

OTVORENÁ VEDA V PRAXI:

Otvorené vedecké dáta

OPAKOVANIE

Otázky, ktoré sa môžu vyskytnúť na písomnej skúške

**Aký je význam otvorenosti vo výskume a čo
môžeme robiť s otvorenými materiálmi?**

Aký je význam otvorenosti vo výskume a čo môžeme robiť s otvorenými materiálmi?

- Otvorenosť zvyšuje **efektivitu** výskumu (časovú, ekonomickú), **transparentnosť, integritu, spravodlivosť, etiku, viditeľnosť a dosah** výskumu, podporuje **kultúru spolupráce a zdieľania** v akademickej komunite;
- K otvoreným materiálom máme **prístup** (môžeme si ich prezerať, stiahnuť si ich do počítača), môžeme ich **používať** (na výskum, vyučovanie, a ak to umožňuje licencia, tak napr. aj na podnikanie), **pozmeniť** (napr. preložiť, aktualizovať, vylepšiť) a **zdieľať** (šíriť, dať si to na stránku) bez zbytočných bariér (technických, právnych).

**Čo všetko môžu vedci v rámci výskumu otvorene
zdieľať?**

**Uvedte aspoň dva príklady vhodných miest, kde to
môžu zdieľať.**

Čo všetko môžu vedci v rámci výskumu otvorene zdieľať? Uvedte aspoň dva príklady vhodných miest, kde to môžu zdieľať.

Čo zdieľať?

- Metodiky,
- Publikácie,
- Vzdelávacie materiály,
- Dáta,
- Recenzné posudky.

Kde zdieľať?

- Publikácie, dáta, posudky: **v OA časopisoch** (zoznam napr. v DOAJ) a **publikačných platformách ako ORE**,
- Publikácie, dáta, vzdelávacie materiály, metodiky: **v repozitároch**, napr. Zenodo (všeobecný), alebo **protocols.io** (metodiky),
- Úzko odborne zamerané dáta: **v špecializovaných medzinárodných databázach** (napr. GBIF – global biodiversity information facility - pre dáta o biodiverzite).

Čo je reprodukovateľnosť a replikovateľnosť?

Prečo sú pre vedu dôležité?

Čo je reprodukovateľnosť a replikovateľnosť? Prečo sú pre vedu dôležité?

- **Reprodukovateľnosť:** realizácia výskumu a komunikácia výsledkov takým spôsobom, ktorý **umožňuje zopakovať identické analýzy a s použitím tých istých dát** získať **konzistentné výsledky**;
- **Replikovateľnosť:** realizácia výskumu a komunikácia výsledkov takým spôsobom, ktorý umožňuje **zopakovať celý výskum, pomocou identických** (alebo v rámci možností podobných) **metód nazbierať nové dáta**, rovnako ich **analyzovať** a získať **konzistentné výsledky**;
- **Význam:** Reprodukovateľnosť a replikovateľnosť **zvyšujú spoľahlivosť vedeckých výsledkov**, pretože **umožňujú vedecké štúdie zopakovať a overiť**.

**Uvedte aspoň päť opatrení, ktoré pomáhajú
zvýšiť reprodukovateľnosť a replikovateľnosť
výskumu.**

Uvedte aspoň päť opatrení, ktoré pomáhajú zvýšiť reprodukovateľnosť a replikovateľnosť výskumu.

