

INSTITUTUL DE DEZVOLTARE A SOCIETĂȚII INFORMAȚIONALE

**Proiectul „Stimularea angajamentului Republicii Moldova în Știința
Deschisă: suport metodologic și aplicativ ”**

Atelier de lucru ”Principii FAIR în managementul datelor de cercetare”

Date științifice deschise

Principii FAIR

Mihai Greco
Institutul de Dezvoltare a Societății Informaționale

Chișinău, 19 mai 2022



Paradigme ale cercetării științifice

“One of the greatest challenges for 21st-century science is how we respond to this new era of data-intensive science. This is recognized as a new paradigm beyond experimental and theoretical research and computer simulations of natural phenomena—one that requires new tools, techniques, and ways of working.”

Douglas Kell, University of Manchester

Paradigma I

Știința experimentală

Paradigma II

Știința teoretică

Paradigma III

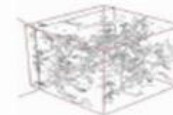
Știința computațională (recentă)

Paradigma IV

Știința datelor - analiza seturilor masive de date



$$\left(\frac{a}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$



Sursă: https://www.microsoft.com/en-us/research/wp-content/uploads/2010/07/Tony-Hey_Oct18.pdf

Schimbare de paradigmă

Cercetarea bazată pe date

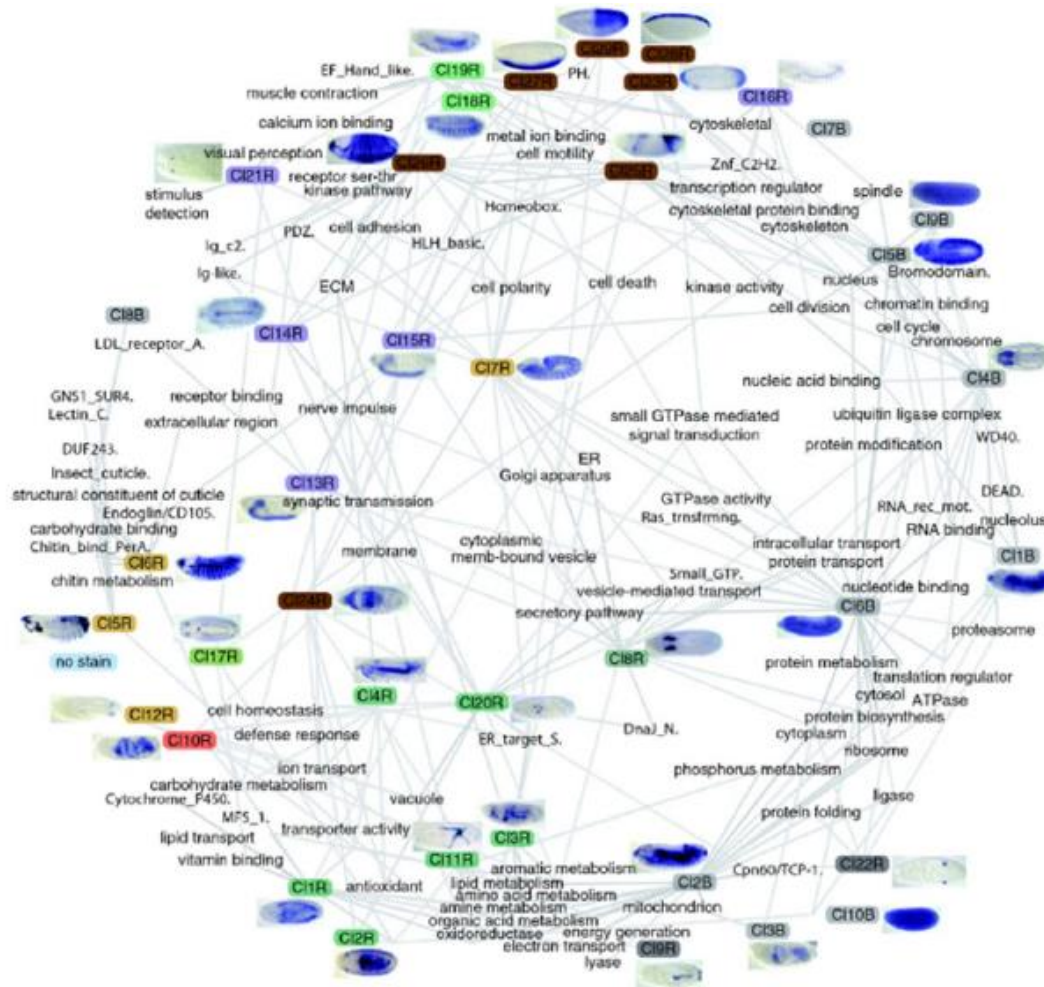
Oamenii de știință și cercetarea sunt supuși unei schimbări de paradigmă. Evoluțiile software-urilor statistice, instrumentelor și științelor bazate pe date, cum ar fi biologia computațională și chimia computațională, au condus la o nouă generație de oameni de știință care se concentrează pe analiza și interpretarea datelor obținute

