

Otázky:

TAČR (obsahuje základní otázky, na které je třeba vždy odpovědět, jednotlivé otázky jsou číslované, pod nimi jsou vysvětlivky či nápověda či vzorový text):

Popis dat:

1. Popište, jaká data budete vytvářet nebo sbírat.

Z jakých přístrojů, v jakém formátu, v jakém SW, jen nová data či stará, jen vlastní data či přejatá, jaká bude odhadovaná velikost dat, počet souborů (výhodou je, pokud použijete cizí otevřená data - vždy uvést!). Pokud můžete vytvořit tabulku, tak ji vytvořte, např.:

Type of data	File format(s)	How will data be collected (for re-used data: source and terms of use)?	Purpose of processing	Storage location	Who will have access to the data
3D scan of the complete face and/or facial features, without texture mesh	.stl files	3D scanning	To use as benchmark for other scanning technologies	Surf drive	The project team
Age, gender, disease	.csv files	Survey	To place the results into context and understand the differences between different subjects	Surf drive	The project team
Photo's and video's of participants face	.mp4 and .jpg files	Digital camera	To use as input for photogrammetry 3D reconstruction	Surf drive	The project team

Zdroj:

https://dmponline.dcc.ac.uk/plans/111030/export.pdf?export%5Bquestion_header%5D=true

2. Popište, jaká data budete zpracovávat a jakým způsobem. Uveďte, zda se bude jednat o data vlastní či externí a jakým způsobem je získáte.

Kde budou uložena data během projektu, jak budou zabezpečena či případně otevřená a kdo k nim bude mít přístup?¹ Jak budou analyzována (v jakých konkrétně programech - komerčních či volně dostupných či už zakoupených? Totéž ohledně metadat! Totéž jak to bude na závěr či po ukončení projektu. (Je vhodné uvést, jak budou data použita, jaký bude výstup. Pokud bude jen publikace, pak zmínit, že je v plánu publikace.)

¹ doporučení pro užívání úložišť viz <https://it.muni.cz/prehledy/doporuceni-pro-uzivani-ulozist>

Správa dat:

3. Jak zajistíte, že data budou vyhledatelná? Budou data ve strojově čitelném formátu? Budou mít perzistentní identifikátor? Opatříte data metadaty ve standardním formátu? Budou i metadata vyhledatelná?

Popsat formáty dat a komplikace s nimi. (Výhodný je otevřený formát, např. CSV, ale pokud by došlo při převodu z původního formátu ke ztrátě informace, pak tuto informaci zdůraznit a napsat, že z tohoto důvodu se použije původní formát dat! Výhodné jsou formáty, které se dají zpracovávat ve volně dostupném software, opět zdůraznit, opět pokud by byl nutný komerční software, tak je vhodné uvést důvod). Data budou opatřena metadaty dle směrnice TUL..., která předepisuje metadatový popis dle normy ČSN ISO 690:2022 a metadatový standard DataCite Metadata Schema 4.4. Data budou opatřena identifikátorem (rozhodnout dopředu a uvést konkrétní - buď DOI či handle, který přidělí Univerzitní knihovna TUL - Markéta Trykarová, nebo DOI, který přidělí zvolený repozitář). Metadata budou indexovaná a vyhledatelná standardními nástroji (Google Scholar aj.²)

