIR nu trebuie să fie obligatoriu deschise, putând fi partajate și în condițiile acțiunii anumitor restricții.

CUM SĂ VĂ FACEȚI DATELE FAIR

Transpunerea principiilor FAIR în practică poate diferi în funcție de disciplinele la care se aplică, dar următoarele linii directoare sunt general valabile:

- Faceți datele identificabile asigurându-vă că au un identificator persistent și metadata bogate, că sunt ușor de căutat și de descoperit online.
- Faceți datele accesibile asigurându-vă că pot fi preluate online utilizând protocoale standardizate și impunând restricții specifice dacă este necesar.
- Faceți datele interoperabile prin utilizarea formatelor și standardelor comune și prin raportarea la vocabulare predefinite.
- Faceți datele reutilizabile asigurându-vă că sunt bine documentate, că prezintă informații clare despre licență, despre proveniența informațiilor și despre modul de generare a datelor.
- Folosiți abordări și instrumente de cercetare inovative.

Identificabile

Primul pas pentru a reutiliza datele este găsirea acestora. Datele și metadatale trebuie să fie ușor de identificat, atât de oameni cât și de calculator. Metadatale interpretate de calculator sunt esențiale pentru descoperirea automată a seturilor de date și a serviciilor, fiind deci o componentă esențială a procesului de FAIR-ificare.

Accesibile

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

Interoperabile

De obicei, datele trebuie să fie integrate cu alte date. În plus, trebuie să fie interoperabile cu aplicații sau fluxuri de lucru pentru a fi analizate, stocate și procesate.

Reutilizabile

Scopul ultim al FAIR este optimizarea reutilizării datelor. Pentru realizarea acestui obiectiv, metadatale și datele trebuie să fie bine descrise, astfel încât să poată fi replicate și/sau combinate în diferite configurații.



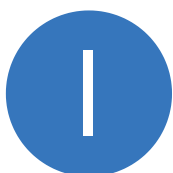
Pași de urmat către date Identificabile

- Metadatelor le sunt asignat un identificator persistent unic (PID)
- Datele sunt descrise cu metadate bogate
- Metadatele includ în mod clar și explicit identificatorul dateilor pe care îl descriu
- Metadatele sunt înregistrate și indexate într-o resursă care permite efectuarea de căutări.



Pași de urmat către Accesibilitate

- Metadatele pot fi recuperate cu ajutorul identificatorului lor, utilizând un protocol de comunicații standardizat
- Protocolul este deschis, gratuit și implementabil universal
- Protocolul permite o procedură de autentificare și autorizare, când este necesar
- Metadatele sunt accesibile, chiar dacă datele nu mai sunt disponibile.



Pași de urmat către Interoperabilitate

- Metadatele folosesc un limbaj pentru reprezentarea cunoștințelor formal, accesibil, partajat și aplicabil pe scară largă
- Metadatele utilizează vocabulare care respectă principiile FAIR
- Metadatele includ referințe calificate către alte metadate.



Pași de urmat către Reutilizabilitate

- Metadatele sunt descrise cu o multitudine de atribute corecte și relevante
- Metadatele sunt expuse cu o licență clară și accesibilă de utilizare a datelor
- Metadatele sunt asociate cu elemente care detaliază proveniența
- Metadatele respectă standardele relevante în comunitate.

DE CE SĂ VĂ FACEȚI DATELE FAIR?

Asigurarea caracterului FAIR al datelor de cercetare are multiple beneficii pentru cercetători, comunități de cercetare, facilități din infrastructura de cercetare și organizații de cercetare. Datele FAIR:

- facilitează maximizarea potențialului datelor și a impactului cercetării prin creșterea vizibilității și a citărilor
- îmbunătățesc reproductibilitatea și fiabilitatea cercetării
- facilitează alinierea la abordările și standardele internaționale
- permite integrarea în noi parteneriate cu cercetători, oameni de afaceri și comunități mai largi
- permite furnizarea de răspunsuri la noi întrebări de cercetare
- utilizează instrumente și paradigme de cercetare inovative.

EOSC ȘI FAIR

Cloud-ul European pentru Știință Deschisă (The European Open Science Cloud - 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.

EOSC va furniza o platformă pentru cercetarea europeană, inclusiv o rețea de date și servicii de cercetare.

CÂT DE FAIR SUNT DATELE DUMNEAVOASTRĂ?

În efortul de a determina cât de FAIR sunt datele lor de cercetare și de a identifica modalități prin care să crească acest nivel, cercetătorii pot consulta instrumente deja existente, precum:

- “How FAIR are your data” checklist (<http://doi.org/10.5281/zenodo.1065991>), sau
- “A design framework and exemplar metrics for FAIRness” (<https://www.biorxiv.org/content/10.1101/225490v3>)

BIBLIOGRAFIE

- 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/>



www.ni4os.eu



NI4OS-Europe a primit finanțare prin programul de cercetare și inovare Horizon 2020 al Uniunii Europene, contractul de grant nr. 857645