Proiecte de cercetare bazate pe date:

- Large Hadron Collider
- Hubble Telescope
- Human Genome



Complexitate în creștere în biologia celulară



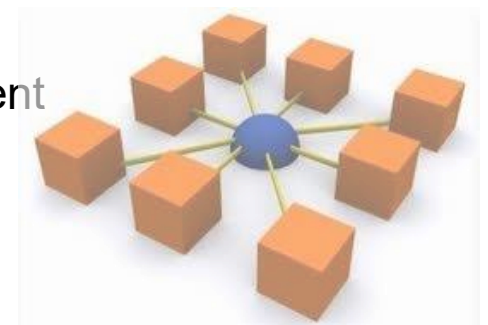
Tomancak et al. Genome Biology :R145.

<http://docplayer.net/8988022-Dixa-a-data-infrastructure-for-chemical-safety-jos-kleinjans-dept-of-toxicogenomics-maastricht-university.html>

Situația actuală privind datele de cercetare

Situația actuală privind datele de cercetare este determinată de eforturi care vizează:

- Dezvoltarea de infrastructuri specifice cercetării în cadrul unor discipline
- Foi de parcurs și politici la nivel înalt
- Interoperabilitatea infrastructurilor de date
- Raționalizarea proceselor de dezvoltare a infrastructurilor de date
- Noi modele de date/metadate
- Instrumente avansate și noi sisteme de management al datelor



Datele de cercetare în contextul accesului deschis la informațiile științifice



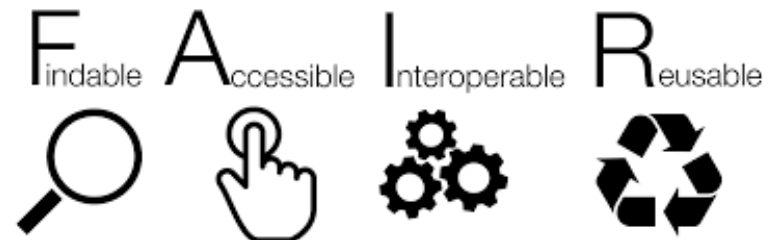
Sursa: http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm

Principii FAIR privind datele științifice

Provocarea științei intensive a datelor este de a facilita descoperirea cunoștințelor asistând oamenii și mașinile (calculatoarele) în descoperirea, accesul, integrarea și analiza datelor științifice

FAIR – este un set de principii directoare pentru a face datele

- Regăsibile (*Findable*)
- Accesibile (*Accessible*)
- Interoperabile (*Interoperable*)
- Re-utilizabile (*Reusable*)



Principii **FAIR**

F
indable



Datele științifice trebuie să fie Regăsibile (*Findable*):

- F1 (meta)datelor li se atribuie identificatori unici și persistenți
- F2 datele sunt descrise cu metadate suficient de bogate
- F3 (meta)datele sunt înregistrate/ indexate într-o resursă care poate fi căutată
- F4 metadata specifică identificatorul unei date

Principii FAIR



Datele științifice trebuie să fie Accesibile (*Accessible*):

- A1 (Meta) datele sunt recuperabile prin identificatori folosind protocoale standardizate de comunicații
 - A1.1 Protocolul este deschis, gratuit și universal implementabil
 - A1.2 Protocolul permite o procedură de autentificare și autorizare, acolo unde este necesar
- A2 Metadatele sunt accesibile, chiar și atunci când datele nu mai sunt disponibile

Principii FA/R



Datele trebuie să fie Interoperabile (*Interoperable*):