4. Budou data přístupná bez omezení komukoliv? Pokud ne, uveďte důvod a podrobněji omezení popište. Uložíte data do repozitáře? Jakého?³ Jak dlouho budou data dostupná? Budete uplatňovat časové embargo na zveřejnění dat? Pokud ano, jak bude dlouhé a jaký pro něj bude důvod? Předpokládáte nějaké další omezení sdílení dat (např. poskytnutí dat na žádost)?
 - a. Otevřená data - Nejlepší je uvést konkrétní oborový repozitář, který poskytuje DOI (vyhledávač viz <https://www.re3data.org/>), dále Národní repozitář pro data (viz <https://data.narodni-repozitar.cz/>, přiděluje DOI⁴), následně Zenodo (<https://zenodo.org/>, přiděluje DOI); zdůraznit, že cílíme na trvalou archivaci bez časového embarga a s úplným otevřením dat.
 - b. Data se nedají ihned otevřít - Uložíme a otevřeme metadata, ovšem samotná data plánujeme zveřejnit až po embargo XXX měsíců, z důvodu právních a z důvodu ochrany know-how firmy (ochrany duševního vlastnictví formou patentu). Případně uvést, že poskytneme data na vyžádání. Nejlepší je vybrat konkrétní oborový repozitář, který poskytuje DOI (vyhledávač viz <https://www.re3data.org/>), dále Národní repozitář pro data (viz <https://data.narodni-repozitar.cz/>, přiděluje DOI), následně Zenodo (<https://zenodo.org/>, přiděluje DOI); zdůraznit, že cílíme na trvalou archivaci a že metadata budou dostupná.

² Lze přidat další důležité vyhledávače, podle toho, kde je konkrétní repozitář indexovaný - např. OpenAIRE (<https://explore.openaire.eu/search/find?qf=true&active=result>), BASE (<https://www.base-search.net/>), DataCite (<https://search.datacite.org/>).

³ více info viz <https://openscience.cuni.cz/OSCI-67.html>

⁴ návod viz https://du.cesnet.cz/cs/navody/narodni_repozitar/start

5. Jakými zásadami pro zajištění interoperability se budete řídit? Jaké použijete slovníky, standardy, formáty nebo metodiky pro umožnění šíření a opětovného použití dat ve vašem oboru i mimo něj?

Dodržíme pravidla FAIR a “as open as possible and as closed as necessary”. Pro zajištění interoperability metadat i dat použijeme oborový slovník/tezaurus/standard XXX (uvést ideálně nejužívanější terminologický slovník v daném oboru, např. medicína Medical Subject Headings⁵, zemědělství AGROVOC⁶, pedagogika ERIC Thesaurus⁷, umění tezaury Getty⁸, pro biologické vědy existuje vyhledávač terminologií - <https://bioportal.bioontology.org/recommender>, existují také další vyhledávače tezurů viz <https://bartoc.org/>, <https://lov.linkeddata.es/dataset/lov/vocabs>. Obecné tezaury např. NASA⁹ či WikiData¹⁰. Doporučuji zmínit také konkrétní oborové ontologie viz https://en.wikipedia.org/wiki/Ontology_%28computer_science%29#Published_examples, ontologie z oboru medicíny viz <https://bioportal.bioontology.org/ontologies>). Dodržíme metadatový standard dle směrnice univerzity (viz předchozí bod), který je kompatibilní s normou ČSN ISO 690:2022 a schématem DataCite Metadata Schema 4.4. Zpřístupníme data v otevřených formátech (vynechat, pokud není možné toto dodržet), dbát budeme na identifikátory autorů a projektů, linkování na publikace a další výstupy vč. systémových požadavků a vztahů k dalším výsledkům.

6. Budou data volně dostupná pro jejich opětovné použití třetí stranou, zejména po skončení projektu? Opatříte vaše data standardními licencemi pro opětovné použití?
- Otevřená data - Po skončení projektu zpřístupníme všechna data. Licenci primárních dat předpokládáme CC-BY ve verzi 4.0 nebo CC0 všude tam, kde je to účelné, opakované využití zveřejněných dat tedy bude bez nutnosti individuální autorizace přístupu.
 - Data se nedají otevřít - zdůvodnit proč (komercializace, ochrana know-how, speciální formáty dat, které jsou nepoužitelné aj.!

