

Analýza a návrh postupů k naplňování FAIR principů v rámci NRP/NDI

Základní informace

Název dokumentu	Analýza a návrh postupů k naplňování FAIR principů v rámci NRP/NDI
Registrační číslo projektu	CZ.02.01.01/00/23_014/0008787
Název KA	5.1
Číslo a název výstupu	D5.1.2
Zodpovědná osoba	Ilona Trtíková
Perzistentní identifikátor	
Licence dokumentu	CC BY 4.0 https://creativecommons.org/licenses/by/4.0/

Historie verzí

Verze	Datum	Popis
1.2	15. 02. 2026	Zpřehlednění textu podle připomínek, ke kapitolám doplněny souhrny a upřesněné vysvětlující texty v tabulce s FAIR principy. Upravena a zpřesněna terminologie (glosář). Aktualizovány informace o službách a nástrojích, které jsou dostupné v různých úrovních provozu (produkční či pilotní) k datu vydání verze 1.12
1.1	24. 11. 2025	Zpracovány připomínky k zveřejněné zprávě a aktualizovány informace o službách a nástrojích, které jsou dostupné v různých úrovních provozu (produkční či pilotní) k datu vydání verze 1.1
1.0	31. 03. 2025	První veřejná verze zprávy.



Autoři

Instituce	Jméno a příjmení	ORCID	Email
CESNET	Ilona Trtíková	0000-0001-6491-2947	Ilona.trtikova@cesnet.cz
KNAV	Petra Bártková	0000-0001-6484-4710	bartkova@knave.cz
BTÚ AV ČR	Lada Biedermannová	0000-0001-6504-6403	lada.biedermannova@ibt.cas.cz
KNAV	Jindřich Fejfar	0000-0002-2048-3682	fejfar@knave.cz
ČVUT	Robert Pergl	0000-0003-2980-4400	perglr@fit.cvut.cz
KNAV/ARÚ	David Novák	0000-0001-6592-6291	novak@arup.cas.cz
KNAV	Tomáš Čížek	0000-0001-7734-2947	tomas.cizek@soc.cas.cz

Seznam použitých zkratk

Zkratka	Popis
ARK	Archival Resource Key – identifikátor využívaný v digitálních archivech
CARDS	Czech Academic and Research Discovery Services
CAT	Compliance Assessment Toolkit
CCMM	Czech Core Metadata Model
CC REL	Creative Commons Rights Expression Language
CDIF	Cross-Domain Interoperability Framework
CODATA	Committee on Data of the International Science Council
CTS	CoreTrustSeal – mezinárodní organizace udělující certifikaci repozitářům
DANS	Data Archiving and Network Services – Nizozemské národní odborné centrum a úložiště výzkumných dat
DCAT	Data Catalog Vocabulary – základní datový model evropského standardu pro katalogizaci dat
DCAT-AP	Data Catalog Vocabulary Application Profile
DDI	Data Documentation Initiative Alliance – metadatový standard pro sociální vědy
DDI-CDI	Data Documentation Initiative Alliance – Cross-Domain Integration
DMP	Data Management Plan – plán správy dat
DOI	Digital Object Identifier – perzistentní identifikátor pro identifikaci digitálních objektů
DSW	Data Stewardship Wizard
FAIR	Findable, Accessible, Interoperable, Reusable
FDMM	FAIR Data Maturity Model
FER	FAIR-Enabling Resource
FIP	FAIR Implementation Profile
FORCE11	Zásady FAIR výzkumných dat, definované komunitou na workshopu "Future of Research Communication and e-Scholarship" v roce 2011
F-UJI	Automated FAIR Data Assessment Tool
GO	Gene Ontology – databáze znalostí genové ontologie
GO FAIR	Mezinárodní iniciativa zaměřená na podporu principů FAIR
HDF5	Hierarchical Data Format version 5
JSON-LD	JavaScript Object Notation for Linked Data – formát pro reprezentaci sémantických dat, kompatibilní s webovými technologiemi
LDAP	Lightweight Directory Access Protocol
MeSH	Medical Subject Headings
MIAME	Minimum Information About a Microarray Experiment
NDI	Národní datová infrastruktura
NMA	Národní metadatový adresář
NRP	Národní repozitářová platforma pro výzkumná data
NTK	Národní technická knihovna
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OAuth	Web Authorization Protocol – Open standard for access delegation
ODRL	Open Digital Rights Language Ontology
ORCID	Open Researcher and Contributor ID
OSI	Výzva Open Science I
OSII	Výzva Open Science II

OWL	Web Ontology Language – jazyk pro tvorbu ontologií, používaný v sémantickém webu
PID	Persistent identifier
PROV-O	PROV Ontology
PURL	Persistent Uniform Resource Locator – přesměrovací služba umožňující uchování stabilních odkazů na digitální objekty
RDA	Research Data Alliance
RDF	Resource Description Framework – standardní model pro reprezentaci propojených dat
REST API	Representational State Transfer – Application Programming Interface
ROR	Research Organization Registry – perzistentní identifikátor výzkumných organizací
SAML	Security Assertion Markup Language
SFTP	SSH File Transfer Protocol
SKOS	Simple Knowledge Organization System – formát pro reprezentaci, sdílení a publikování heslářů, taxonomií, klasifikačních schémat, tezaurů v rámci sémantického webu vycházející ze standardů RDF a RDFS
SPARQL	Simple Protocol and RDF Query Language
SPDX	Software Package Data Exchange
XKOS	Extended Knowledge Organization System – rozšíření standardu SKOS o funkcionality potřebné pro statistická a ekonomická data
XML	Extensible Markup Language

Terminologie (glosář)

Základní objekty

- **Datová sada (dataset)** – Konceptuální entita představující publikovaná data.
- **Distribuce (distribution)** – Konkrétní reprezentace datové sady. Datová sada může být dostupná v různých serializacích, které se mohou navzájem lišit různými způsoby, mimo jiné při-rozeným jazykem, media-typem či formátem, schematickou organizací, časovým a prostorovým rozlišením, úrovní detailu či profily (které mohou specifikovat některé či všechny tyto rozdíly).
- **Metadatový záznam** – Strukturované údaje („data o datech“) popisující digitální objekt za úče-lem jeho identifikace, vyhledatelnosti a správy. Zahrnují např. informace o autorství, původu, datu a typu objektu. Mohou být popisná, technická, strukturální či administrativní a řídí se stan-dardy jako Dublin Core, DataCite aj.
- **Landing page** – Strojově i člověkem čitelná specializovaná webová stránka, která poskytuje informace o konkrétním výzkumném objektu (např. datové sadě) a přístup k němu; obsahuje podrobnosti o daném objektu, včetně metadat, přístupových odkazů a informací o omezeních jeho užívání.
- **Repozitář** – organizace, která uchovává, spravuje a zpřístupňuje digitální materiály různých formátů a typů. Materiály v online repozitářích jsou spravovány tak, aby umožňovaly jejich vy-hledávání, objevování a opětovné použití. Musí být zajištěna odpovídající kontrolní opatření, aby digitální materiály zůstaly dlouhodobě autentické, spolehlivé, přístupné a použitelné.

Integrace a publikace

- **Harvesting** (sklizeň metadat) – automatizované získávání metadat z repozitářů do centrální služby (typicky přes OAI-PMH). Cílem harvestování je obvykle vytvořit centrální index metada-tových záznamů a umožnit tak vyhledávání z jednoho vyhledavače nad více zdroji, popřípadě vytvoření velké tematické kolekce obsahu z různých zdrojů.
- **Signposting** – použití HTTP Link hlaviček pro strojově čitelnou navigaci mezi landing page, metadaty a distribucemi.
- **Tombstone page** – cílová stránka (landing page) pro staženou/odstraněnou datovou sadu; obsahuje minimálně metadata a informaci o důvodu či náhradě.

Identifikátory a autority

- **PID (trvalý identifikátor)** – Trvalý digitální odkaz na objekt, který poskytuje informace o tomto objektu bez ohledu na to, co se s objektem stane.
- **DOI** – nejběžnější PID pro výzkumná data a publikace.
- **ORCID / ROR** – mezinárodní identifikátory pro osoby (ORCID) a organizace (ROR), které za-jišťují jednoznačné přiřazení autorů a institucí.

