

## Plán správy dat projektu FerrMion

### Účel a principy plánu správy dat

Plán správy dat projektu FerrMion (dále “plán”) slouží jako vodítko pro nakládání s výzkumnými datovými sadami generovanými v rámci plnění výzkumných úkolů projektu FerrMion v průběhu jejich celého životního cyklu (dále jen “data”). Cílem správy dat je evidovat výzkumná data, chránit je před ztrátou nebo poškozením a umožnit jejich sdílení. Správa dat se řídí pravidly [FAIR](#), principem „otevřená jak jen možno, uzavřená jen jak nutno“, respektuje institucionální postupy správy dat partnerů projektu a omezení vyplývající z oprávněných zájmů a omezení specifikovaných v Odstavci 6.

### Typ spravovaných dat a účel jejich sběru

V průběhu projektu jsou generována a sbírána výzkumná data především tří typů: i) experimentální data, ii) fotografie a snímky (případně videa) ze zobrazovacích přístrojů, iii) data generovaná v numerických simulacích. Data prvních dvou typů vznikají především jako výsledek charakterizace anorganických vzorků pevné a kapalné fáze za různých fyzikálních a chemických podmínek a různými metodami základního výzkumu. Data třetího typu jsou především výstupem numerických simulací zkoumaných systémů na různých časových a prostorových škálách. V projektu se nepředpokládá sběr vědeckých dat týkajících se osob nebo lidských či zvířecích vzorků; žádná osobní nebo citlivá data tedy nejsou sbírána ani shromažďována. Typ a formát syrových dat je závislý na přístrojovém a softwarovém vybavení zapojených výzkumných pracovišť a využitých výzkumných infrastrukturách (.dat, .bin, .mat, .odb, .jpg, apod.). Celkovou průměrnou velikost očištěných dat generovaných v projektu lze odhadnout v řádu desítek až stovek TB za rok.

Data jsou průběžně využívána k ověřování/vyvracení vědeckých hypotéz o fyzikálních a chemických procesech probíhajících ve zkoumaných materiálech a o jejich vnitřní struktuře, validaci teoretického popisu materiálů, materiálových systémů a jejich interakcí, a k dalším obvyklým činnostem související s prováděným výzkumem.

### Sběr, ukládání a dokumentace dat

Výzkumná data jsou sbírána a ukládána průběžně; kvalita dat je zajištěna oborově standardními mechanismy (např. opakovaná měření, srovnání s již publikovanými daty, atd.). K ukládání a zálohování dat je využívána stávající, potřebám projektu postačující infrastruktura zapojených pracovišť. K dodatečnému zálohování dat mohou být po konzultaci s data stewardem využita důvěryhodná a zabezpečená externí datová úložiště (např. cloud národní e-infrastruktury pro vědu a výzkum [CESNET](#)), a to pokud bude zajištěna patřičná ochrana a budou jasně vymezena přístupová práva. Instituce žadatele a partnerů jsou zodpovědné za poskytování technických nástrojů a finančních zdrojů pro ukládání, zálohování a zabezpečení dat vzniklých činnostmi výzkumníků ze svých pracovišť po dobu trvání projektu i po jeho ukončení. Vědečtí pracovníci jsou odpovědní za vytvoření a uchování dokumentace k získaným experimentálním datům, např. prostřednictvím laboratorních deníků; elektronická forma laboratorních deníků bude preferována s ohledem na možnosti jejich budoucího zpřístupnění v rámci postupů otevřené vědy.

Vědečtí pracovníci jsou odpovědní za vytvoření a uchování dokumentace k získaným experimentálním datům, např. prostřednictvím laboratorních deníků; elektronická forma laboratorních deníků bude preferována s ohledem na možnosti jejich budoucího zpřístupnění v rámci postupů otevřené vědy.

### Přístup k datům, zpracování dat, odstranění nepotřebných dat

Pravidla přístupu k datům na jednotlivých pracovištích partnerů projektu se řídí jejich vnitřními pravidly a normami. Data jsou průběžně čištěna a zpracovávána do různých forem (tabulky, obrázky, grafy,

schémata, apod.). V tomto procesu je dodržován etický kodex příslušné výzkumné instituce. Vědeckí pracovníci si povedou záznam o postupech a nástrojích (např. software) využitých při čištění a zpracování dat. Pokud se po zpracování dat ukáže, že jejich některá část je nepotřebná (např. technická chyba na měřicím přístroji, chybná specifikace vstupních dat simulace), může vědecký pracovník navrhnout odstranění takovýchto dat z datových úložišť. O tomto návrhu rozhoduje vedoucí výzkumného záměru (VZ), případně jím pověřený člen týmu VZ.

### **Harmonogram otevření dat**

O dokončení fáze zpracování a interpretace dat informuje odpovědný výzkumný pracovník vedoucího výzkumné aktivity a další členy týmu podílející se na výzkumné aktivitě. Na společné schůzi je sjednán předběžný způsob zveřejňování výsledků (především formou vědeckých publikací) a zpřístupňování příslušných výzkumných dat (především formou uložení do datového repozitáře) a připraven harmonogram tohoto zpřístupňování. O tomto harmonogramu bude informován vedoucí VZ, nejpozději na následujícím pravidelném jednání týmu VZ, a data stevard. Průběžnou kontrolu a případnou aktualizaci a úpravu harmonogramu provádí vedoucí výzkumné aktivity na návrh odpovědného výzkumného pracovníka; o úpravě neprodleně informuje vedoucího VZ, v případě nutnosti data stevarda.

