

Postavenie a úloha knižníc v manažmente otvorených dát

OTVORENÁ VEDA V PRAXI

Simona Hudecová,

Centrum vedecko-technických informácií SR



Táto prezentácia je šírená pod licenciou
[Creative Commons 4.0 Attribution](https://creativecommons.org/licenses/by/4.0/)



EURÓPSKA ÚNIA
Európske štrukturálne a investičné fondy
OP Integrovaná infraštruktúra 2014 – 2020



MINISTERSTVO
DOPRAVY A VÝSTAVBY
SLOVENSKEJ REPUBLIKY



MINISTERSTVO
ŠKOLSTVA, VEDY,
VÝSKUMU A ŠPORTU
SLOVENSKEJ REPUBLIKY



NISP **EZ** ^{IV}

Investícia do Vašej budúcnosti

Obsah

- Spravovanie dát – význam a úžitok, výhody
- Research data management a data governance
- Správcovstvo dát a knižnice
- Dáta na medzinárodnej úrovni, dátová infraštruktúra



[Zdroj: https://hri.fi/en_gb/support/what-is-open-data/]

Obsah

- Spravovanie dát – význam a úžitok, výhody



[Zdroj: https://hri.fi/en_gb/support/what-is-open-data/]

Otvorené dáta a nové výzvy pre knižnice

the **Declaration of the Independence of Cyberspace** – (Davos, CH, 1996), potreba novej profesie „správca dát“.

„Libraries as stewards, educators, publishers and partners: from traditional functions to new roles“

Open Science podporovanie prostredníctvom

OpenGlam sa snaží zabezpečiť otvorený prístup ku kultúrnemu dedičstvu

Open (research) Data - sprístupnenie dát pre vedcov – vzdelávanie ako a kde sprístupniť dáta:

podporovanie výskumníkov a tvorby dátovej infraštruktúry

- Poradenstvo pri uchovávaní výsledkov výskumu a súvisiacich datasetov
- Poradenstvo v oblasti správy dát vrátane dostupnosti, šírenia
- Podpora výskumníkov v oblasti tvorby, uchovávania a zdieľania dát
- Znalosti a poradenstvo v oblasti tzv. data mining, zdroje pre rôzne vedné disciplíny, trendy v hľadaní a získavaní dát
- Spolupráca pri tvorbe projektov
- Schopnosť vytvoriť schému popisných dát a poradenstvo v oblasti štandardov a postupov disciplíny/predmetu pre jednotlivé výskumné projekty.
- Stratégia 2023 – 27, LIBER, [What role for academic libraries in Open Science?](#)



Správcovstvo dát - terminológia

Designing public data commons - **a**/vytvorenie a správa verejných spoločných dát v dôveryhodnom priestore je predmetom verejného záujmu;
b/ekosystém dát;
c/kolaboratívne riadenie;
d/digitálna emancipácia.

Digital commons - objekty (dáta) aj subjektivita (dôvera, vzájomná pomoc), odklon od prísne technického chápania správy dát

Data Mapping - mapovanie dát, proces vytvárania mapovaní dátových prvkov medzi dvoma odlišnými dátovými modelmi. Je to proces porovnávania polí jednotlivých databáz, cieľom je uľahčenie migrácie dát, integrácie údajov a iných úloh správy dát.

Zásady:

- **Primerané a vyvážené pravidlá na lepšiu reguláciu digitálneho ekosystému, transparentnosť a zodpovednosť pre online platformy.**
- **Kolaboratívne riadenie ako systémová spätná väzba**

Digitálne inkluzívna spoločnosť: nová digitálna ekonomika, „kognitívny kapitalizmus“: systematický proces privatizácie informácií v rôznych formách umožňuje maximalizáciu ziskov ako napr. dátový kolonializmus (Bauwens a kol., 2019; Fumagalli a kol., 2019), dáta sa používajú na podriadenie a transformáciu sociálnych vzťahov (Culdry & Mejjias, 2019). Potrebné aktivity: ochrana údajov, boj proti ich zneužitiu, zabránenie rušivému vývoju v oblastiach dolovania údajov, školenie umelej inteligencie (AI) na základe veľkých dát (Isin & Ruppert, 2015).

Princípy otvorených dát

- **Verejná doména:** Údaje musia byť verejnými informáciami, aby mohli byť otvorené pre verejné použitie. Citlivé alebo inak kompromitujúce informácie nie je možné zverejňovať z dôvodu ochrany práv jednotlivcov na súkromie a všeobecnej bezpečnosti spoločnosti.
- **Technická dostupnosť:** Údaje boli publikované v dostupnom formáte, ktorý uľahčuje použitie pri vývoji nových webových služieb, publikovanie údajov vo formátoch CSV alebo XLS alebo poskytnutie prístupu k pôvodnému zdroju údajov prostredníctvom aplikačného programovacieho rozhrania (API).
- **Vol'ný prístup:** Dáta dostupné bez poplatkov, možnosť pre experimentovanie s údajmi bez potreby rozpočtovej byrokracie.
- **Licencia povoľujúca opätovné použitie:** vydavateľ povoľuje opätovné použitie a jasne to vyjadruje v licencií priloženej k údajom. V opačnom prípade môže byť zistenie podmienok používania také namáhavé, že údaje zostanú nevyužité.
- **Vyhľadateľnosť:** Existencia a umiestnenie zdroja údajov by mali byť verejne známe, publikovanie údajov vo verejnom katalógu údajov a pod.
- **Zrozumiteľnosť:** príslušné dáta majú zrozumiteľnú štruktúru, obsahujú popis a informácie o tom, ako boli zdokumentované, inak ich nemožno správne interpretovať a zostávajú nepoužiteľné.