1. **Predregistrácia** výskumu, **plánovanie manažmentu dát**,
2. Vedci sa pri zbere dát starajú o **dôslednú dokumentáciu**, v čom ich podporujú inštitúcie aj financovatelia,
3. Vedci pri publikovaní **zdieľajú dáta, softvér a detailné metodiky**,
4. Vedci sa **vzdelávajú** a nezanedbávajú **tréning a mentoring v oblasti integrity a reprodukovateľnosti** pre študentov a mladých pracovníkov,
5. Časopisy v rámci svojich politík vyžadujú, aby vedci **dôsledne citovali všetky použité zdroje** (vrátane dát, softvéru a pod.) a opísali, ako a **kde budú dostupné dáta/softvéry/iné materiály**, ktoré v rámci výskumu vytvorili (napr. Transparency and Openness Promotion od Open Science Framework),
6. Financovatelia v rámci svojich politík / mandátov vyžadujú zdieľanie otvorených publikácií a dát, plánovanie manažmentu dát a pri určitých druhoch výskumu aj predregistráciu,
7. Dohľad na dodržiavanie kritérií je nezávislý,
8. **Financovanie je dlhodobé**, aby umožnilo výskum do hĺbky problematiky,
9. Inštitúcie a financovatelia **podporujú nielen nový výskum, ale aj replikovanie a overovanie existujúceho výskumu** a otvorenú vedu všeobecne,
10. Inštitúcie na svojej pôde organizujú pre študentov aj starších vedeckých pracovníkov **pravidelné tréningy. VŠ majú štandardne v osnovách vzdelávanie o otvorenej vede**, manažmente dát, integrite a reprodukovateľnosti / replikovateľnosti výskumu, majú systémy, zdroje a nástroje na zabezpečenie kvality výskumu,
11. **Systém hodnotenia vedy odmeňuje nielen publikácie, ale aj zdieľanie dát, mentoring a ďalšie praktiky prospievajúce replikovateľnosti výskumu.**

**Uved'te aspoň tri všeobecné alebo konkrétne
príklady výskumných dát**

Uveďte aspoň tri všeobecné alebo konkrétne príklady výskumných dát

- **primárne dáta:** pôvodné údaje (meranie, text, obraz, zvuk, video, dotazníky atď.) zozbierané alebo vytvorené na účely realizácie výskumného projektu
- **existujúce dáta** zozbierané alebo skopírované na okamžité alebo budúce použitie v projektoch (najmä administratívne alebo štatistické údaje, digitalizovaný obsah zo zbierok, údaje dostupné v databázach výslovne sprístupnených výskumnej komunite);
- **všetky nové (sekundárne) dáta**, ktoré sú výsledkom spracovania (analýzy, agregácie, transformácie atď.) primárnych dát.

**Uved'te aspoň dve výhody otvoreného
publikovania výskumných dát**

Uveďte aspoň dve výhody otvoreného publikovania výskumných dát

- zabezpečenie dlhodobého, trvalého prístupu k vedeckým údajom
- zvyšovanie dosahu a viditeľnosti výskumu
- zvýšenie efektivity výskumu
- rozpoznanie a uznanie kreditu pre pôvodného autora/ov
- podpora reprodukovateľnosti a transparentnosti

**Uvedte aspoň dve kritéria pri výbere
spoľahlivého dátového repozitára**

Uvedte aspoň dve kritéria pri výbere spoľahlivého dátového repozitára

- dlhodobý udržateľný
- ukladá dáta bezpečným spôsobom
- dáta vyhľadateľné (napr. pomocou PID), prístupné a opakovaná použiteľné
- používa prijaté štandardy metadát
- vkladateľ môže určiť licenciu, ktorou sa riadi prístup k dátam a ich opakované použitie

Na čo slúžia dátové časopisy?

Na čo slúžia dátové časopisy?

- publikácie, ktorých hlavným cieľom je zverejňovať súbory údajov

Uved' aspoň jednu službu, ktorú poskytuje

DataCite

Uved' aspoň jednu službu, ktorú poskytuje DataCite

- registrácia trvalých identifikátorov (DOI)
- platforma pre vyhľadávanie dát
- vytvorenie citácií
- extrahovanie metadát z DOI
- štatistické prehľady
- zdieľanie dát

Čo znamená skratka FAIR?

Čo znamená skratka FAIR?

Findable – nájdniteľné

Accessible – dostupné

Interoperable – interoperabilné

Reusable – opakovane využiteľné

Na čo slúži ORCID?

Na čo slúži ORCID?

Slúži na jednoznačnú identifikáciu vedecko-
výskumných pracovníkov.

**Ktoré osobné údaje považujeme za citlivé?
Uvedte aspoň 3.**

Ktoré osobné údaje považujeme za citlivé?

- osobné údaje, ktoré odhaľujú **rasový alebo etnický pôvod, politické názory, náboženské alebo filozofické presvedčenie**
- **genetické údaje, biometrické údaje** spracúvané len na účely identifikácie ľudskej bytosti
- **zdravotné údaje**
- údaje týkajúce sa **sexuálneho života alebo sexuálnej orientácie** danej osoby

**Akými spôsobmi sa dajú ochrániť citlivé
údaje?**

Akými spôsobmi sa dajú ochrániť citlivé údaje?