FAIR a oborové profily

- **FAIR principy** – soubor zásad, aby data šla najít, zpřístupnit, propojit a znovu **použít**.
- **FAIRness** – míra souladu digitálního objektu (datové sady, metadat) s FAIR principy.
- **Metadatový profil** (Metadata Profile) – konkrétní sada povinných/volitelných prvků a validač-ních pravidel pro metadata (včetně slovníků a kontrolovaných hodnot) používaná v repozitáři či komunitě.
- **FER (FAIR Enabling Resource)** – konkrétní služba, specifikace nebo datová politika, která je slouží pro zavedení principů FAIR do praxe, tj. pro jejich uplatnění. FER poskytuje funkci po-třebnou k dosažení určitého aspektu chování FAIR a je výslovně spojena s jedním nebo více principy FAIR. Příkladem může být “DataCite DOI” jakožto FER pro zajištění F1.

- **FIP (FAIR Implementation Profile)** – seznam deklarovaných implementačních rozhodnutí (FER), jejichž cílem je naplňovat jednotlivé zásady FAIR a které byly přijaty jako společné rozhodnutí oborové komunity. Soubor otázek ke zjištění konkrétních zdrojů (slovníků, identifikátorů, standardů, technologií, schémat atd.), použitých k naplnění jednotlivých zásad FAIR.
- **CCMM (Czech Core Metadata Model)** – základní metadatový model pro popis výzkumných dat v ČR, vytvořený v rámci projektu CARDS. Vychází zejména ze standardů DCAT-AP a DataCite, částečně také z DCAT-AP-CZ a Dublin Core.

Práva, přístup a verze

- **Licence** – podmínky, za kterých smí být data používána. Vždy musí být uvedena odkazem na přesné znění.
- **Podmínky přístupu** – zda jsou data otevřená, v embargu nebo na vyžádání, a jak o přístup požádat.
- **Verze** – označení vydání datové sady. Nová hlavní verze má mít jasnou vazbu na předchozí (ideálně i nový PID); má být srozumitelně popsáno, co se změnilo.

Role a odpovědnosti

- **Správce repozitáře** – zodpovídá za technickou konfiguraci, provoz a integraci služeb.
- **Datový kurátor** – dohlíží na kvalitu metadat, soulad s datovou politikou repozitáře a pomáhá sjednocovat terminologii. stanovuje pravidla pro ukládání a uchování dat v repozitáři a rozhoduje o konkrétních datových sadách. Zajišťuje harmonizaci metadat podle oborových profilů a jejich interoperabilitu s dalšími systémy (např. NMA) a podílí se na stanovení oborového metadatových profilu.
- **Data steward** – pracuje na straně výzkumného týmu nebo komunity, pomáhá s tvorbou DMP, výběrem standardů a přípravou dat pro uložení.
- **Datová politika repozitáře (Data Policy)** – soubor veřejně dostupných pravidel, která vymezují podmínky přijímání, správy, zpřístupňování a dlouhodobého uchování dat v repozitáři. Upravuje zejména odpovědnosti vkladatelů a repozitáře, pravidla verzování a změn, licenční podmínky, úroveň přístupu, dobu uchování a přidělování perzistentních identifikátorů (PID).

Obsah

1	Úvod.....	10
1.1	Mezinárodní standardy.....	10
2	Architektura implementace FAIR v NRP/NDI.....	12
2.1	Technická infrastruktura pro FAIRifikaci.....	12
2.2	Prostředky podpory FAIR datové správy v NRP/NDI.....	13
2.2.1	Integrace metadatových standardů	13
2.2.2	Podpora správy licencí	14
2.2.3	Řízení přístupu k datům a podpora pro citlivá data.....	14
2.2.4	Automatizace tvorby a sběru metadat	14
2.3	Znázornění implementace FAIR v prostředí NRP/NDI.....	14
2.3.1	Přehledové konceptuální schéma implementace FAIR v NRP/NDI	15
2.3.2	Ekosystém repozitáře: aktéři, pravidla a procesy.....	16
3	FAIR Principy a jejich výklad v kontextu NRP/NDI.....	18
3.1	Implementace FAIR principů v NRP/NDI.....	19
4	Služby podporující implementaci FAIR v rámci oborů.....	45
4.1	FAIR Implementation Profiles.....	45
4.2	FIP Wizard	45
4.3	Integrované řešení pro plánování správy dat	46
5	Nástroje a modely pro posouzení naplnění FAIR principů	48
5.1	Sebehodnotící modely a nástroje.....	48
5.2	Automatizované hodnocení FAIR.....	49
5.2.1	Klíčová role repozitáře v automatizovaném hodnocení.....	49
5.2.2	Příklady nástrojů pro automatické hodnocení.....	49
6	Závěr.....	52
	Příloha 1 – hodnocení vybraných datasetů v repozitářích pomocí nástroje F-UJI	56
	Postup hodnocení	56
	Repozitář ASEP.....	56
	Příklad hodnoceného záznamu	56
	Výsledek hodnocení	57
	Komentované výsledky hodnocení	57
	LINDAT/CLARIAH-CZ.....	58
	Příklad hodnoceného záznamu	58
	Výsledek hodnocení	58
	Komentované výsledky hodnocení	58

Český sociálněvědní datový archiv (ČSDA)	58
Příklad hodnoceného záznamu	58
Výsledek hodnocení	58
Komentované výsledky hodnocení	59
Příloha č. 2 – popis uzlů schémat	60

Shrnutí

Tento dokument popisuje, jak v prostředí NRP/NDI praktick¹ – jaké minimální požadavky musí splnit repozitáře a centrální služby, aby byla data dohledatelná, přístupná, interoperabilní a znovupoužitelná. Stěžejní pasáží je tabulka FAIR požadavků v kap. 3, která definuje požadavky na NRP, repozitáře zahrnuté do NDI a (tam, kde je relevantní) související³.

Minimum pro zapojení repozitáře do NDI (pro základní plnění FAIR):

- PID datové sady (preferovaně DOI) a landing page jako cílový přístupový bod.
- Základní metadata v CCMM a jejich zpřístupnění pro centrální sklizeň (NMA).
- Odkazy na distribuce (stabilní URL/endpointy) a jasně uvedená licence (minimálně jako URI).

Doporučená nadstavba (pro oborovou hloubku a pokročilé integrace v souladu s FAIR):

- Bohatá oborová metadata v repozitáři + strojově čitelná reprezentace na landing page (např. JSON-LD) a/nebo API.
- Signposting (HTTP Link v hlavičce) tam, kde to platforma umožňuje.
- Oborové profily/FIP pro standardní definici oborových schémat, slovníků a validačních pravidel a dalších požadavků.

Jak se orientovat v dokumentu:

- Kap. 2 – architektura a služby NRP/NDI.
- Kap. 3 – tabulka požadavků a doporučených postupů pro FAIR (NRP/repozitáře/FIP).
- Kap. 4–5 – podpora oborů (FIP) a nástroje pro posouzení FAIRness.

¹ <http://www.go-fair.org/fair-principles/>

² WILKINSON, Mark D.; DUMONTIER, Michel; AALBERSBERG, IJsbrand Jan, APPLETON, Gabrielle; AXTON, Myles et al. *The FAIR Guiding Principles for scientific data management and stewardship*. Online. Scientific Data, vol. 3 (2016), article no. 160018. Dostupné z: <https://doi.org/10.1038/sdata.2016.18>.

³ <https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/>

1 Úvod

V rámci implementace iniciativy EOSC⁴ v ČR je budována společná platforma pro sdílení, správu a přístup k datům a výpočetním zdrojům pro výzkumné účely, tzv. **Národní datová infrastruktura (NDI)**. Její součástí je **Národní repozitářová platforma pro výzkumná data (NRP)** vznikající v rámci výzvy **Open science I (OSI)**⁵.