⁵ <https://www.nlm.nih.gov/mesh/meshhome.html>

⁶ <https://agrovoc.fao.org/>

⁷ <https://eric.ed.gov/?faq-thesaurus>

⁸ <https://www.getty.edu/research/tools/vocabularies/>

⁹ <https://www.sti.nasa.gov/nasa-thesaurus/>

¹⁰ <https://www.wikidata.org/>

Ukládání dat:

7. Popište jakým způsobem budete data ukládat během řešení projektu a po jeho skončení. Uveďte, do jakého druhu (oborový, institucionální, mezinárodní, apod.) repozitáře plánujete data ukládat, popř. uveďte konkrétní repozitář.

Je rozepsáno v předchozích bodech.

8. Popište, jakým způsobem budou data zabezpečena (např. citlivé či osobní údaje).

V projektu nevzniknou a ani nebudou zpracovávána žádná citlivá data vyžadující omezení či anonymizaci. Data i metadata budou ukládána v certifikovaných repozitářích, důkladně bude overěna a nastavena možnost přístupu pro jednotlivé typy uživatelů během projektu a po projektu. Ideální je uvést např. konkrétní zabezpečení u místních a online úložišť.

Náklady:

9. Jaké budou náklady na správu dat? Jak tyto náklady pokryjete?

Zamyslete se nad všemi náklady - na hardware, software, repozitář, změny formátů dat, osobní náklady na správce dat (data stewarda) aj. Uveďte zdroje financování těchto nákladů.¹¹

Odpovědná osoba:

10. Kdo je odpovědný za správu dat? Uveďte prosím jeho pozici/roli v projektu. Bude tato osoba součástí řešitelského týmu?

Pro správnou správu dat zřídíme pozici data stewarda, který bude odpovídat za dodržování běžné mezinárodní praxe v oblasti datového stewardshipu.

Závěr:

11. Pokud byste chtěli dodat jakékoli další informace vztahující se k Plánu správy dat, uveďte je prosím níže.

Doporučuji uvést, jak Vaše otevřená data budou moci být využita v budoucnu, jakou budou mít hodnotu, zda jsou unikátní aj. Např. takto:
Zveřejněná data budou nadále užitečná pro statistické vyhodnocení a optimalizaci firemních/vědeckých postupů a technologií v rámci týmu i u dalších výzkumných institucí a firem z oboru.

¹¹ Pokud využijete Národní repozitář pro data, pak jeho využití je zdarma v rámci infrastruktury CESNET. Zenodo je také zdarma.

Oblast otevřené vědy se bouřlivě vyvíjí. Vzhledem k aktuálně probíhajícím a připravovaným aktivitám (např. plánované směrnici pro otevřená data na TUL, postupnému zavádění Národní datové infrastruktury) budeme tuto oblast pečlivě monitorovat a v případě kladného vyhodnocení využívat aktuální infrastrukturu (např. Národní datový repozitář, či přidělování perzistentních identifikátorů) dle aktuálních směrnic a doporučení.

XX

Horizon (obsahuje detailní otázky, opět číslované; pokud znáte odpovědi, je výhodné je uvést):

Shrnutí dat:

12. Použijete znovu některá stávající data? K čemu je znovu použijete? Uveďte důvody, pokud bylo opakované použití stávajících dat zvažováno, ale zamítnuto.
13. Jaké typy a formáty dat budou v rámci projektu generovány nebo znovu použity?
14. Za jakým účelem se budou data generovat nebo opětovně používat ve vztahu k cílům projektu?
15. Jaká je předpokládaná velikost dat, které hodláte generovat nebo znovu použít?
16. Jaký je původ dat, ať už vytvořených nebo znovu použitých?
17. Pro koho mohou být vaše data užitečná ("užitečnost dat") kromě vašeho projektu?