Anonymizácia

Pseudonymizácia

Enkrypcia/Kódovanie

Čo znamená skratka GDPR?

Čo znamená skratka GDPR?

General Data Protection Regulation

Všeobecné nariadenie o ochrane údajov

Ako často sa môže meniť/modifikovať DMP?

Ako často sa môže meniť/modifikovať DMP?

V priebehu projektu – podľa potreby.

Aký je ideálny formát na uchovávanie dát?

Uvedte aspoň 3 vlastnosti.

Aký je ideálny formát na uchovávanie dát? Uvedte aspoň 3 vlastnosti.

- Ne proprietárny
- Nešifrovaný
- Ne komprimovaný
- Bežne používaný výskumnou komunitou
- V súlade s otvorenými štandardami

Zálohovanie - čo je to pravidlo 3-2-1 ?

Zálohovanie - čo je to pravidlo 3-2-1 ?

Zásada zálohovania v skratke – **koľko záloh** vytvárať, na akých **typoch médií** a ako a **kde uchovávať pamäťové zariadenia**.

- aspoň **3 záložné kópie dát**
- aspoň **2 rôzne typy pamäťových médií**
- pamäťové zariadenia uchovávať na **oddelených miestach** a aspoň **1 mimo lokality**,
- pravidelne kontrolujte, či fungujú
- poznať a dodržiavať postupy

**Čo je to konvencia pomenovania/file naming
convention?**

Čo je to konvencia pomenovania/file naming convention?

File Naming Conventions (FNC) je **rámec na systematické pomenovanie súborov** spôsobom, ktorý opisuje, čo obsahujú a ako súvisia s inými súbormi.

File Naming Conventions (FNC) je **stratégia/dohoda o vhodnom pomenovaní priečinkov, zložiek, súborov.**

**Aké licencie na dáta poznáte (okrem CC0 a
CC BY)?**

Uved'te aspoň dve

Aké licencie na dáta poznáte (okrem CC0 a CC BY)? Uveďte aspoň 2.

- **Open Data Commons Open Database License (ODbL)** – “Attribution Share-Alike for data/databases”
- **Open Data Commons Attribution License (ODC-By)** – “Attribution for data/databases”
- **Open Data Commons Public Domain Dedication and License (PDDL)**
- **Open Government Licence (OGL)**
- *Non-Commercial Government Licence for public sector information (NGL)*
- **Community Data License Agreement - CDLA Permissive 2.0**
- **Community Data License Agreement - CDLA Sharing V-1.0**

Na čo slúži trvalý identifikátor/Persistent Identifier,
aké sú jeho funkcie?

Uvedte aspoň 3 funkcie.

Na čo slúži trvalý identifikátor/Persistent Identifier, aké sú jeho funkcie?

Funkcie trvalého identifikátora – možné odpovede:

- Identifikácia (jednoznačné označenie objektu), lokácia, spĺňa požiadavku dlhodobej dostupnosti prostredníctvom odkazu, objekt s PID je možné vždy nájsť, a to aj v prípade zmeny názvu a umiestnenia
- PID podporujú implementáciu dátovej politiky a efektívne prispievajú k manažmentu informácií on-line prístup k digitálnemu objektu
- Možnosť zabrániť vytvoreniu nefunkčných odkazov alebo nenájdeniu stránky
- Podporujú viditeľnosť a dostupnosť dát a prepojenosť PID – viditeľnosť na webe, v sociálnych médiách a iných e-zdrojoch (profil ORCID, publikačná činnosť)
- Jednoduché a včasné zdieľanie dát a výstupov výskumu - tj. zdieľanie informácií medzi rôznymi systémami

Uvedte aspoň tri trvalé identifikátory?

Otázka: uveďte aspoň tri trvalé identifikátory?

Odpoveď - možnosti:

- International Standard Name Identifier (ISNI)
- Open Researcher and Contributor Identifier (ORCID ID)
- WOS ResearcherID
- The International Standard Serial Number (ISSN)
- The International Standard Book Number (ISBN)
- Archival Resource Key (ARK)
- Digital Object Identifier (DOI)
- Persistent Uniform Resource Locator (PURL)
- Funder ID v databáze Crossref
- Global Research Identifier Database (GRID) ID transformácia do
- Research Organization Registry (ROR) ID
- Ringgold -Identifikátor Ringgold 2003, (USA, UK)
- DataCite – prideľuje trvalé identifikátory po registrácii v DataCite Fabrica

Na čo slúži Research Organization Registry (ROR)?