Platforma má zajistit dlouhodobé uchování výzkumných dat a jejich zpřístupnění tak, aby byla v souladu s principy FAIR⁶ – tedy vyhledatelná, přístupná, interoperabilní a opakovaně použitelná. Vzniká proto řada služeb a systémů, které umožní bezpečně ukládat, spravovat a sdílet výzkumná data s plnou kontrolou nad jejich zpřístupněním, a to v souladu s mezinárodně uznávanou dobrou praxí. Z těchto důvodů je klíčové se v rámci přípravy a následného využití platformy NRP zabývat plněním a podporou FAIR principů, které byly vědeckou komunitou definovány právě pro zlepšení sdílení a opakovaného využívání dat.

Za kvalitu dat nesou odpovědnost především jejich tvůrci, avšak klíčovou podmínkou pro naplňování FAIR principů je vhodně nastavená infrastruktura – zejména pravidla, procesy a služby repozitářů. Odpovědnost za FAIR data je tak sdílená. Tato analýza se proto zaměřuje na všechny klíčové aktéry – správcům repozitářů a kurátorům poskytuje vodítka pro nastavení technických a organizačních opatření; vědcům a data stewardům objasňuje, jak efektivně využívat služby NRP/NDI pro správu a sdílení výzkumných dat v souladu s dobrou praxí.

V rámci NRP i NDI jsou nebo budou nabízeny služby, které podporují naplňování FAIR principů. Tento text mapuje nároky FAIR principů a zasazuje je do kontextu prostředí NRP/NDI. Stěžejní částí textu je výklad FAIR principů nejen jako mezinárodního standardu, ale především ve vztahu k možnostem, které nabízí nebo bude nabízet NRP/NDI. Zároveň jsou zde uvedeny požadavky na NRP, které ze standardů vyplývají.

Dokument rovněž představuje plánované aktivity v rámci NRP, které jsou nyní ve fázi implementace a jejichž cílem je systematicky podporovat uživatele NRP/NDI při naplňování požadavků na FAIR data. Důležitým prvkem této implementace je architektura NRP/NDI, která poskytuje technické a organizační prostředí pro ukládání a správu FAIR dat. Schémata znázorňující klíčové vztahy v rámci této infrastruktury napomáhají porozumění tomu, jak jsou FAIR principy v prostředí NRP/NDI realizovány a jak mohou jednotliví aktéři efektivně spolupracovat na jejich naplnění. Analýza pomáhá všem zainteresovaným stranám orientovat se v problematice FAIR dat a popisuje nástroje a metodiky pro jejich efektivní správu a hodnocení.

1.1 Mezinárodní standardy

V tomto dokumentu vycházíme z FAIR principů, jak je formulovali Wilkinson et al. (2016)⁷, tj. aby data byla dohledatelná, přístupná, interoperabilní a znovupoužitelná. Sladění pravidel a postupů uplatňovaných v českém prostředí s mezinárodními standardy umožňuje širší využití a sdílení výzkumných dat napříč vědeckou komunitou. Při analýze a budoucí implementaci jsou proto zohledňovány výstupy řady iniciativ. Mezi východiska tak patří zejména iniciativy:

⁴ <https://www.eosc.cz/>

⁵ https://opjak.cz/vyzvy/vyzva-c-02_23_014-open-science-i/

⁶ Principy FAIR <http://www.go-fair.org/fair-principles/>

⁷ WILKINSON, Mark D.; DUMONTIER, Michel; AALBERSBERG, IJsbrand Jan, APPLETON, Gabrielle; AXTON, Myles et al. *The FAIR Guiding Principles for scientific data management and stewardship*. Online. Scientific Data, vol. 3 (2016), article no. 160018. Dostupné z: <https://doi.org/10.1038/sdata.2016.18>.

GO FAIR⁸ – mezinárodní iniciativa zaměřená na podporu principů FAIR. Iniciativa byla založena v roce 2017 jako společná aktivita Nizozemska, Německa a Francie s cílem podporovat otevřenou vědu a implementaci FAIR principů. Koordinaci zajišťuje GO FAIR International Support and Coordination Office (GFISCO), které sídlí v Göttingenu (Německo). V současnosti je iniciativa GO FAIR podporována Evropskou komisí v rámci implementace EOSC, přičemž důraz je kladen na koordinaci se standardy a komunitními nástroji vyvíjenými v evropském výzkumném prostoru. Na tvorbě standardů a vývoji FAIR nástrojů spolupracuje vědecká komunita, univerzity i výzkumné ústavy.

FAIRsFAIR (Fostering FAIR Data Practices in Europe)⁹ - projekt financovaný Evropskou unií v rámci programu Horizon 2020, který podporuje implementaci principů FAIR v Evropě. FAIRsFAIR poskytuje FAIR nástroje, služby, doporučení a pokyny v rámci EOSC. Projekt se zaměřuje na praktická řešení FAIR v celém životním cyklu dat.

Cross-Domain Interoperability Framework (CDIF)¹⁰ - projekt cílí na zlepšení interoperability a opětovné použitelnosti FAIR dat napříč různými vědními obory a institucemi. CDIF poskytuje konkrétní doporučení a referenční standardy, které podporují interoperabilitu a opakovanou použitelnost dat v několika klíčových oblastech (metadata, řízené slovníky, podmínky přístupu a licence). Především se jedná o standardy pro interoperabilitu dat a jejich integraci, dále standardy pro popis podmínek přístupu k datům a jejich povoleného použití, doporučení pro publikaci kontrolovaných slovníků a další podklady. Tato doporučení jsou klíčovým základem pro nastavení postupů a podpory pro naplňování FAIR principů v rámci NRP/NDI.

WorldFAIR¹¹ - globální projekt zaměřený na podporu implementace principů FAIR pro sdílení výzkumných dat. Cílem projektu bylo vytvořit doporučení, rámce interoperability a pokyny pro hodnocení FAIR dat napříč různými vědními obory na příkladu 11 případových studií. Projekt je koordinován organizací CODATA (Committee on Data of the International Science Council)¹² ve spolupráci s RDA (Research Data Alliance)¹³.

Základní terminologie (datová sada, distribuce, landing page, PID, FIP, signposting aj.) je uvedena v úvodu dokumentu v části **Terminologie (glosář)**.

⁸ <https://www.go-fair.org/go-fair-initiative/>

⁹ <https://doi.org/10.3030/831558>

¹⁰ <https://worldfair-project.eu/cross-domain-interoperability-framework/>

¹¹ <https://worldfair-project.eu/the-project/>

¹² <https://codata.org/>

¹³ <https://www.rd-alliance.org/>

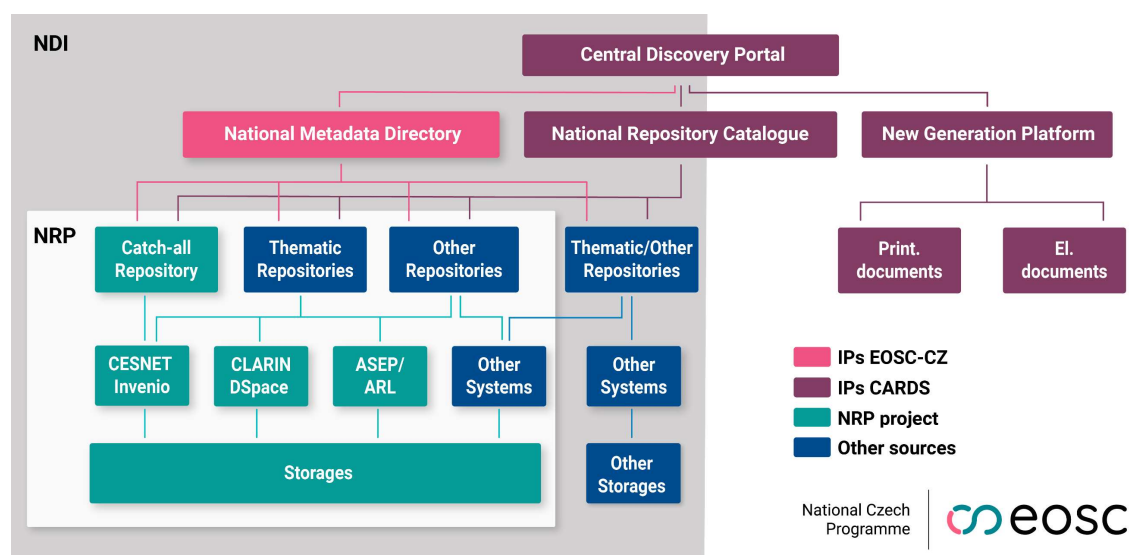
2 Architektura implementace FAIR v NRP/NDI

Tato kapitola představuje architekturu technické infrastruktury NRP/NDI s důrazem na implementaci FAIR principů. Zahrnuje také schémata, která znázorňují vztahy mezi uživateli infrastruktury a jednotlivými službami a nástroji, které jsou dostupné nebo v přípravě.