Výhody otvorených dát pre knižnice

- **Transparentnosť a demokratická kontrola:** o.i. uľahčuje odkazovanie na odborné informácie, overovanie kvality
- **Inovácie:** Poskytnutie bezplatného prístupu k otvoreným dátam podporuje efektívnosť a inovácie, vedci rôzneho zamerania môžu nájsť nové spôsoby využitia informácií a rýchlejšie zvoliť efektívne riešenia.
- **Harmonizácia postupov** správy otvorených dát prináša aj finančné úspory a uľahčuje prenos znalostí medzi organizáciami. Otvorené údaje môžu tiež pomôcť pri hľadaní postupov, kde sa potenciál digitálnych informácií ešte nerealizoval.
- **Nové produkty a služby**
- **Nové poznatky z kombinovaných dátových zdrojov vo veľkých objemoch dát**
- **Efektívnosť on-line vzdelávania** čerpá zo zdrojov pre otvorené dáta

PODMIENKY :

údaje možno nájsť na internete - údaje sú prístupné (jasné práva a licencie) - údaje sú v použiteľnom formáte;

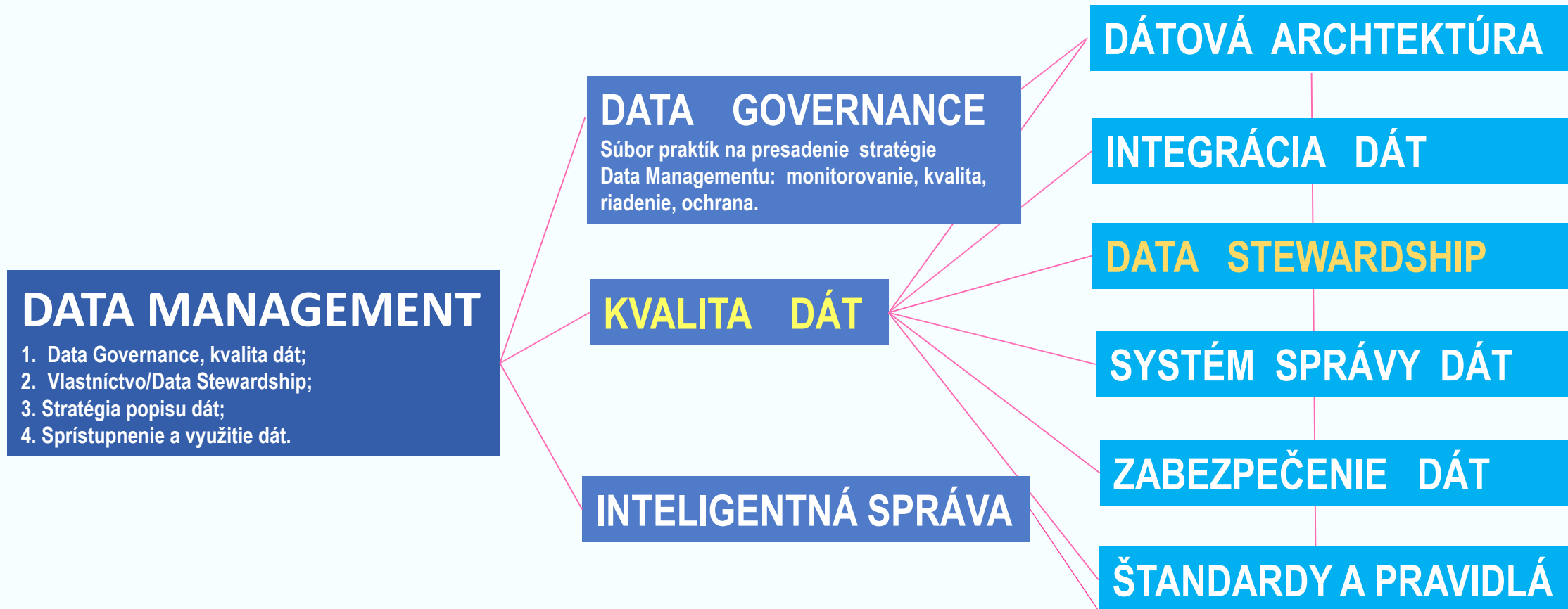
údaje sú spoľahlivé - údaje sú identifikované jedinečným a trvalým spôsobom, aby bolo možné na ne odkazovať.

- Research data management a data governance



[Zdroj: https://hri.fi/en_gb/support/what-is-open-data/]

Research data management



Data Governance

Úspešný program správy dát súbor princípov, štandardov a praktík, ktoré zaručia spoľahlivosť, dôveryhodnosť a konzistentnosť dát aj do budúcnosti. Stanovuje zásadné ciele celej stratégie, systém rozhodovacích práv a zodpovedností za procesy súvisiace s informáciami, tj. stanovuje pravidlá riadenia, právne požiadavky pre automatizované a ľudské procesy na presadzovanie týchto pravidiel.

Makroúroveň - riadenie cezhraničných tokov dát medzi krajinami, nazýva sa medzinárodné riadenie dát.

Mikroúroveň – individuálny subjekt, inštitúcia, spoločnosť a pod.

Koncepcia Data Governance - jednotný súbor pravidiel a procesov na zhromažďovanie, ukladanie a používanie dát:

- Vízia a široký strategický cieľ
- Úlohy (priority, vzťahy - sieť pracovníkov: riadiaci a výkonný výbor, sieť správcov dát, Líder správy dát/Data Governance Leader)
- Technológia správy dát – technologické riešenia správy dát, dátové inžinierstvo
- Definovať a dokumentovať štandardy a normy, indikátory kvality, klasifikácia dát zodpovednosť a vlastníctvo dát

Data Governance strategické zásady:

- zodpovednosť za kvalitu dát, informácií a vlastníctvo;
- organizačné úlohy a zodpovednosti;
- štandardy zberu a overovania dát;
- usmernenia o bezpečnosti informácií a ochrane dát;
- uchovávanie a prístup k dátam a ich používanie;
- ochrana a zabezpečenie dát;
- zásady archivácie dát.

Data Governance a priestor pre knižnice

Data Governance manažment kvality dát

- stratégia riadenia kvality dát – manažment životného cyklu dát
- tvorba terminológie,
- priebežná analýza, monitorovanie, audit a overovanie kvality dát
- hlboké disciplinárne znalosti
- presnosť (zahŕňajúca správnosť a konzistentnosť),
- použiteľnosť, integrita, úplnosť, zrozumiteľnosť

Komunikácia a zdieľanie:

- spätná väzba – odbornosť vedných disciplín, stratégia, koncepcia, aktualizácia procesov,
- monitorovanie – analýza rozhodnutí
- priebežné expertízy kvality služieb

Data Governance stanovuje: ciele, procesy, metriky a štandardy. Stanovuje úlohy a súvisiace zodpovednosti, definuje procesy riadenia a prípadných zmien pri správe dát; je spojená s riadením informačných technológií.