Na čo slúži Research Organization Registry (ROR)?

Odpoveď:

Research Organization Registry (ROR) ID – 2021, sú „globálne jedinečné, trvalé, objaviteľné a resolver“ identifikátory pre výskumné organizácie.

Identifikátory ROR uchovávajú popisné dáta o organizáciách (alternatívne mená/skratky, externé adresy URL a iné identifikátory, ako sú Wikidata, ISNI a Open Funder Register).

Aký je rozdiel medzi Data Governance a Data Stewardship?

Aký je rozdiel medzi Data Governance a Data Stewardship?

Odpoveď:

Data Governance stanovuje: ciele, procesy, metriky a štandardy. Stanovuje úlohy a súvisiace zodpovedností, 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 správcom, takže správcovia dát predstavujú záujmy ostatných a zabezpečujú, aby sa práca súvisiaca s dátami vykonávala v súlade so zásadami a postupmi určenými prostredníctvom správy. Cieľom je spracovanie a konsolidácia dát v súlade s medzinárodnými normami a štandardmi.

Dlhodobé uchovávanie dát:

1. Ktoré dáta uchovávať dlhodobo?

2. Ktoré dáta sa môžu vymazať?

3. Ako dlho po skončení projektu dáta uchovávať?

Dlhodobé uchovávanie dát:

1)Ktoré dáta uchovávať dlhodobo?

všetky podkladové dáta súvisiace s publikáciou;
dáta, ktorých zber je náročný alebo nákladný;
dáta, ktoré sa nedajú nahradiť/znovu vygenerovať

2)Ktoré dáta sa môžu/musí vymazať?

biologický materiál, ktorý sa časom znehodnotí;
dôverné údaje alebo osobné údaje, ktoré nemožno anonymizovať

3)Ako dlho uchovávať po ukončení projektu

V závislosti od nastavenia politiky/usmernenia inštitúcie alebo legislatívy /
dodržiavať právne predpisy, príp. nariadenia platné pre váš typ dát, pravidlá či
usmernenia inštitúcie.

Čo znamená skratka EOSC?

Čo znamená skratka EOSC?

European Open Science Cloud

Ako sa volá legislatívny dokument Európskej komisie, ktorý vytvára rámec fungovania dátového trhu v Európskej únii?

Ako sa volá legislatívny dokument Európskej komisie, ktorý vytvára rámec fungovania dátového trhu v Európskej únii?

European Data Act

**Aké požiadavky majú granty na vedecké projekty
programu Horizon Europe na dáta?**

Aké požiadavky majú granty na vedecké projekty programu Horizon Europe na dáta?

- Vytvorenie dokumentu Plán manažmentu dát (Data Management Plan) a jeho napĺňanie pri realizácii projektu
- Uplatnenie FAIR princípov na dáta
- Tak otvorene ako je možné, tak zatvorené ako je potrebné

**Musím dáta z projektov Horizon Europe
plne sprístupniť?**

Musím dáta z projektov Horizon Europe plne sprístupniť?

Nie, uplatňuje sa princíp:

“As open as possible, as closed as necessary”

Vedec sám určí, do akej miery sú jeho dáta prístupné, prístup sa môže obmedziť. Treba však uviesť dôvod.

Čo sú metadáta a na čo slúžia?

Čo sú metadáta a na čo slúžia?

- Dáta o dátach.
- Pomáhajú **identifikovať, lokalizovať, citovať a použiť výskumné dáta** aj v budúcnosti akýmkoľvek používateľom.

Uved'te aspoň tri metádatové štandardy

Uved'te aspoň tri metádatové štandardy

Napríklad:

- DublinCore,
- Marc 21.....
- [Darwin Core](#)
- [DDI](#)
- [EML](#)
- [VRA Core](#)

Čo sú riadené slovníky a na čo slúžia?

Čo sú riadené slovníky a na čo slúžia?

- Riadené slovníky sú **štandardizované a organizované usporiadania slov a fráz.**
- Poskytujú **konzistentný spôsob opisu dát** a **umožňujú zdieľať dáta jednoduchšie** s výskumníkmi z rovnakej disciplíny.