2.1 Technická infrastruktura pro FAIRifikaci

Implementace FAIR v rámci EOSC-CZ vychází z architektury NRP/NDI, která tvoří základní technickou a organizační vrstvu pro ukládání a správu výzkumných dat. Klíčovými komponentami jsou federované repozitáře, které umožňují uložení a vyhledávání výzkumných dat s podporou persistentních identifikátorů (PID) a standardizovaných metadatových schémat. Schéma (obr. 1) znázorňuje hlavní služby ekosystému NRP/NDI a jejich logické vazby z pohledu správy a sdílení metadat a dat; nejde o detailní technickou architekturu ani úplný výčet komponent.

Jak číst schéma: Repozitáře jsou zdrojem dat a metadat; centrální služby (např. NMA/NKR) slouží k agregaci a orientaci napříč repozitáři; podpůrné služby (např. PID/licence) doplňují publikační workflow. Poznámka o financování má pouze orientační charakter pro lepší znázornění implementační strategie.



Obr. 1 – Ekosystém služeb NRP/NDI relevantní pro publikaci dat a práci s metadaty (konceptuální přehled)

Ekosystém NRP/NDI je komplexní, jednotlivé služby jsou logicky propojené a jejich výstupy na sebe navazují. Z hlediska řešení problematiky FAIR principů a sdílení informací o datech jsou v rámci NRP/NDI klíčové následující služby:

Národní metadatový adresář (NMA)

NMA (National Metadata Directory)¹⁴ je k datu vydání verze 1.1 dostupný jako produkční služba; současně probíhá postupné zapojování dalších repozitářů a průběžné rozšiřování a ověřování vybraných funkcí (zejména nových integrací).

¹⁴ <https://nma.eosc.cz/>

Do budoucna má NMA fungovat jako centrální služba pro agregaci, ukládání a vyhledávání metadat o výzkumných datech z projektů realizovaných na českých výzkumných institucích, a to z různých domácích i zahraničních zdrojů. Jeho cílem je zlepšit dohledatelnost českých výzkumných dat a umožnit jejich efektivní vyhledávání napříč obory.

Národní katalog repozitářů (NKR)

NKR (National Repository Catalogue) bude evidovat a zpřístupňovat přehled repozitářů zapojených do NRP/NDI, včetně informací o použitých metadatových schématech, ontologiích a řízených slovnících, což napomůže standardizaci a usnadní integraci repozitářů do širšího ekosystému EOSC. Správci repozitářů budou zodpovědní za poskytování těchto informací, což zajistí aktuálnost a správnost katalogu.

Národní centrum PID

Národní centrum PID¹⁵ zajišťuje správu a koordinaci využívání perzistentních identifikátorů (PID) v České republice, zejména prostřednictvím členství v konsorciu DataCite a poskytování služeb DOI a dalších identifikátorů. Díky této službě mají české instituce možnost přidělovat DOI a další podporované typy PID, což usnadní citování, sledování a dlouhodobou dostupnost dat.

2.2 Prostředky podpory FAIR datové správy v NRP/NDI

Následující text shrnuje oblasti podpory v NRP/NDI, které prakticky přispívají k naplňování FAIR principů: **Findable** a **Interoperable** (standardizace a integrace metadat), **Accessible** (řízení přístupu včetně citlivých dat) a **Reusable** (licenční informace a podmínky využití, provenance). Detailní mapování na konkrétní FAIR principy (F1–R1.3) je uvedeno v tabulce v kapitole 3; zde uvádíme přehled hlavních oblastí podpory:

- **Integrace metadatových standardů:** podpora vyhledatelnosti a interoperability (CCMM, sklizeň, profily).
- **Podpora správy licencí:** podpora opětovné použitelnosti (jasné licenční podmínky v metadatech a workflow).
- **Řízení přístupu a citlivá data:** podpora přístupnosti při respektování omezení (autorizace, policy, specializovaná řešení).
- **Automatizace tvorby a sběru metadat:** podpora kvality, provenance a opětovné použitelnosti (metadata z přístrojů, ELN).

2.2.1 Integrace metadatových standardů

Jedním z hlavních pilířů NRP/NDI je automatizovaná agregace metadat ze všech zapojených repozitářů, která podpoří vyhledatelnost, interoperabilitu a opakované využití dat. Klíčovým prvkem této integrace je základní metadatový model (CCMM), který zajistí jednotné jádro pro strukturování metadat.

Správci repozitářů a odborné komunity mají k dispozici Dataspecer, nástroj pro tvorbu a správu metadatových profilů (datových specifikací) používaných pro popis dat v jednotlivých repozitářích. Dataspecer umožňuje profily verzovat, dokumentovat, exportovat (např. do SHACL/JSON Schema) a připravovat mapování mezi schémata (např. vůči CCMM, DCAT-AP nebo DataCite) za účelem zvýšení interoperability. Samotná dostupnost nástroje však nezaručuje vznik úplného přehledu používaných schémat napříč repozitáři – ten vznikne pouze

¹⁵ <https://www.techlib.cz/cs/84700-narodni-centrum-pro-perzistentni-identifikatory>

tehdy, pokud repozitáře budou profily v nástroji skutečně udržovat, vhodně publikovat a bude nastaven provozní a governance model (odpovědnosti, proces změn, pravidla publikace a správa oprávnění). Předpoklady pro efektivní správu profilů zahrnují zejména: správu verzí schémat, export do SHACL/JSON Schema, podporu mapování na DCAT-AP/DataCite, řízený proces změn (governance) a nástroje pro testování shody.

V NRP/NDI rozlišujeme dvě úrovně metadat a dva způsoby jejich zpřístupnění. Pro centrální vyhledávání v Národním metadatovém adresáři (NMA) repozitáře poskytují minimální, jednotné metadatové jádro (CCMM), které NMA sklízí přes OAI-PMH. Detailní oborová metadata zůstávají v repozitářích; k jejich plnému popisu se uživatelé i strojové služby dostanou přes landing page a/nebo API repozitáře. Záznam v NMA proto kromě CCMM obsahuje i odkaz na landing page (a kde je to dostupné, i na rozhraní pro strojový přístup k bohatším metadatům).

2.2.2 Podpora správy licencí

NRP poskytne nástroje a metodickou podporu pro výběr a přidělování licencí k datovým sadám. Tímto budou mít vědci a správci dat k dispozici prostředky pro jasné a jednotné nastavení podmínek přístupu, sdílení a využití výzkumných dat. Plánováno je zveřejnění interaktivního průvodce (Public Licence Selector), který uživatele provede výběrem vhodné licence a poskytne strojově čitelné vyjádření licence (minimálně URI). Cílem je zpřístupnit tuto službu samostatně (web/API) a současně – tam, kde to repozitářové platformy umožní – ji integrovat do publikačního workflow repozitářů; konkrétní forma integrace bude upřesněna pro jednotlivé platformy v implementačních profilech/FIP.

2.2.3 Řízení přístupu k datům a podpora pro citlivá data

I když by FAIR data měla být co nejotevřenější, ne vždy je to možné kvůli etickým, právním nebo bezpečnostním omezením. Proto NRP vyvíjí nástroje pro řízení přístupu, které umožní detailní nastavení oprávnění na úrovni datových sad i jednotlivých distribucí, včetně podpory pro data s omezeným nebo podmíněným přístupem. Pro správu citlivých dat, která vyžaduje speciální přístup, se plánuje vybudování specializovaných repozitářů určených pro tato data, vytvoření metodik, které budou reflektovat požadavky jednotlivých vědních disciplín, a implementace technologií, umožňujících bezpečné ukládání a sdílení citlivých datových sad.