FAIR data:

18. Budou data opatřena perzistentním identifikátorem?
19. Budou poskytnuta bohatá metadata umožňující vyhledávání? Jaká metadata budou vytvořena? Jaké disciplinární /oborové nebo obecné standardy budou dodržovány? V případě, že ve vašem oboru neexistují metadatové standardy, uveďte, jaký typ metadat bude vytvořen a jakým způsobem.
20. Budou v metadatach uvedena klíčová slova pro vyhledávání, aby se zlepšila možnost vyhledávání a následného potenciálního opakovaného použití?
21. Budou metadata nabízena takovým způsobem, aby je bylo možné automaticky zpracovávat (harvestovat) a indexovat?
22. Budou data uložena v důvěryhodném repozitáři?
23. Prověřili jste vhodná řešení s daným repozitářem, kde budou vaše data uložena?
24. Zajišťuje repozitář, přidělení identifikátoru k datům? Umožní repozitář přiřadit identifikátor k digitálnímu objektu?
25. Budou všechna data volně dostupná? Pokud určité soubory dat nelze sdílet (nebo je třeba je sdílet za podmínek omezeného přístupu), vysvětlete proč, a jasně oddělte právní a smluvní důvody pro záměrné omezení. Mějte na paměti, že v projektech s více příjemci je také možné, aby si konkrétní příjemci ponechali data uzavřená, pokud je jejich zpřístupnění dat v rozporu s jejich oprávněnými zájmy nebo jinými omezeními podle grantové dohody.

26. Pokud je uplatňováno embargo, které má poskytnout čas na publikování nebo na ochranu duševního vlastnictví (např. patentů), uveďte, proč a jak dlouho bude časové embargo platit, s ohledem na to, že významná data by měla být zpřístupněna co nejdříve.
27. Budou data přístupná prostřednictvím bezplatného a standardizovaného přístupového protokolu?
28. Pokud existuje omezení k užití dat, jak bude zajištěn přístup k datům? Jak bude nastaven přístup v průběhu projektu, a jak po jeho skončení?
29. Jak bude ověřena totožnost osoby, která dostane k datům přístup?
30. Je potřeba zřídit komisi pro přístup k datům (např. pro hodnocení/schvalování žádostí o přístup k osobním/citlivým údajům)?
31. Budou metadata otevřeně dostupná a licencovaná pod veřejnou licenci CC-0, jak je uvedeno v grantové dohodě? Pokud ne, objasněte, prosím, proč. Budou metadata obsahovat informace, které uživateli umožní přístup k datům?
32. Jak dlouho budou data dostupná a dohledatelná? Bude zaručeno, že metadata zůstanou dostupná i poté, co data už nebudou dostupná?
33. Bude zahrnuta dokumentace nebo odkaz na jakýkoliv software, který bude potřebný pro přístup k datům nebo jejich prohlížení?
34. Bude možné zahrnout příslušný software (např. v otevřeném zdrojovém kódu)?
35. Jakými slovníky, standardy, formáty nebo metodikami dat a metadat se budete řídit, aby vaše data byla interoperabilní a umožňovala šíření a opakované použití dat v rámci oborů i mezi nimi? Budete dodržovat osvědčené postupy pro interoperabilitu, které jsou podporovány (vědeckou) komunitou? Které to budou?
36. V případě, že je nevyhnutelné používat neobvyklé ontologie nebo generovat ontologie a slovníky specifické pro projekt, poskytnete mapování běžněji používaným ontologiím? Budete otevřeně publikovat generované ontologie nebo slovníky, aby se dali znovu použít, upřesnit nebo rozšířit?
37. Budou vaše data obsahovat odkazy na jiná data (např. jiná data z vašeho projektu nebo soubory dat z předchozího výzkumu)?
38. Jak poskytnete dokumentaci potřebnou k ověření analýzy dat a usnadnění opětovného použití dat (např. soubory readme s informacemi o metodice, workflow, čištění dat, analýze dat, definicích proměnných, měrných jednotkách atd.)?
39. Budou vaše data volně dostupná, aby bylo umožněno jejich co nejširší opakované použití? Budou vaše data licencována pomocí standardních licencí pro opětovné použití v souladu s povinnostmi stanovenými v grantové dohodě?
40. Bude umožněno použití dat třetí straně během, a hlavně po skončení projektu?
41. Bude původ dat důkladně zdokumentován pomocí příslušných standardů?
42. Popište všechny relevantní procesy zajištění kvality dat.