- I1 (Meta)datele utilizează un limbaj formal, accesibil, comun și aplicabil pentru reprezentarea cunoștințelor
- I2 (Meta)datele folosesc vocabulare care respectă principiile FAIR
- I3 (Meta)datele includ referințe calificate la alte (meta)date

Principii FAIR



Datele științifice trebuie să fie Reutilizabile (*Reusable*):

- R1 (Meta)datele au o pluralitate de atribute precise și relevante
 - R1.1 (Meta)datele sunt eliberate cu o licență clară și accesibilă de utilizare
 - R1.2 (Meta)datele sunt asociate cu proveniența lor
 - R1.3 (Meta)datele respectă standardele comunitare relevante pentru domeniu

Planuri de management al datelor de cercetare



Planul de management al datelor (PMD) - un document care descrie datele proiectului de cercetare, modul în care vor fi gestionate, descrise, analizate și stocate, mecanismele utilizate la păstrarea și partajarea datelor.

Scopul PMD - de a lua în considerare diferite aspecte ale gestionării datelor, generării de metadate, conservării datelor și analizei lor.

PMD asigură că datele sunt bine gestionate în prezent și sunt pregătite pentru conservare în viitor.

Contribuie la buna formalizare a proceselor de cercetare, identificarea punctelor slabe, asigurarea evidenței în cadrul cercetărilor.

Structura unui Plan de management al datelor (PMD)



Descrierea
proiectului



Datele existente



Date care vor apărea
pe parcurs



Organizarea datelor



Cadrul normativ și
codul de etică



Arhivarea, partajarea
și publicarea datelor



Responsabilități și
obligațiuni



Costuri și resurse

Proiectul e-IDSM

Titlul proiectului	Elaborarea cadrului conceptual și metodologic pentru e-infrastructura de date în sfera CDI din Republica Moldova (e-IDSM)
Direcția strategică	Patrimoniu national si dezvoltare a societatii
Termen de realizare	2018 - 2019
Programul de Stat 18.07.10A/PS50	Dezvoltarea e-Infrastructurii de date din sfera cercetării, dezvoltării și inovării din Republica Moldova
Obiectiv general	Instituirea și dezvoltarea unui sistem cuprinzător de management al datelor de cercetare

e-IDSM: Viziune. Scop

Viziune

O infrastructură de date de cercetare într-un cadru conceptual nou care să confere o nouă dimensiune socială și organizatorică actului cercetării științifice prin crearea unui ecosistem științific digital deschis, extensibil și evolutiv care să faciliteze valorificarea datelor și cunoștințelor acumulate, să fie un intensificator al cercetărilor și descoperirilor științifice.

Scop

e-IDSM este dezvoltată pentru a putea oferi servicii fiabile și durabile care să răspundă nevoilor generice și specifice privind managementul datelor de cercetare.

e-IDSM: Obiective specifice

- Infrastructură de date pentru un management sustenabil și cercetare interoperabilă
- Procese și proceduri comune, abordare standard a datelor științifice
- Servicii fiabile și sustenabile pentru managementul datelor de cercetare
- Standarde de metadate pentru utilizarea a datelor de cercetare într-un context larg
- Corelarea cu infrastructuri și platforme internaționale de date de cercetare
- Reutilizabilitatea datelor și infrastructurilor de date
- Bază tehnologică și metodologică comună privind securitatea și calitatea datelor

Componente de structură organizațională e-IDSM



Consiliul e-IDSM



Comisii de experți pe domenii



Instituție directoare, coordonare managementul datelor



Infrastructuri instituționale de date CDI



Centre de date științifice



Stewarzi de date din instituții și proiecte de cercetare



Consortii de date pe discipline și domenii

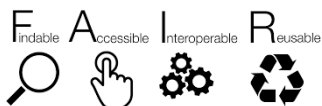


Planuri de management al datelor de cercetare

Documente normative în cadrul e-Infrastructurii datelor de cercetare (e-IDSM)



Recomandări la documentele de politici ale instituțiilor de cercetare privind datele de cercetare



Recomandări privind transpunerea principiilor FAIR în cadrul Infrastructurii datelor de cercetare



Recomandări pentru elaborarea Planului de management al datelor în proiectele de cercetare



Recomandări privind organizarea seturilor de date și a serviciilor de date în cadrul infrastructurii de date CDI



Act normativ privind rolul și responsabilitățile părților participante la dezvoltarea infrastructurii de date CDI