2.2.4 Automatizace tvorby a sběru metadat

Součástí budoucího rozvoje infrastruktury NRP/NDI je také podpora automatizovaného sběru dat a metadat z výzkumných zařízení a/nebo datových akvizičních pipeline, například prostřednictvím sběrných služeb a API. To umožní okamžitý sběr a strukturovaný popis experimentálních dat přímo při jejich vzniku, což je zásadní pro zajištění jejich kvality a FAIR kompatibility. Metadata extrahovaná z přístrojů mohou být následně automaticky synchronizována s repozitáři a katalogy, což zjednoduší procesy FAIR správy dat.

2.3 Znázornění implementace FAIR v prostředí NRP/NDI

Výzkumníci, data stewardi a datoví kurátoři spolupracují na zajištění kvality a interoperability dat s využitím specializovaných nástrojů, standardů a automatizovaných procesů v rámci infrastruktury NRP/NDI. Tato kapitola obsahuje dvě schémata, která přehledně znázorňují předpokládaný způsob ukládání a správy vědeckých dat v ekosystému NRP/NDI s důrazem na implementaci FAIR principů. Klíčové vztahy mezi aktéry v rámci NDI jsou představeny v **Přehledovém konceptuálním schématu** ([obr. 2](#)), které ukazuje propojení mezi uživateli, datovými sadami, jejich metadaty, dílčími nástroji NDI a odbornými komunitami. Naproti tomu schéma **Ekosystém repozitáře: aktéři, pravidla a procesy** ([obr. 3](#)) se zaměřuje na strukturu vztahů v repozitáři, kde jsou definovány role uživatelů, správců a nástrojů zajišťujících kvalitu

a kontrolu dat. Ukazuje, jakým způsobem repozitář formuje procesy správy dat a metadat, včetně interakce s externími systémy.

Tato schémata jsou zásadní pro ujasnění, v jakých fázích workflow a v jakých částech datové infrastruktury je potřebné zohledňovat doporučení pro zvýšení souladu s FAIR principy u ukládaných datových sad. V Příloze je uveden podrobný popis jednotlivých uzlů ve schématech (viz [Příloha 2](#)). Ve schématech jsou aktéři znázorněni červeně, služby a nástroje žlutě a výstupy či produkty zeleně.

2.3.1 Přehledové konceptuální schéma implementace FAIR v NRP/NDI

Výzkumník nebo data steward (End User/Data Steward) připravuje a spravuje výzkumné datové sady (Datasets) a jejich metadatové záznamy (Metadata), které následně ukládá do repozitáře (Repository). V některých oborech mohou být výzkumná data a metadata generována výzkumnými zařízeními a/nebo datovými akvizičními pipeline (Research Instruments), jež zajišťují automatizovaný přenos dat do repozitáře prostřednictvím API. Koncový uživatel v takovém případě automaticky generovaná data a metadata obohacuje o další relevantní údaje, provádí jejich správu a řídí jejich uložení do repozitáře.

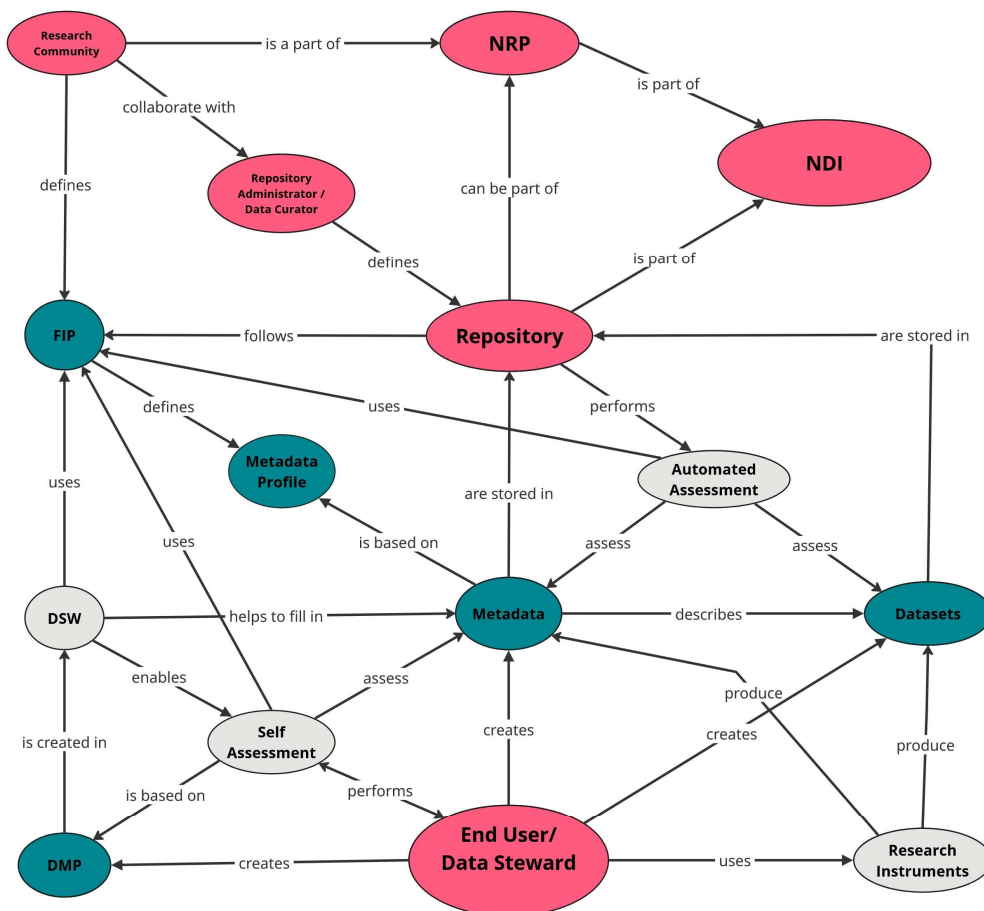
Pro efektivní správu dat využívá uživatel plán správy dat (DMP), který je optimálně vytvářen v nástroji podporujícím strojové zpracování, jako je například Data Stewardship Wizard (DSW)¹⁶. Tento nástroj umožňuje nejen systematickou správu výzkumných dat, ale také provádění sebehodnocení FAIRness (Self Assessment) na základě zadaných informací ještě před samotným vytvořením nebo uložením dat.

Jako člen oborové komunity (Research Community) se koncový uživatel může podílet na definování komunitních standardů a z nich odvozených FAIR implementačních profilů (FIP), které určují oborově specifické požadavky na naplňování FAIR principů, zejména bohaté a jednotné metadatové popisy datových sad (Metadata Profile). FAIR implementační profily (FIP) zajišťují oborově specifický soulad s FAIR principy a určují vhodné ontologie, řízené slovníky, formátové standardy a další kritéria nezbytná pro správnou správu, interpretaci a opakované využití dat.

Repozitář slouží k ukládání, správě a publikaci citovatelných datových sad v souladu s FAIR principy. Kromě toho zajišťuje technické, personální a procesní podmínky pro dlouhodobou archivaci dat.

Správce repozitáře (Repository Administrator) zajišťuje technické a provozní aspekty repozitáře. Určuje strukturu dat a datové standardy prostřednictvím konfigurace repozitáře. Integruje požadavky odborné komunity na využití FIP, čímž naplňuje odpovídající komunitní standardy. Uplatnění FIP a DMP podporuje nasazení automatizovaných nástrojů pro hodnocení FAIRness datových sad (Automated Assessment) v repozitářích. Správce repozitáře úzce spolupracuje s datovým kurátorem (Data Curator), který stanovuje pravidla pro ukládání dat a správu metadat. Kurátor zajišťuje kontrolu kvality a compliance datových sad a metadat nad rámec automatizovaných kontrol, schvaluje a zveřejňuje záznamy a harmonizuje metadata pro zajištění interoperability s dalšími systémy v rámci Národní datové infrastruktury (NDI).