V případě dat získaných na funkčních vzorcích/prototypch nebo jiných entitách vzniklých v rámci projektu, u kterých je zvažována možnost patentové ochrany, nebo v jiných případech, kdy vzniká legitimní předpoklad, že datová sada nebo její část vyžaduje utajení nebo uzavření (viz Odstavec 6), je procedura zpřístupňování výsledků pozastavena. O dalším postupu zpřístupňování dat bude rozhodnuto nejpozději na následujícím pravidelném jednání týmu VZ, na které bude přizván odborný manažer projektu, viz Odstavec 6.

### **Otevřený přístup k datům a jeho odůvodněné omezení**

Na základě harmonogramu zpřístupňování jsou v součinnosti s data stevardem data uložena v důvěryhodném datovém repozitáři, který umožňuje uveřejnění v licenci CC BY nebo jí ekvivalentních. Žadatel projektu se zapojil do přípravy [projektu pro vybudování, implementaci a správu repozitáře](#) pro oborově-vědní klastr *Materiálové vědy a technologie* v ČR a standardizaci metadatových struktur a jiných datových nástrojů pro obory sdružené v klastru, viz též [e-INFRA.cz](#). Do okamžiku spuštění národního oborového repozitáře jsou využívány především oborové datové repozitáře [Materialscloud](#), [MDR](#), institucionální repozitář [ASEP](#) spravovaný AVČR, případně univerzální repozitář [Zenodo](#), nebo – v případě nutnosti a po konzultaci s data stevardem – jiné repozitáře splňující kritéria kvality dle [Příručky postupů otevřené vědy OP JAK, s. 6](#). Záznam je provázán s co nejpodrobnějším metadatovým popisem odpovídajícím oborovým standardům a je ověřeno přiřazení persistentního identifikátoru. V rámci snahy o kvalitní a bohatý metadatový popis zajistí odpovědný člen týmu prostřednictvím těchto repozitářů maximální možné propojení mezi metadaty datové sady a dalšími souvisejícími publikacemi a výstupy (např. publikovaný článek, zpráva z výzkumu, algoritmus pro zpracování dat, popis postupů zpracování dat, apod.). V případě dat vzniklých na výzkumné infrastruktuře spravované mimo partnerská pracoviště projektu (ESRF, ESS, apod.) zohledňují postupy zpřístupnění dat pravidla nastavená provozovatelem této infrastruktury.

V odůvodněných případech lze omezit přístup k některým nebo všem výzkumným datům. Takovými důvody jsou zejména: právo na ochranu soukromí, ochrana osobních údajů, důvěrnost údajů a dat, oprávněné obchodní zájmy, obchodní tajemství, a práva duševního vlastnictví třetích stran, bezpečnost státu, rozpor s oprávněnými zájmy příjemce, včetně komerčního využití dat; nebo jiné oprávněné zájmy a jiná oprávněná omezení daná mj. interními předpisy žadatele a partnerů. O takovém omezení přístupu k datům rozhoduje vedoucí VZ po konzultaci s odborným manažerem projektu a data stevardem. O tomto rozhodnutí je informován Řídící výbor projektu nejpozději na svém nejbližším zasedání a jeho platnost je omezena na dobu do dalšího zasedání Řídícího výboru (zpravidla jeden rok). Na každém dalším zasedání zhodnotí vedoucí VZ aktuálnost důvodů pro omezení přístupu. Pokud pominou důvody k omezení přístupu k datům, vedoucí VZ po konzultaci s odborným manažerem projektu a data stevardem přistoupí k sestavení nového harmonogramu otevření dat dle Odstavce 5.

### **Data stevard**

V projektu bude ustanovena funkce data stevarda. Data stevard zodpovídá zejména za:

- Nastavení workflow managementu dat z různých zdrojů a dohled nad celým životním cyklem dat, identifikace případných problémů a jejich řešení.
- Nastavení a zajištění metadatového popisu, uchování, zálohování a zabezpečení dat, s ohledem na jejich citlivost a objem.
- Podpora výzkumníků a jejich práce s daty, dohled nad dodržováním etických a právních požadavků při ukládání a sdílení dat, pomoc při výběru licence a stanovení přístupnosti dat.
- Dodržení standardů a nejlepší praxe podle zvyklostí daného oboru, sledování vývoje v oblasti otevřených dat a navrhování případných úprav metodik a postupů v rámci projektu.

Data stevard také spolupracuje s odděleními informačních technologií žadatele a partnerů při výběru a nákupu softwarového a hardwarového vybavení sloužícího pro správu dat v rámci projektu.

### **Aktualizace plánu v průběhu projektu a jeho zveřejnění**

Tento plán bude v průběhu projektu pravidelně revidován (alespoň po ukončení prvního a třetího roku projektu) s ohledem na vývoj potřeb a možností žadatele a partnerů projektu za současného respektování pokynů uvedených v [Příručce postupů otevřené vědy OP JAK](#) a v [Šabloně plánu pro správu dat](#). Za úpravy plánu odpovídá data stevard, aktualizace plánu podléhá schválení odborného manažera projektu. S aktualizovaným plánem jsou bezodkladně seznámeni všichni členové týmu projektu. Plán může být v souladu s doporučením [Příručky postupů otevřené vědy OP JAK, s. 16](#) zveřejněn prostřednictvím uložení do specializovaného repozitáře (např. [DMPonline](#)).

### **Závěrečná ustanovení**

Tento plán byl vytvořen při podání žádosti o podporu projektu FerrMion, byl schválen navrhovatelem projektu a je přiložen jako povinná součást této žádosti. V případě schválení podpory projektu vstupuje tento plán v platnost prvním dnem trvání projektu.