Reglementări și recomandări europene de referință

- Guidelines on FAIR Data Management in Horizon 2020
- Recomandarea Comisiei privind accesul la informațiile științifice și conservarea acestora
- Implementation Roadmap for the European Open Science Cloud
- Commission recommendation on access to and preservation of scientific information
- EOSC Declaration Action List
- Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020
- European Commission Open Science Policy Platform Recommendations
- The Vienna Declaration on the European Open Science Cloud
- Comunicarea Comisiei Europene "Către un spațiu european comun al datelor"
- Directiva CE privind reutilizarea informațiilor din sectorul public
- Regulamentul Comisiei cu privire la metadate

PARTHENOS



PARTHENOS

Pooling Activities, Resources and Tools
for Heritage E-research Networking,
Optimization and Synergies

PARTHENOS - infrastructură de cercetare pentru consolidarea coeziunii cercetării în domeniile conexe asociate umanității:

- studii lingvistice
- patrimoniu cultural
- istorie
- arheologie

Infrastructuri de cercetare existente precum ARIADNE (arheologie), CLARIN (lingvistică) și DARIAH (arte și științe umane) sunt membri ai PARTHENOS.

PARTHENOS dezvoltă standarde comune, coordonează activități comune, armonizează definirea și implementarea politicilor, partajează servicii și soluții comune pentru probleme de cercetare

ARIADNE - Infrastructură de cercetare în domeniul arheologiei



ARIADNE reunește și integrează infrastructuri de date de cercetare arheologice existente, astfel încât cercetătorii să poată utiliza diverse seturi de date distribuite și tehnologii noi și performante ca o componentă integrantă a metodologiei de cercetare arheologică.

Acces transnațional la centre de date, instrumente și ghiduri; crearea de noi servicii Web bazate pe depozite de date, interfețe comune, disponibilitatea seturilor de date de referință și utilizarea tehnologiilor inovatoare.

Datele se regăsesc în diverse colecții, seturi de date, rapoarte, publicații. Ele contribuie la crearea unei noi comunități de cercetători care să exploateze contribuția tehnologiei informației și să o încorporeze în metodologia de cercetare arheologică.

CLARIN: Infrastructură de date cu caracter lingvistic



Resurse lingvistice digitale pentru cercetători din toate disciplinele, în special din științele umaniste și sociale.

Soluții pe termen lung și servicii tehnologice pentru implementarea, conectarea, analizarea și susținerea datelor și instrumentelor digitale.

CLARIN susține oamenii de știință în cercetări bazate pe date, contribuind la crearea unui spațiu european de cercetare cu adevărat multilingv.

Misiunea CLARIN este realizată printr-o federație în rețea a centrelor operaționale în multe țări europene: depozite de date lingvistice, centre de servicii și centre de cunoștințe cu acces unic pentru toți membrii comunității academice din toate țările participante. Instrumentele și datele din diferitele centre CLARIN sunt interoperabile, astfel încât colecțiile de date pot fi combinate, iar instrumentele din diferite surse pot fi înlănțuite pentru a efectua operații complexe și pentru a sprijini cercetătorii în activitatea lor.

Experiențe finlandeze

Depozite/ servicii date	Conținut
AILA Data Service	Arhivă de date pentru științe sociale
IDA	Depozit și servicii de date de cercetare - Ministerul educației și culturii
FIN-CLARIN	Depozit de resurse și date lingvistice
The Language Bank of Finland	Depozit date: corpusuri de text și de vorbire, servicii de căutări versatile
LAJI.FI	Depozit de date. Informații privind biodiversitatea
LODI4DH	Servicii naționale centralizate de date corelate (Linked Data) pentru știința deschisă
<u>re3data.org</u>	Registrul depozitelor de date de cercetare
<u>AVAA</u>	Platformă de publicare a datelor deschise - Ministerul educației și culturii
<u>FinnONTO</u>	Infrastructură de ontologie la nivel național în Finlanda
<u>YSO</u>	Ontologie finlandeză generală

Mulțumesc!

Mihai Grecu
Institutul de Dezvoltare a Societății Informaționale

R. Moldova, 2028, or. Chișinău, str. Academiei, 5A

Telefon: (+373-22) 28 98 39,

Fax: (+373-22) 28 98 49

E-mail: iinfo@idsi.md, mihai.grecu@idsi.md