¹⁶ <https://ds-wizard.org/>



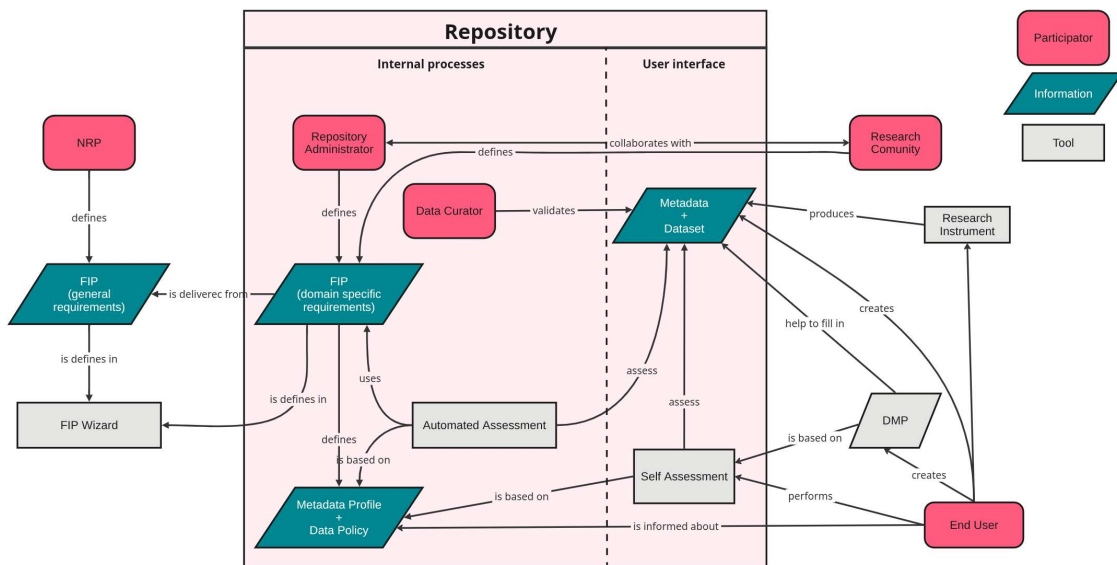
Obrázek 2: Přehledové konceptuální schéma implementace FAIR v NRP/NDI

2.3.2 Ekosystém repozitáře: aktéři, pravidla a procesy

Toto schéma ukazuje, jakým způsobem repozitář řídí proces ukládání dat a metadat, aby byl v souladu s FAIR principy. Uživatel je při své práci ovlivněn několika klíčovými prvky:

- **Definice pravidel:** Práci uživatele vymezuje konfigurace repozitáře, kterou nastavuje správce a kurátor na základě tří vzájemně provázaných typů pravidel:
 - **Metadatový profil (Metadata Profile):** určuje *jaká pole* jsou povinná/volitelná, jaké jsou kontrolované slovníky a validace (např. povinná licence, typ dat, vazba na projekt/grant).
 - **Datová politika repozitáře (Data Policy):** definuje *podmínky přijímání a správy dat* (přístupová práva, licencování, verzování, doba uchování, postup pro stažení dat, přidělování PID).
 - **FAIR Implementation Profile (FIP, volitelně):** doplňuje specifická rozhodnutí a doporučení komunity užívající repozitář (např. preferované ontologie/slovníky, oborová schémata a formáty) a slouží jako vodítko pro konfiguraci repozitáře a kurátorské postupy v dané komunitě.

- **Praktické nástroje a procesy:** Během vkládání dat je uživatel podporován automatickými kontrolami, přednastavenými ontologiemi a řízenými slovníky a následným kurátorským procesem.
- **Nástroje pro zpětnou vazbu:** Přímo v uživatelském rozhraní by měly být dostupné nástroje pro ověření souladu nahrávaných dat, konkrétně:
 - **Rychlá kontrola souladu (Self Assessment)**
 - **Automatizované hodnocení (Automated Assessment)**
- **Strojově čitelné vyjádření pravidel (cílový stav):** Cílem je, aby klíčová pravidla repositáře byla dostupná také ve strojově čitelné podobě a mohla být využita externími nástroji (např. pro plánování správy dat/DMP). Minimální rozsah zahrnuje: (i) zveřejnění metadatového profilu ve strukturované podobě (např. SHACL/JSON Schema), (ii) strojově jednoznačnou informaci o licencích a přístupových podmínkách (minimálně jako URI v metadatech) a (iii) seznam podporovaných formátů/typů distribucí. Konkrétní způsob publikace (API/registry/dokumenty) bude sjednocen v implementačních profilech a doporučeních pro repositářové platformy.



Obrázek 3: Ekosystém repositáře: aktéři, pravidla a procesy

3 FAIR Principy a jejich výklad v kontextu NRP/NDI

Tabulka v této kapitole přináší systematický přehled FAIR principů a jejich interpretace v kontextu NRP/NDI. Obsahuje také jejich propojení na požadavky CoreTrustSeal (CTS) a doporučení Cross-Domain Interoperability Framework (CDIF). Tabulka vychází z oficiálního dokumentu FAIRsFAIR Data Object Assessment Metrics (v0.5)¹⁷ a dále jsou zohledněny i závěry formulované v normativním článku "*FAIR Guiding Principles for scientific data management and stewardship*" (Wilkinson, M. et al. 2016). Tabulka slouží jako nástroj pro porozumění jednotlivým FAIR principům a jejich praktickému naplnění v rámci NRP/NDI.

Struktura tabulky a vysvětlení jednotlivých sloupců

Sloupec	Vysvětlení
Princip (CZ)	Český název daného FAIR principu, který shrnuje jeho základní význam.
Popis	Detailní vysvětlení principu a jeho důležitosti pro správu a sdílení vědeckých dat.
Doporučení CDIF	Relevantní doporučení CDIF, která pomáhají zajistit mezioborovou interoperabilitu dat.
Požadavky CTS	Požadavky certifikace CoreTrustSeal (CTS), které souvisejí s daným FAIR principem a podporují jeho naplnění na úrovni repozitáře.
Naplnění v NRP	Jak je daný princip implementován v rámci NRP, včetně technických a organizačních opatření.
Požadavky na NRP	Specifické požadavky, které musí splňovat NRP pro naplnění daného FAIR principu.
Požadavky na repozitáře	Požadavky na jednotlivé repozitáře zapojené do NDI, k dosažení souladu s FAIR principy.
Požadavky na FIP	Konkrétní implementační kroky a standardy, které mají být definovány ve FIP pro jednotlivé vědní obory pro naplnění FAIR principů.

Kromě samotných FAIR principů tabulka zahrnuje dvě důležité doplňkové dimenze:

- **Požadavky CoreTrustSeal (CTS)** – Certifikace CTS slouží k vyhodnocení, zda repozitář splňuje standardy pro důvěryhodné uchovávání a správu dat. Zaměřuje se na aspekty jako je dlouhodobá dostupnost a integrita dat, způsob správy metadat, právní a etické aspekty uchovávání dat, politika, udržitelnost, organizace a řízení datové infrastruktury a podpora FAIR principů na úrovni repozitářů a procesů. Jedná se o důležitý pilíř dobré praxe, který doprovází a posiluje FAIR principy. Další informace o CTS a jeho aktuálních požadavcích jsou dostupné na oficiálním webu CoreTrustSeal¹⁸.
- **Doporučení Cross-Domain Interoperability Framework (CDIF)**¹⁹ je rámec pro mezioborovou interoperabilitu dat, který poskytuje praktická doporučení pro správu metadat, přístupová práva a licencování, řízené slovníky a klasifikace a integraci a propojení dat napříč obory. Každému FAIR principu jsou v tabulce přiřazena relevantní doporučení CDIF, která podporují mezioborovou interoperabilitu metadat, licencí, řízených slovníků i přístupových politik v rámci širšího výzkumného ekosystému.

¹⁷ <https://doi.org/10.5281/zenodo.6461229>

¹⁸ Aktuální požadavky CoreTrustSeal (2023–2025) <https://www.coretrustseal.org/why-certification/requirements/>

¹⁹ <https://worldfair-project.eu/cross-domain-interoperability-framework/>

Další informace o CDIF a jeho implementaci v rámci EOSC jsou dostupné na webu EOSC²⁰ případně v souhrnném přehledu CDIF doporučení²¹.