Stanovuje zásady na najvyššej úrovni na ochranu pred stratou, poškodením, krádežou alebo zneužitím dát.

Data Stewardship zameranie na realizáciu stanovených cieľov, noriem, pravidiel a zásad, zabezpečenie kvality a pridanej hodnoty dát. Starostlivosť o dáta, ktoré nepatria samotným správcovi, takže správcovia dát predstavujú záujmy ostatných a zabezpečujú, aby sa práca súvisiaca s dátami vykonávala v súlade so zásadami a postupmi určenými prostredníctvom správy.

DATA STEWARDSHIP organizácia

Spracovanie a konsolidácia dát: štandardy, normy



DATA GOVERNANCE LEADER

Data steward

- Integrita, archivácia dát
- Pôvod dát
- Kvalita popisných dát

Vedecký steward

- Hodnotenie, verifikácia, kontrola kvality dát
- kvalita popisných dát, využiteľnosť

IT steward

- Softvérové požiadavky
- Modely spracovania dát
- Integrita, dostupnosť, viditeľnosť dát

Projektový steward

- Pravidlá kvality dát
- Definície
- Spolupráca s IT

Obchodný steward

- Kvalita dát
- Manažment dát

Kustód dát

- Správa a uchovávanie dát, IT
- Implementácia pravidiel
- Dostupnosť dát

Vlastníci dát



Používatelia

- Správcovstvo dát a knižnice



[Zdroj: https://hri.fi/en_gb/support/what-is-open-data/]

Správcovstvo dát a knižnice

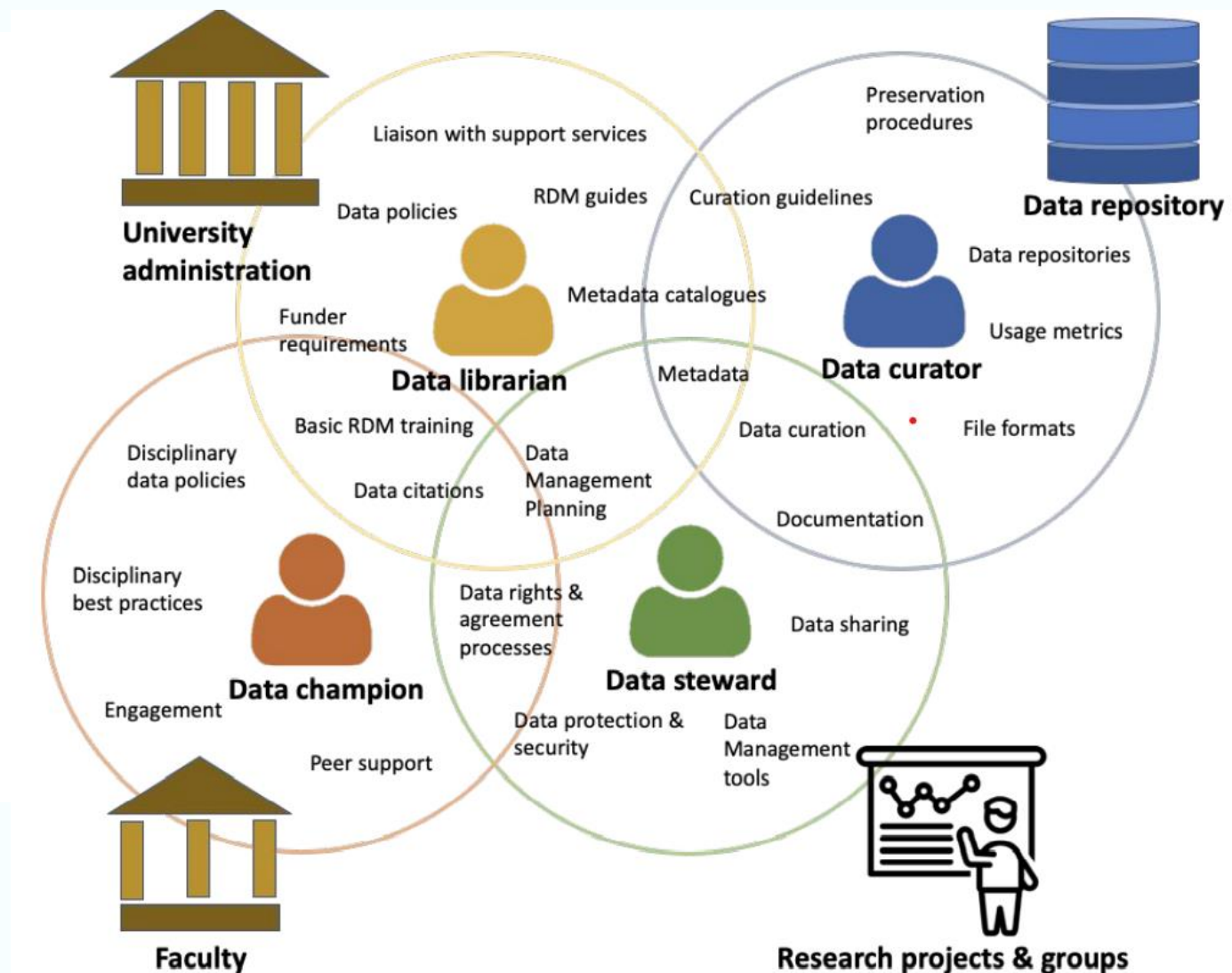
Príklad z praxe: [Helsinki University Library](https://zenodo.org):

Data librarian – informačný špecialista, sprostredkovateľ komunikácie v oblasti servisu a správy výskumných dát (IT, výskumná kancelária atď.);

Data curator – správa a uchovávanie dát a popisných dát poskytuje školenia a poradenstvo;

Data steward - manažment správy dát výskumných projektov, rozvoj pracovných postupov a postupov vo výskume, znalosti o správe dát v oblasti výskumu;

Data champion - poradca v oblasti správy dát, je dôležitý pri zapájaní sa do výskumu a marketingu



Research Data Management a knižnice

LIBER 10 Recommendation on Getting Started in RDM (2012)