43. Kromě zásad FAIR by se DMP měly rovněž zabývat jinými výstupy výzkumu než daty a měly by pečlivě zvážit aspekty související s přidělováním zdrojů, bezpečností dat a etickými aspekty.

Další výstupy:

44. Kromě správy dat by příjemci měli také zvážit a naplánovat správu dalších výstupů výzkumu, které mohou být generovány nebo znovu použity v rámci jejich projektů. Tyto výstupy mohou být digitální (např. software, pracovní toky, protokoly, modely atd.) Nebo fyzické (např. nové materiály, protilátky, činidla, vzorky atd.).
45. Příjemci by měli zvážit, které z otázek týkajících se výše uvedených FAIR standardů se mohou vztahovat na řízení dalších výstupů výzkumu a měly by se snažit poskytnout dostatečné podrobnosti o tom, jak budou jejich výstupy z výzkumu řízeny a sdíleny nebo zpřístupněny k opětovnému použití v souladu se zásadami FAIR.

Alokace zdrojů:

46. Jaké budou náklady na vytvoření FAIR dat nebo jiných výstupů výzkumu ve vašem projektu (např. přímé a nepřímé náklady související s úložištěm, archivací, opětovným použitím, zabezpečením atd.)?
47. Jak budou tyto náklady pokryty? Upozorňujeme, že náklady související se správou dat o výzkumu/výstupech jsou způsobilé jako součást grantu Horizon Europe (pokud jsou v souladu s podmínkami grantové dohody)
48. Kdo bude zodpovědný za správu dat ve vašem projektu?
49. Jak bude zajištěno dlouhodobé uchování? Diskutujte o nezbytných zdrojích, jak toho dosáhnout (náklady a potenciální hodnota, kdo a jak rozhoduje, jaká data budou uchovávána a jak dlouho)?

Bezpečnost dat:

50. Jaká opatření jsou nebo budou zavedena pro zabezpečení dat (včetně obnovy dat, zabezpečeného úložiště/archivace a přenosu citlivých dat)?
51. Budou data bezpečně uložena v důvěryhodných úložištích pro dlouhodobé uchování a udržování?

Etika:

52. Existují nebo mohly by existovat nějaké etické nebo právní překážky, které mohou mít dopad na sdílení dat? Lze o nich také diskutovat v kontextu etického přezkumu. Je -li to relevantní, zahrňte do popisu akce Descriptio of Action (DoA) odkazy na etické výstupy a kapitolu etiky.
53. Bude informovaný souhlas se sdílením dat a dlouhodobým uchováváním zahrnut do dotazníků zabývajících se osobními údaji?

Ostatní:

54. Využíváte, nebo budete využívat jiné národní/grantové/sectorové/institucionální postupy pro správu dat? Pokud ano, jaké (uveďte prosím jejich seznam a stručně je popište).

XX

Užitečné odkazy:

plán správy dat TAČR:

<https://www.tacr.cz/dokumenty/plan-spravy-dat-ta-cr>

plán správy dat Horizon:

https://www.tacr.cz/wp-content/uploads/documents/2022/11/23/1669191818_PI%C3%A1n%20spr%C3%A1vy%20dat%20-%20Horizont%20Evropa.docx

(trvalý link <https://doi.org/10.48813/sstg-4g21>)

repozitáře:

<https://www.re3data.org/>

ontologie z oboru medicíny:

<https://bioportal.bioontology.org/ontologies>

další intologie viz např.

https://en.wikipedia.org/wiki/Ontology_%28computer_science%29#Published_examples

<https://fairsharing.org/search?fairsharingRegistry=Standard>

metadata:

<https://www.dcc.ac.uk/guidance/standards/metadata/list>

webináře a školení:

<https://data.gov.cz/vzd%C4%9BI%C3%A1v%C3%A1n%C3%AD/>

další užitečné linky:

<https://www.wg-rdm.cz/domovska-stranka/uzitecne-zdroje>

<https://openscience.cuni.cz/>

Kontakty:

Bc. Ludvík Steiner (ludvik.steiner@tul.cz)

Mgr. Marta Zizienová (marta.zizienova@tul.cz)

Doporučené nástroje pro DMP:

- <https://dmponline.dcc.ac.uk/> (nejjednodušší, britský nástroj)
- <https://argos.openaire.eu/home> (Evropská unie, slovenská verze viz <https://argos.openaire.eu/about>)
- <https://ds-wizard.org/> (nejpropracovanější, FIT ČVUT a Ústav organické chemie a biochemie Akademie věd, kontakt <https://ds-wizard.org/contact>, obsahuje kalkulátor nákladů na otevřená data)

Poznámka k otevřeným datům pro rok 2023 (vhodné uvést vždy):

Oblast otevřené vědy se bouřlivě vyvíjí. Vzhledem k aktuálně probíhajícím a připravovaným aktivitám (např. plánované směrnici pro otevřená data na TUL, postupnému zavádění Národní datové infrastruktury) budeme tuto oblast pečlivě monitorovat a v případě kladného vyhodnocení využívat aktuální infrastrukturu (např. národní datový repozitář či přidělování perzistentních identifikátorů) dle aktuálních směrnic a doporučení.

Vzorové anonymizované texty bez konkrétních poskytovatelů:

Projekt 1:

V rámci projektu budeme vytvářet a zpracovávat data z měřících přístrojů (popis technického zabezpečení projektu viz. příloha č.1 projektu, odstavec 4.1). Půjde o nová data z TUL a z firmy XXX; formáty tabulkové (xls, txt, dat) i obrazové (jpg, png, bmp, tif) i textové (txt, doc); celkem cca XXX GB. Data nebudou obsahovat žádné citlivé (osobní) údaje, anonymizace dat nebude nutná.

Během projektu budou data ukládána na osobních počítačích žadatelů, dále na počítačích každého z přístrojů a na virtuálním cloudovém úložišti zřízeném v rámci infrastruktury CESNET, kam je omezený přístup (pouze pro členy virtuální organizace). Data budou dostupná jen řešitelům projektu. Metadata nicméně plánujeme zveřejňovat průběžně (po etapách) již v době trvání projektu s možností o data zažádat před jejich zveřejněním.

Analýzy dat a diskuse výsledků proběhnou v již koupeném komerčním software (Microsoft Office, Adobe software, Excel, TIBCO Statistica, Matlab, Grammarly aj.). Výsledky poslouží k publikování (jeden impaktovaný článek a dvou článků do Scopus) a pro ochranu duševního vlastnictví (patentování ve formě užitečného vzoru, dvou ověřených technologií a dvou funkčních vzorků).

Z důvodu právních a z důvodu ochrany know-how firmy plánujeme úplné zveřejnění dat až po ukončení projektu, půjde o původní naměřená data vč. bohatých metadat (údajů o přístrojích aj.).

Použijeme Národní repozitář pro data, který by měl být v čase trvání projektu již plně funkční. Dodržíme metadatový standard DataCite Metadata Schema v4, interoperabilitu (terminologický slovník meSH, formáty dat, identifikátory autorů a projektů, linkování na publikace a další výstupy vč. systémových požadavků a vztahů k dalším výsledkům). Trvat budeme na přidělení identifikátoru DOI a především na dlouhodobé archivaci.

Licenci primárních dat předpokládáme CC-BY ve verzi 4.0 nebo CC0 všude tam, kde je to účelné, opakované využití zveřejněných dat tedy bude bez nutnosti individuální autorizace přístupu.

Skripty pro analýzy zveřejňovat neplánujeme, výsledky analýz budou publikovány.

Pro správnou správu dat zřídíme pozici data stewarda, který bude odpovídat za dodržování běžné mezinárodní praxe v oblasti datového stewardshipu.