Tabulka má sloužit jako referenční nástroj pro další klíčové aktivity NRP/NDI, správce repozitářů a další aktéry v oblasti výzkumných dat. Může být využita pro mapování aktuálního stavu FAIR implementace v NRP/NDI a v repozitářích a k identifikaci oblastí pro zlepšení, podporu mezioborové interoperability prostřednictvím aplikace doporučení CDIF a pro vytvoření konkrétních plánů na implementaci FAIR principů v rámci jednotlivých vědních komunit prostřednictvím FIP. Tabulka poskytuje komplexní pohled na implementaci FAIR principů v rámci NRP/NDI, přičemž zohledňuje jak technické aspekty FAIR dat, tak organizační požadavky na repozitáře. Naplnění FAIR principů a splnění požadavků a doporučení CoreTrustSeal a CDIF umožní robustní a důvěryhodnou správu výzkumných dat, která odpovídá mezinárodním standardům a osvědčeným praktikám. Tato tabulka je tedy klíčovým nástrojem pro plánování a řízení FAIRifikace dat a repozitářů v rámci NRP/NDI.

3.1 Implementace FAIR principů v NRP/NDI

Princip (EN)	F1: (Meta) data is assigned a globally unique and persistent identifier
Princip (CZ)	F1: (Meta)datům je přiřazen globálně jedinečný a trvalý identifikátor
Popis	<p>Tento princip upozorňuje, že každá datová sada a její metadata mají být jednoznačně identifikovatelné prostřednictvím globálně jedinečného a trvalého identifikátoru (PID). PID lze v případě potřeby přidělit i jednotlivým distribucím. Použití PID v rámci metadat umožňuje jejich efektivní vyhledávání, správné citování a dlouhodobou udržitelnost.</p> <p>Nejčastěji používané PID v oblasti výzkumných dat zahrnují:</p> <ul style="list-style-type: none"> • DOI (Digital Object Identifier) – nejrozšířenější systém pro identifikaci vědeckých dat. DOI je nadstavba nad Handle System; v praxi je preferováno pro publikace a výzkumná data. • Handle System – systém pro perzistentní identifikátory, často využívaný repozitáři. • ARK (Archival Resource Key) – identifikátor využívaný v digitálních archivech. • PURL (Persistent Uniform Resource Locator) – přesměrovací služba umožňující uchování stabilních odkazů na digitální objekty. <p>Další příklady jedinečných identifikátorů jsou IRI, URI.</p>
Doporučení CDIF	Relevantní CDIF doporučení: žádná
Požadavky CTS	CoreTrustSeal R12 vyžaduje, aby repozitáře umožnily uživatelům vyhledávat datové sady a odkazovat na ně trvalým způsobem prostřednictvím správného citování. To znamená, že každá uložená datová sada musí mít PID, který je indexovatelný a dlouhodobě dostupný. Oba požadavky (FAIR F1 a CoreTrustSeal R12) se shodují v klíčovém aspektu: nutnosti trvalého a globálně unikátního identifikátoru pro metadata a data.

Implementace CDIF v rámci EOSC <https://eosc.eu/roadmap/cross-domain-interoperability-framework/>

²¹ <https://worldfair-project.eu/cross-domain-interoperability-framework/>

<p>Naplnění v NRP</p>	<p>V cílovém stavu NRP vyžaduje, aby každá publikovaná datová sada měla přidělený perzistentní identifikátor (PID), typicky DOI (případně Handle). PID je povinnou součástí metadat a musí být resolvovatelný na landing page. U převzatých/legacy záznamů, které PID dosud nemají, je jako minimální přechodné řešení akceptována stabilní URL; doplnění PID je součástí procesu onboarding/migrace.</p> <p>Přidělování DOI bude v NRP prioritně realizováno prostřednictvím služeb Národního centra DOI (NC DOI) koordinovaného Centrem PID v NTK, které institucím zajišťuje smluvní a metodickou podporu a přístup k registrační agentuře DataCite (včetně zřízení repozitářového účtu a přidělení DOI prefixu). Samotné DOI následně přiděluje (mintuje) příslušný repozitář/instituce při publikaci záznamu; alternativně mohou repozitáře využít i jiné uznávané systémy PID (např. Handle) dle svých potřeb a možností.</p>
<p>Požadavky na NRP</p>	<ul style="list-style-type: none"> • Zajištění PID pro publikované datové sady: Každá publikovaná datová sada v NRP musí mít přiřazený perzistentní identifikátor (PID), který bude součástí metadat. U převzatých záznamů bez PID musí být minimálně stabilní URL a definovaný postup doplnění PID. • Podpora typů PID (rozsah): NRP bude garantovat podporu přidělování DOI pro datové sady (prostřednictvím služeb NC DOI / DataCite). V metadatech NMA bude možné evidovat i jiné uznávané PID (např. Handle, ARK), pokud je přiděluje konkrétní repozitář/instituce; NRP je však centrálně nezajišťuje a jejich dostupnost závisí na použité repozitářové platformě. • Zajistit mapování metadatových polí pro identifikátory (tj. kde je PID uveden a jaký má význam) mezi minimálním schématem NRP/CCMM a exportními profily používanými pro interoperabilitu (např. DataCite a DCAT-AP), aby se identifikátory při sklizni/exportu přenesly bez ztráty významu. • Integrace PID do metadatového schématu: Metadatový záznam v NMA musí obsahovat PID datové sady jako identifikátor záznamu a tento PID musí být resolvovatelný na landing page konkrétní datové sady. Záznam zároveň musí obsahovat jednoznačné odkazy na distribuce (např. download URL / access URL nebo API endpoint); pokud je některé distribuci přidělen samostatný PID, musí být v metadatech uveden jako identifikátor dané distribuce a být k ní explicitně přiřazen. • Zajištění dlouhodobé dostupnosti PID: PID musí být trvale dostupný a spravovaný tak, aby nedocházelo k jeho zneplatnění nebo ztrátě.
<p>Požadavky na repozitáře</p>	<ul style="list-style-type: none"> • Přidělení PID při publikaci: Repozitář/instituce zajistí přidělení (mintování) PID každé publikované datové sadě (preferovaně DOI; alternativně jiný uznávaný PID dle možností platformy). • PID v metadatech a resolvovatelnost: PID je uveden jako identifikátor datové sady v metadatovém záznamu a resolvuje na landing page. Identifikace distribucí je řešena dle principů F3/F4 (stabilní URL; v odůvodněných případech samostatný PID distribuce). • Podpora správy a aktualizace PID: Repozitář musí zajistit dlouhodobou resolvovatelnost PID (tj. PID musí vždy vést na platnou landing page; v případě odstranění dat na tombstone page). Při změnách infrastruktury/URL repozitář zajistí aktualizaci cílové URL PID (např. v registrační službě pro DOI) a potřebná přesměrování.