- Podpora pri správe údajov z výskumu vrátane plánov správy údajov pre žiadosti o grant, poradenstvo v oblasti práv duševného vlastníctva a informačné materiály. Spolupráca s fakultou pri tvorbe plánu správy údajov a integráciou správy údajov do učebných osnov.
- Vývoj a vzdelávanie v oblasti popisných údajov, dátových štandardov a poskytovanie odborných súvisiacich služieb pre výskumné dáta.
- Špecializácia dátových knihovníkov a rozvíjanie profesionálnych zručností zamestnancov pre dátové knihovníctvo.
- Aktívny podiel na vývoji politiky inštitucionálnych výskumných údajov, vrátane plánov zdrojov.
- Partnerstvo s výskumníkmi, výskumnými skupinami, dátovými archívmi a dátovými centrami – vytvorenie infraštruktúry pre prístup k dátam.
- Podporovanie životného cyklu výskumných dát poskytovaním služieb na ukladanie, vyhľadávanie a trvalý prístup.
- Vzdelávanie v oblasti citovania výskumných dát s použitím trvalých identifikátorov na výskumné údaje.
- Tvorba inštitucionálneho dátového repozitáru.
- Zapojenie inštitucionálneho katalógu/repozitáru do infraštruktúry dátových zdrojov.
- Spolupráca pri tvorbe a správe dát pre špecifické vedné disciplíny.
- Spolupráca pri bezpečnom ukladaní dynamických a statických výskumných dát v spolupráci s inštitucionálnymi IT jednotkami a/alebo hľadání vhodných cloudových služieb.

Koordináčné činnosti knižnice

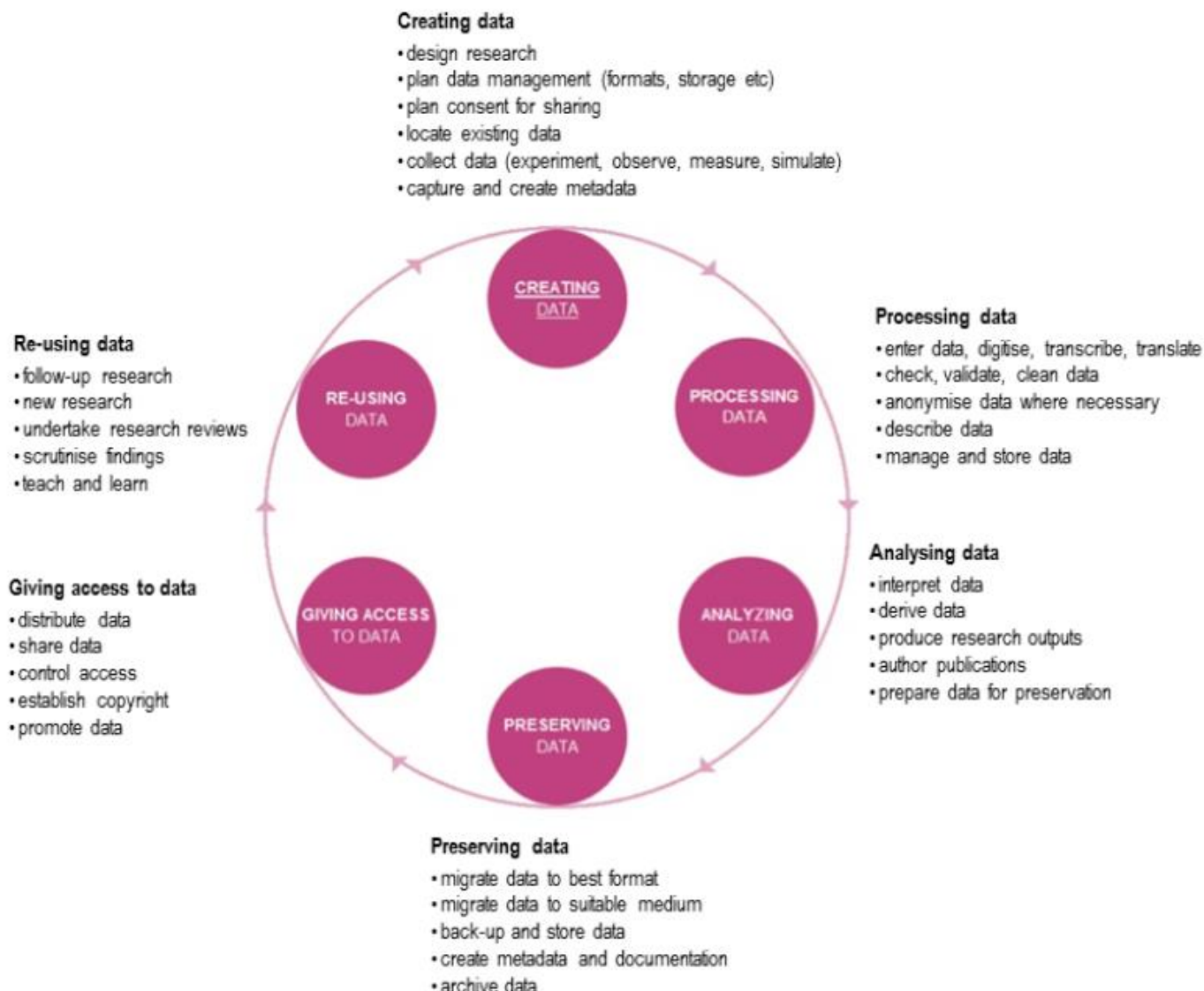
- plnenie cieľov Data Managementu a Data Governance Framework;
- Kvalita dát — popisné údaje, normy, štandardy;
- Aktuálne trendy v oblasti spracovania dát;
- Etické, právne aspekty, vlastníctvo dát;
- Prístupy k open access zdrojom;
- Sieť Data Stewardov a kustódov — transparentnosť, konzistentnosť, spolupráca;
- Priebežná kontrola a aktualizácia dát a súvisiacich služieb;
- Odozva používateľov a skvalitňovanie procesov Data Managementu.