Náklady na správu dat budou odpovídat výši nákladů na pozici Data Steward. Cloudové úložiště máme díky členství v infrastruktuře CESNET zdarma.

Zveřejněná data budou nadále užitečná pro statistické vyhodnocení a optimalizaci firemních/vědeckých postupů a technologií v rámci týmu i u dalších výzkumných institucí a firem z oboru.

Oblast otevřené vědy se bouřlivě vyvíjí. Vzhledem k aktuálně probíhajícím a připravovaným aktivitám (např. plánované směrnici pro otevřená data na TUL, postupnému zavádění Národní datové infrastruktury) budeme tuto oblast pečlivě monitorovat a v případě kladného vyhodnocení využívat aktuální infrastrukturu (např. národní datový repozitář či přidělování perzistentních identifikátorů) dle aktuálních směrnic a doporučení.

Projekt 2:

V rámci projektu budeme vytvářet a zpracovávat data z měřících přístrojů (popis technického zabezpečení projektu viz. příloha č.1 projektu, odstavec 4.1). Půjde o nová data z TUL a z firmy XXX; formáty tabulkové (spa, jdf, csv, xls, txt, dat) i obrazové (jpg, png, bmp, tif) i textové (txt, doc); celkem cca do XXX GB. Data nebudou obsahovat žádné citlivé (osobní) údaje, anonymizace dat nebude nutná.

Během projektu budou data ukládána na osobních počítačích žadatelů, dále na počítačích každého z přístrojů a na virtuálním cloudovém úložišti zřízeném v rámci infrastruktury CESNET, kam je omezený přístup (pouze pro členy virtuální organizace). Data budou dostupná jen řešitelům projektu. Metadata nicméně plánujeme zveřejňovat průběžně (po etapách) již v době trvání projektu s možností o data zažádat před jejich zveřejněním.

Analýzy dat a diskuse výsledků proběhnou v již koupeném komerčním software (Microsoft Office, Adobe software, MNova, Omnic, FluorCam 7, TA universal analysis 2000) nebo freeware (Past 2.17c). Export dat do otevřených formátů je možný, nikoli však bezetrátově. Výsledky poslouží pro přípravu ověřené technologie.

Z důvodu právních a z důvodu ochrany know-how firmy plánujeme úplné zveřejnění dat až po ukončení projektu, půjde o původní naměřená data vč. bohatých metadat (údajů o přístrojích aj.).

Použijeme Národní repozitář pro data, který by měl být v čase trvání projektu již plně funkční. Dodržíme metadatový standard DataCite Metadata Schema v4, interoperabilitu (formáty dat, identifikátory autorů a projektů, linkování na publikace a další výstupy vč. systémových požadavků a vztahů k dalším výsledkům). Trvat budeme na přidělení identifikátoru DOI a především na dlouhodobé archivaci.

Licenci primárních dat předpokládáme CC-BY ve verzi 4.0 nebo CC0 všude tam, kde je to účelné, opakované využití zveřejněných dat tedy bude bez nutnosti individuální autorizace přístupu.

Skripty pro analýzy zveřejňovat neplánujeme, výsledky analýz budou publikovány.

Pro správnou správu dat zřídíme pozici data stewarda, který bude odpovídat za dodržování běžné mezinárodní praxe v oblasti datového stewardshipu.

Náklady na správu dat budou odpovídat výši nákladů na pozici Data Steward. Cloudové úložiště máme díky členství v infrastruktuře CESNET zdarma.

Zveřejněná data budou nadále užitečná pro statistické vyhodnocení a optimalizaci firemních/vědeckých postupů a technologií v rámci týmu i u dalších výzkumných institucí a firem z oboru.

Oblast otevřené vědy se bouřlivě vyvíjí. Vzhledem k aktuálně probíhajícím a připravovaným aktivitám (např. plánované směrnice pro otevřená data na TUL, postupnému zavádění Národní datové infrastruktury) budeme tuto oblast pečlivě monitorovat a v případě kladného vyhodnocení využívat aktuální infrastrukturu (např. národní datový repozitář či přidělování perzistentních identifikátorů) dle aktuálních směrnic a doporučení.