Plán realizácie Data Stewardshipu v knižnici:

- Vytvorenie Data Stewardship Framework;
- Popis práce dátových odborníkov v knižniciach;
- Náčrt požadovaných zručností pre Data Stewardov;
- Definovanie štruktúry siete Data Stewardov a popis kompetencií;
- Implementačná stratégia pre knižnice na zlepšenie zručností zamestnancov;
- Koordinácia odborných činností podľa cieľov Data Managementu;
- Vytvorenie optimálnej siete Data Stewardov a kustódov s ohľadom na štruktúru a činnosti inštitúcie;
- Spolupráca s tvorcami Data Managementu a Data Governance Framework a účasť na školeniach.

Otvorené dáta a nové výzvy pre knižnice

Research Data Management (RDM) – Data lifecycle model



DOSTUPNOSŤ dát pre výskumníkov Integrujte súbory dát do vyhľadávacích služieb

VYHĽADATEĽNOSŤ – trvalé identifikátory, uplatnenie noriem pre jednotlivé aktivity, tvorba schém popisných dát, citačných postupov

INTERPRETOVATEĽNOSŤ – prelinkovanie medzi zdrojmi dát a súvisiacimi publikáciami, kontextové vzťahy, opätovná použiteľnosť

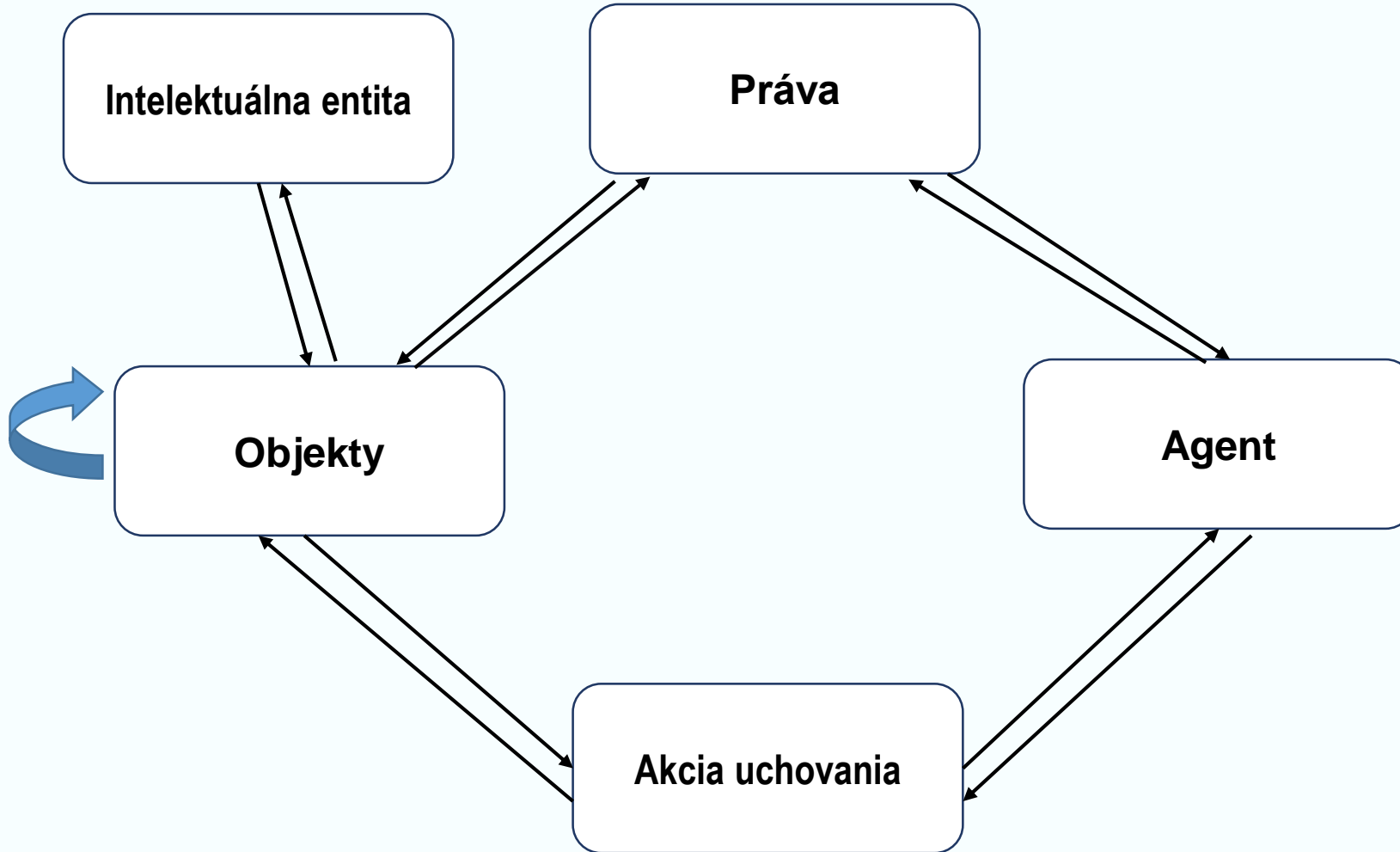
SPRÁVA A UCHOVÁVANIE SÚBOROV – archívny softvér, transparentné podmienky pre opätovné použitie súborov dát

CITOVATEĽNOSŤ - vytváranie jednotných štandardov citovania dát, podporovanie trvalých identifikátorov

KONZERVÁCIA DÁT – postupy spracovania dát, pre jednotlivé disciplíny pri vytváraní údajov (formáty uchovávania, dokumentácia experimentu,...)

spolupráca s tvorcami dát,

Dáta, dokumentácia a trvalé identifikátory



Obohatenie pôvodných údajov MARC a záznamov konvertovaných v prepojených údajoch o identifikátory z externých zdrojov.

Ak zdroje údajov spravujú tretie strany, ale sa sprístupňujú prostredníctvom inštitúcie, informácie môže poskytnúť tretia strana v dohodnutej forme, ktorá je v súlade s usmerneniami inštitúcie alebo v natívnom formáte dodávateľ'a.

Trvalé identifikátory: DOI, ROR, EISSN, EISBD, ResearcherID, ORCID, FunderID...

[The Virtual International Authority File](#)

medzinárodný súbor menných autorít. Ide o spoločný projekt, na ktorom sa podieľ'a viac národných knižníc a súborných katalógov, spravovaný Online Computer Library Center. Prvotný impulz vzišiel od Nemeckej národnej knižnice a americkej Knižnice Kongresu.

Ako sprístupniť dáta – „urobiť dáta sociálnymi“

- **Pravidlá:** jednoduchosť prístupnosti súborov, inovácie – učenie zo skúseností ostatných, úzke kontakty s potencionálnymi/používateľmi
- **Kroky:** **a/** výber súborov dát pre sprístupnenie
 - b/** použiť otvorenú licenciu: zistiť, aké práva duševného vlastníctva existujú v dátach, použiť vhodnú „otvorenú“ licenciu, ktorá licencuje všetky tieto práva a podporuje definíciu otvorenosti
 - c/** sprístupnenie dát vo vhodnom formáte, aj alternatívne, napríklad prostredníctvom rozhrania API
 - d/** urobiť dáta ľahko objaviteľným – na webe, v katalógu, sociálne médiá a siete, diskusné fórum a pod. podujatia pre prezentáciu zdrojov dát: [Open Knowledge Events](#), EU: [Open data events](#)
- **Odporúčané licencie pre otvorené dáta:** <http://opendefinition.org/licenses/>
- **Sprístupnenie dát (technická stránka):** otvorený strojovo čitateľný formát, webstránka, portál tretej strany, API, DataHub

Open data day(s)– granty pre knižnice



Open Data Day - je každoročná oslava otvorených dát na celom svete. Skupiny z celého sveta organizujú miestne podujatia v deň, kedy budú využívať otvorené dáta vo svojich komunitách. Ako spôsob pre zapojenie rôznych kultúr na celom svete, v tomto vydaní ponúkame organizáciám príležitosť usporiadať podujatie Open Data Day v najlepšom termíne pre ich miesto medzi **4. a 10. marcom**.

Tematické zameranie na rok 2023:

“Open Data to AI” úloha otvorených dát v oblasti umelej inteligencie, strojového učenia a ďalších nových technológií.

- Ako vyzerá cesta od otvorených dát k AI?
- Ako bola alebo môže byť použitá AI na riešenie naliehavých problémov?
- Ako môžu otvorené údaje a AI ilustrovať naliehavosť v rôznych podkategóriách, ako sú klimatické zmeny, transparentnosť rozpočtu, vlastníctvo, mapovanie a mnohé ďalšie?

Podrobnosti: <https://blog.okfn.org/2023/02/07/open-data-days-2023-will-take-place-from-4th-to-10th-march/>

<https://opendataday.org/#what>

<https://blogs.ifla.org/faife/2020/03/06/libraries-and-open-data/>

Open data a AI – etický prístup k jednotlivým etapám životného cyklu AI

Tvorba dát: tvorba podľa stanoveného modelu, hodnotenie AI modelu

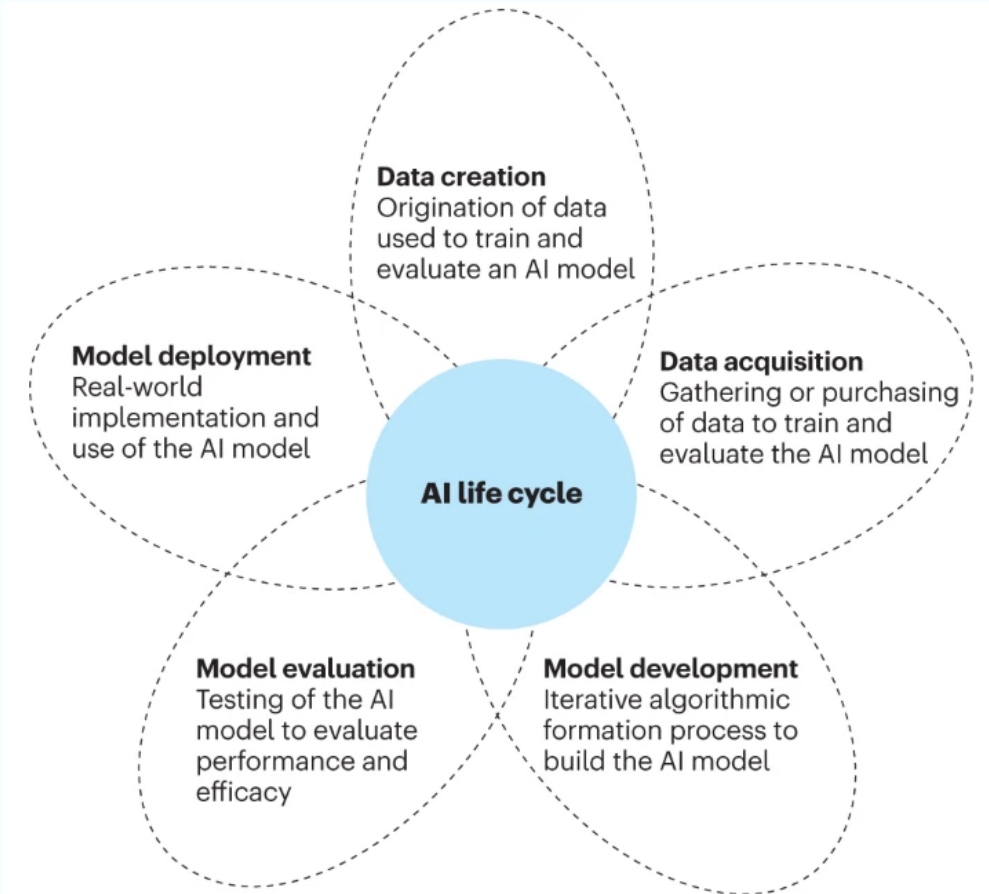
Zber dát: skúmanie pôvodu dát, hodnotenie modelu AI, informácie o získavaní údajov pre algoritmy AI, etika, transparentnosť, vhodnosť dát

Vývoj modelu AI: jasná definíciu problému alebo otázky, ktorú má algoritmus AI riešiť, koncového používateľa a modelu a príjemcu modelu. Model musí byť tiež vyladený tak, aby fungoval podľa metód hodnotenia a cieľov nasadenia. Podrobné informácie týkajúce sa technických špecifikácií modelu. (Mnohé časopisy zaviedli požiadavku na otvorený zdrojový kód, tj. aby výskumník opísal architektúru modelu AI včetně vybraných metód a parametrov a zároveň poskytol verejne dostupný zdrojový kód.)