Projekt 3:

V rámci projektu budeme vytvářet a zpracovávat data z měřících přístrojů. Půjde o nová data z TUL a z firmy XXX; formáty tabulkové (csv, xls), obrazové (jpg, png, tif) i textové (txt, doc); celkem cca do XXX GB. Data nebudou obsahovat žádné citlivé osobní údaje, anonymizace nebude nutná.

Na TUL budou v průběhu projektu data ukládána na osobních počítačích žadatelů, dále na počítačích přístrojů a na virtuálním cloudovém úložišti zřízeném v rámci infrastruktury CESNET, s omezeným přístupem pouze pro členy virtuální organizace. Firma XXX bude ukládat data do vlastního sdíleného replikovaného úložiště NAS s RAIDovým polem zabezpečeného UPS s přepěťovou ochranou a zabezpečeným přístupem. Data budou dostupná jen řešitelům projektu.

Metadata plánujeme zveřejňovat průběžně (po etapách) již v době trvání projektu s možností o data zažádat před jejich zveřejněním.

Analýzy dat a diskuse výsledků proběhnou v již koupeném komerčním software (Microsoft Office, GraphPad). Export dat do otevřených formátů je možný, nikoli však bezztrátově. Výsledky poslouží pro optimalizaci poloprovozní technologie.

Z důvodu právních a z důvodu ochrany know-how firmy plánujeme úplné zveřejnění dat až po ukončení projektu, půjde o původní naměřená data vč. bohatých metadat (údajů o přístrojích aj.). Vybraná data firmy XXX budou vedena v režimu utajení z důvodu ochrany citlivého know-how.

Použijeme Národní repozitář pro data, který by měl být v čase trvání projektu již plně funkční. Dodržíme metadatový standard DataCite Metadata Schema v4, interoperabilitu (terminologický slovník meSH, formáty dat, identifikátory autorů a projektů, linkování na publikace a další výstupy vč. systémových požadavků a vztahů k dalším výsledkům). Trvat budeme na přidělení identifikátoru DOI a především na dlouhodobé archivaci.

Licenci primárních dat vzniklých na TUL předpokládáme CC-BY ve verzi 4.0 nebo CC0 všude tam, kde je to účelné, opakované využití zveřejněných dat tedy bude bez nutnosti individuální autorizace přístupu.

Skripty pro analýzy zveřejňovat neplánujeme, výsledky analýz budou publikovány.

V rámci řešitelského týmu bude jmenován data steward, který bude odpovídat za dodržování běžné mezinárodní praxe v oblasti datového stewardshipu.

Náklady na správu dat na straně TUL budou nulové. Cloudové úložiště máme díky členství v infrastruktuře CESNET zdarma. Náklady na správu dat firmy XXX budou.....

Zveřejněná data budou nadále užitečná pro statistické vyhodnocení a optimalizaci firemních/vědeckých postupů a technologií v rámci týmu i u dalších výzkumných institucí a firem z oboru.

Oblast otevřené vědy se bouřlivě vyvíjí. Vzhledem k aktuálně probíhajícím a připravovaným aktivitám (např. plánované směrnici pro otevřená data na TUL, postupnému zavádění Národní datové infrastruktury) budeme tuto oblast pečlivě monitorovat a v případě kladného vyhodnocení využívat aktuální infrastrukturu (např. národní datový repozitář či přidělování perzistentních identifikátorů) dle aktuálních směrnic a doporučení.

Pár zahraničních příkladů

https://dmptool.org/plans/63100/export.pdf?export%5Bquestion_headings%5D=true

https://dmptool.org/plans/83986/export.pdf?export%5Bquestion_headings%5D=true

<https://www.unige.ch/researchdata/fr/planifier/dmp-fns/>