Hodnotenie modelu: ako možno využiť model AI pre konkrétny cieľ a úlohu. Obmedzené možnosti hodnotenia môžu spôsobiť skreslenie výsledkov výskumu.

Nasadenie modelu: nasadenie modelu v konkrétnej situácii s cieľom dosiahnutia efektívneho výskumu: transparentnosť dát, miera možnosti zovšeobecnenia poznatkov, dostupnosť informácií, tj, efektívnejší výskum a poznatky.

Prepojenie jednotlivých fáz: model AI slúži ako zdroj údajov pre budúce modely.



Výskumné dáta – reprodukovateľnosť výsledkov výskumu je základným princípom vedy

[Code sharing policy](#) open license, dostupné pre každého, pre výskumníkov aj vydavateľov

- 1. Zdieľanie kódu:** poskytnutie programu na replikáciu akejkoľvek empirickej analýzy, simulácií, numerických výpočtov atď.
- 2. Zdieľanie údajov:** údaje spolu so zdrojovým kódom, v prípade obmedzení copyrightu - súbor pseudoúdajov na ilustráciu formátu súborov
- 3. Výnimky:** napríklad oneskorenie pri zdieľaní kódu, redakciou schválené výnimky uviesť v príspevku
- 4. Podrobnosti o kóde:** podľa potreby komentáre, napríklad podrobnosti o softvéri, programovacom jazyku, o formátoch údajov a zdrojoch údajov.
- 5. Použitie kódu:** používatelia sťahujúci súbory budú musieť potvrdiť, že programy sa budú používať iba na akademický výskum, a zároveň sú povinní vo svojej práci zverejniť tento zdroj
- 6. Experimenty:** pokyny k používaniu priložené v časti Diskusia
- 7. Počítačové programy:** v prílohe v plnom znení
- 8. Nespracované údaje z experimentu:** v prílohe v plnom znení

Príklady: [Review of Finance](#) [Springer Nature](#) [ELSEVIER - Sharing research data](#),
[Journal of Financial Economics](#) – experimenty



- Dáta na medzinárodnej úrovni, dátová infraštruktúra



[Zdroj: https://hri.fi/en_gb/support/what-is-open-data/]

Dáta na medzinárodnej úrovni

- Neutrálne fórum na komunikáciu globálnych pracovných skupín s cieľom podporovať zdieľanie dát a výskum založený na dátach.
- Medzinárodná organizácia založená na členstve zameraná na podporu dátovej infraštruktúry a odstraňovanie bariér dostupnosti dát.
- Pokrýva všetky fázy životného cyklu dát, spája producentov dát, používateľov a správcov, rieši výmenu, spracovanie a ukladanie dát.
- Návrhy stratégie riešenia správy dát, certifikácie dátových úložísk, technologické aspekty.
- Vytvára podmienky pre spoluprácu: záujmové pracovné skupiny na tvorbu globálnej komunity RDA prostredníctvom členstva.
- Organizujte webináre na propagáciu práce záujmovej skupiny s cieľom zvyšovať odbornosť a povedomie o problematike dostupnosti a zdieľania dát.

iRODS - open-source data management software, virtualizuje miesta, kde sú údaje uložené, robí prepojenia na zdroje robustnejšími a umožňuje prístup k údajom bez ohľadu na to, kde sú uložené.

FAIMS (**Field Acquired Information Management System**) - na offline zachytávanie údajov v súlade s FAIR princípmi v terénnom výskume



Dáta na medzinárodnej úrovni

DataCite

medzinárodná nezisková organizácia venujúca sa zlepšeniu citovania dát, vznikla v roku 2009.

Spája popredné výskumné knižnice a informačné centrá.

Ciele:

- vytvoriť jednoduchší prístup k výskumným údajom na internete
- zvýšiť akceptáciu výskumných údajov ako legitímnych, citovateľných príspevkov do vedeckého záznamu
- podpora archivácie údajov, ktorá umožní overenie výsledkov a opätovné použitie pre budúce štúdie.

Základné požiadavky

- členovia DataCite priradujú k súborom údajov trvalé identifikátory, ako sú identifikátory digitálnych objektov (DOI).
- DataCite odporúčaný formát pre citáciu údajov je: Autor (Rok vydania): Názov. Vydavateľ. Identifikátor, alebo alternatívne: Autor (Rok vydania): Názov. Verzia. Vydavateľ. Typ zdroja. Identifikátor.

Možnosť migrácie obsahu do iných služieb, napríklad pre ORCID

Dáta na medzinárodnej úrovni - príklady

Po celom svete existujú stovky portálov s otvorenými údajmi.

<https://opendatainception.io/>

V Európe Európsky dátový portál [the European Data Portal](#)

zhromažďuje popisné dáta na verejných dátových portáloch v európskych krajinách.

data.europa.eu - štúdia

[Open Data Slovakia](#)

Slovensko — príklady

Portál data.gov.sk — ústredný portál verejných služieb ľuďom, [Štatistický úrad SR](#), [Úrad geodézie a kartografie SR](#), [SHMÚ](#), [OBCHODNÝ REGISTER](#) Ministerstvo spravodlivosti SR, [Ministerstvo financií SR](#), [Úrad pre verejné obstarávanie](#),

Príklady z iných krajín:

[Portál otvorených dat](#), [Národní katalog otevřených dat](#), [Český statistický úřad](#), [DBpedia](#)

Medzinárodné projekty pre výskumné dáta - infraštruktúra



3.6 - Strategy for a European Data Infrastructure

Kritériá: nepretržitý prístup a dlhodobé uchovávanie

Infraštruktúru tvoria:

- národné archívy, webové archívy, digitálne repozitáre, dátové centrá a repozitáre

Problematika súvisiaca s archiváciou dát:

- Vhodné technológie, archivácia do cloudu alebo využívajú poskytovateľa SAAS (Software as a Service), webová archivácia
- Zachovanie významu a kontextu údajov, udržiavanie dôvery v dáta, vyrovnanie sa so záplavou dát – výber dát
- Spolupráca tvorcov dátových zdrojov, vytváranie konzorcií
- Súlad s právnymi predpismi: Európsky rámec pre audit a certifikáciu ISO 16363, Štandard pre dôveryhodné digitálne úložiská
- Zachovanie pôvodných digitálnych materiálov ISO 8601
- Normy pre akvizíciu a prenos producent-archív – metodologický abstraktný štandard ISO 20652

štandardom rozhrania medzi tvorcami a archívami ISO 20104

Štandardy katalogizácie a dokumentácie IFLA- USTANOVENIE MEDZINÁRODNÝCH PRINCÍPOV KATALOGIZÁCIE

ISO 16919 Požiadavky na orgány zabezpečujúce audit a certifikáciu dôveryhodných digitálnych úložísk kandidátov (ISO, 2011), ktorá stanovuje požiadavky na akúkoľvek organizáciu, ktorá certifikuje audítorov podľa ISO 16363,

ISO 17021 Požiadavky na orgány zabezpečujúce audit a certifikáciu systémov manažérstva (ISO, 2012a) poskytuje mechanizmus pre akreditačné orgány auditu.

Medzinárodné projekty pre výskumné dáta - infraštruktúra

Kritériá: nepretržitý prístup a dlhodobé uchovávanie

CLOCKSS (Controlled LOCKSS) je nezisková spolupráca medzi knižnicami a vydavateľmi. Ide o archív založený na softvéri LOCKSS, v ktorom obmedzený počet knižníc preberá úlohu archivácie v mene širšej komunity. Poskytuje knižniciam poistenie, že e-časopis a ďalší obsah, ktorý si predplatili, budú dlhodobo uchované. Je opísaná ako „súkromná sieť LOCKSS“.

LOCKSS (Lots of Copies Keep Stuff Safe) poskytuje knižniciam nástroje a podporu s otvoreným zdrojovým kódom, aby mohli prevziať miestnu správu širokej škály materiálov vrátane predplatného a vedeckých aktív s otvoreným prístupom (knihy, časopisy atď.). Čitatelia majú prístup k obsahu chránenému LOCKSS vždy, keď (a z akéhokoľvek dôvodu) nie je možné materiál zobrazit' na serveroch vydavateľa (alebo sprostredkovateľa). Vysoko distribuovaný charakter tohto prístupu má za cieľ zabezpečiť dostatočnú replikáciu na ochranu obsahu napriek akýmkoľvek potenciálnym katastrofám, ktoré by mohli postihnúť jednotlivé inštitúcie LOCKSS.

Open Research Platforms: <https://openscience.eu/open-research-platforms-part-1>
<https://openscience.eu/open-research-platforms-part-2>

Zdroje k téme: Code Sharing Policy

- The American Finance Association:

https://afajof.org/wp-content/uploads/files/policies-and-guidelines/CodePolicy_FAQ.pdf

<https://afajof.org/wp-content/uploads/files/policies-and-guidelines/CodePolicy.pdf>

- European Finance Association: <http://revfin.org/code-sharing-policy/>

- SpringerNature: <https://support.springernature.com/en/support/solutions/articles/6000237619-software-and-code-sharing>
<https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards>

<https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards>

- Center for Global Development: <https://www.cgdev.org/sites/default/files/CGD-Data-code-transparency-policy.pdf>

- PLOS ONE: Materials, Software and Code Sharing, <https://journals.plos.org/plosone/s/materials-software-and-code-sharing>

How does mandated code-sharing change peer review? An interview with Feilim Mac Gabhann, Jason Papin, and Virginia Pitzer

<https://theplosblog.plos.org/2023/09/code-policy-interview/>

- Taylor & Francis journals: [policies on data sharing](#) , <https://authorservices.taylorandfrancis.com/data-sharing-policies/>
[Ethical considerations](#) of data sharing

<https://authorservices.taylorandfrancis.com/wp-content/uploads/2022/10/Data-sharing-understanding-our-basic-data-sharing-policy-infographic.jpg>



Zdroje k téme:

[The European Data Act](#)

[Metodická príručka pre povinné osoby – Ako zverejňovať otvorené dáta](#)

Integračný manuál Open Data API ITMS2014+. Úrad podpredsedu vlády SR pre investície a informatizáciu. 2018. Centrálny koordinačný orgán EU.

The Open Data Handbook, <https://opendatahandbook.org/guide/en/>

[Creative Commons and Data](#)

[DataCite](#) - medzinárodná členská organizácia, ktorej cieľom je: vytvoriť jednoduchší prístup k údajom z výskumu; zvýšiť akceptáciu výskumných údajov ako legitímnych príspevkov vo vedeckom zázname; a na podporu archivácie údajov. DataCite priraduje trvalé identifikátory DOI k súborom údajov.

[Dat Ecosystem](#) - bezplatný softvér na rýchle a šifrované zdieľanie a ukladanie súborov, ktorý umožňuje sledovať zmeny a umožňuje používateľom kontrolovať, kde sa ich súbory uchovávajú.

[Open Data - from SPARC](#) - otvorené dátové zdroje od SPARC, koalície pre vedecké publikovanie a akademické zdroje.

[Open Access Directory -- Data Repositories](#) – zoznam databáz a repozitárov pre otvorené dáta

[r3data.org and the Repository Finder](#) - register úložísk výskumných údajov, spravuje DataCite

[Get the Data](#) - Online komunita

[Google Dataset Search](#) - verejný adresár údajov spoločnosti Google poskytuje overené súbory údajov v grafových a textových formátoch

[UN Data](#) - oficiálne medzinárodné údaje Organizácie Spojených národov.

Zdroje dát v EU: <https://data.europa.eu/en>

[Open Science at the Core of Libraries](#)

Open Data to Support the Sustainable Development Goals, <https://opendatawatch.com/publications/open-data-to-support-the-sdgs/>

[EU AI Act: first regulation on artificial intelligence](#)

Ďakujem za pozornosť

Simona Hudecová, simona.hudecova@cvtisr.sk
Centrum vedecko-technických informácií SR
Odbor podpory otvorenej vedy a výskumu
Lamačská cesta 8/A, 811 04 Bratislava



<https://eiz.cvtisr.sk/>

<https://www.cvtisr.sk/>