Požadavky na FIP	<ul style="list-style-type: none">• Specifikace preferovaných PID pro daný obor: FIP musí definovat, jaké typy PID jsou v daném oboru preferovány a jakým způsobem by měly být implementovány.• Doporučení pro správu PID: FIP by měl obsahovat pokyny, jak zajistit dlouhodobou správu PID a jejich propojení s dalšími metadatovými prvky.
-------------------------	---

Princip	F2: Data are described with rich metadata
Princip (CZ)	F2: Data jsou popsána bohatými metadaty
Popis	<p>Tato zásada vyžaduje, aby metadata poskytovala podrobný a bohatý popis, který umožní datové sady snadno vyhledávat, citovat a opakovaně používat jinými subjekty. Bohatá metadata by měla obsahovat nejen základní popisné informace, jako jsou název, autor a datum vzniku, ale také detailní kontextová metadata, která specifikují původ dat (provenience), jejich strukturu, vztahy k jiným datovým sadám a podmínky použití.</p> <p>Mezi klíčové prvky metadat patří:</p> <ul style="list-style-type: none"> • Základní popisná metadata – název, autor, datum vzniku, přidělený persistentní identifikátor (PID). • Provenienční metadata – informace o metodologii sběru dat, použitých přístrojích a software. • Licenční metadata – podmínky použití dat a jejich případná omezení. • Interoperabilní metadata – využití standardizovaných ontologií a slovníků pro sémantickou interpretaci dat. <p>Tato zásada úzce souvisí se zásadou FAIR R1.</p>
Doporučení CDIF	Discovery profil podporuje použití Schema.org pro klíčová metadata, což přispívá k bohatému popisu dat. Controlled Vocabularies profil doporučuje použití SKOS a OWL pro řízené slovníky a sémantické propojení metadat.
Požadavky CTS	Kvality metadat se přímo či nepřímo týkají požadavky CoreTrustSeal R8, R11, R13 a R14. FAIR F2 a CoreTrustSeal jsou v přímém souladu v tom, že vyžadují bohatá, kvalitní a indexovaná metadata. Metadata musí být dostatečně detailní, aby umožnila správné vyhledávání, citování a opětovné použití dat. Metadata musí být standardizovaná, bohatá a strukturovaná tak, aby byla kompatibilní se standardy používanými danou oborovou komunitou (souvislost s R1.3, dále rovněž souvisí s licenčními metadaty, R1.1)
Naplnění v NRP	<p>V kontextu NRP/NDI je důležité rozlišovat dvě úrovně metadat.</p> <p>Pro centrální vyhledávání v NMA je sklíženo standardizované metadatové jádro (CCMM), které zajišťuje základní interoperabilitu. Plně bohatá a oborově specifická metadata zůstávají v repozitářích a jsou dostupná prostřednictvím jejich rozhraní (landing page, API). Oborové komunity si mohou CCMM pro své potřeby rozšířit nebo poskytovat vedle CCMM i druhý metadatový záznam ve svém oborovém standardu.</p> <p>Minimální požadavky na metadata pro zařazení do NMA jsou definovány jako jednotné metadatové jádro CCMM a jsou udržovány Pracovní skupinou EOSC-CZ Metadata ve spolupráci s provozovatelem NMA a správci repozitářů. CCMM stanovuje povinné prvky a validační pravidla pro centrální vyhledávání; repozitáře mohou vedle CCMM poskytovat i bohatší oborová metadata, která zůstávají v repozitáři. Oborové komunity mohou CCMM rozšířit prostřednictvím metadatových profilů a FIP; NMA v takovém případě sklízí CCMM a v metadatech uvádí odkazy na zdroj kompletního oborového popisu (landing page/API).</p>

	Pro NRP byly definovány minimální požadavky pomocí CCMM schématu ²² , které profiluje metadatová schémata DataCite a DC ²³ . Doménově agnostické vlastnosti schémat DCAT-AP a DataCite umožňují dostatečně flexibilní a konzistentní identifikaci dat pro účely citování a vyhledávání.
Požadavky na NRP	<ul style="list-style-type: none"> • Definování sady minimálních požadavků na metadata (zajišťuje CCMM). • Zajištění toho, aby každá datová sada obsahovala povinná metadata, zahrnující klíčové informace o původu, licencování apod.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Povinné zahrnutí klíčových metadatových atributů: Repozitáře musí zajistit, že každá datová sada obsahuje alespoň minimální metadatový záznam, splňující požadavky NMA, a zajistit jeho mapování na metadatový standard CCMM. • Podpora oborově specifických metadat: Repozitář musí umožnit, aby vedle CCMM bylo možné použít také oborový/komunitní metadatový profil popsaný ve FIP (např. jako samostatné schéma nebo rozšíření formuláře/API). Minimálně musí být ve FIP uvedeno, jaký oborový profil se používá a kde je dostupný plný oborový metadatový záznam (landing page/API), zatímco NMA sklízí CCMM jako společné minimum.
Požadavky na FIP	<ul style="list-style-type: none"> • Definování požadavků na metadata pro daný obor: Každý FIP by měl obsahovat seznam povinných a doporučených metadatových profilů relevantních pro konkrétní vědeckou komunitu.

<https://github.com/techlib/CCMM>

²³ <https://semiceu.github.io/DCAT-AP/releases/3.0.0/>

Princip	F3: metadata clearly and explicitly include the identifier of the data they describe
Princip (CZ)	F3: Metadata jasně a explicitně uvádějí identifikátor dat, která popisují
Popis	Tato zásada FAIR zdůrazňuje nutnost zajistit, že metadata obsahují jednoznačný a explicitní odkaz na datovou sadu, kterou popisují. Landing page datové sady by proto měla obsahovat lidsky i strojově čitelná metadata s příslušnými odkazy na jednotlivé distribuce (datové soubory nebo API) přístupné přes standardní protokoly.
Doporučení CDIF	Všeobecné doporučení CDIF stanovuje, že metadata by měla být uložena v JSON-LD ²⁴ na landing pages nebo ve formě samostatných souborů. To umožňuje jasné propojení metadat s datovou sadou a zajišťuje jednoznačné propojení mezi identifikátorem a odpovídajícími daty. Doporučené je použití signpostingu.
Požadavky CTS	Tento princip úzce souvisí s požadavky CoreTrustSeal, zejména s R13 (Možnost vyhledávání a citování) a R8 (Uložená metadata). Metadata musí jasně obsahovat identifikátor datové sady, aby byla data srozumitelná a trvale propojena s jejich metadaty.
Naplnění v NRP	NRP bude zajišťovat podporu PID, které mohou být přiděleny datovým sadám. Distribuce budou označeny pomocí stabilní URL. V odůvodněných případech jim může být přidělen samostatný PID. NRP bude po repozitářích požadovat jasnou metodiku tvorby odkazu na datovou sadu. Pro strojové zpracování NRP doporučuje signpost ²⁵ (HTTP Link hlavičky) jako preferovaný způsob strojově čitelné navigace mezi landing page, metadaty a distribucemi. Minimálním požadavkem je, aby landing page obsahovala strojově čitelná metadata (např. JSON-LD) s explicitními odkazy na distribuce; tam, kde platforma umožňuje nastavovat HTTP hlavičky (přímo nebo přes webovou vrstvu), je signposting požadován jako preferovaný mechanismus.
Požadavky na NRP	<ul style="list-style-type: none"> Zajištění jednoznačného propojení metadat a datových sad: NRP definuje minimální technickou specifikaci propojení: (i) identifikátor v metadatach (PID/URL) musí vést na landing page datové sady, (ii) landing page musí obsahovat strojově čitelná metadata (např. JSON-LD) s odkazy na distribuce, (iii) preferovaně je toto propojení doplněno signpostingem (HTTP Link hlavičky), pokud jej platforma umožňuje. Podpora standardizovaných identifikačních polí: Metadata musí obsahovat jedinečné identifikátory (např. URL) distribucí dat (datových souborů, API), které umožňují jednoznačnou identifikaci a zpřístupnění konkrétní distribuce.
Požadavky na repozitáře	<ul style="list-style-type: none"> Zahrnutí identifikátoru dat do metadat: Repozitář musí implementovat strojově zjistitelné propojení: metadata → landing page → distribuce (stabilní URL/endpointy). Preferovaným mechanismem je signposting (HTTP Link hlavičky); pokud jej repozitářová platforma neumí nativně, lze jej zajistit na webové vrstvě (např. reverse proxy) nebo minimálně publikací JSON-LD metadat na landing page s explicitními odkazy na distribuce.

²⁴ <https://json-ld.org/>

²⁵ Metoda zpřístupnění metadat a souvisejících zdrojů pomocí HTTP hlaviček a odkazů <https://signposting.org/>

Požadavky na FIP	<ul style="list-style-type: none">• Specifikace, jakým způsobem má být identifikátor dat zahrnut do metadat: FIP musí obsahovat odkaz na metadatový profil, který definuje standardizované pole pro identifikátor dat.
-------------------------	--

Princip	F4: (meta)data are registered or indexed in a searchable resource
Princip (CZ)	F4: (meta)data jsou registrována nebo indexována v prohledávatelných zdrojích
Popis	<p>Tato zásada FAIR zajišťuje, že metadata a data jsou registrována nebo indexována ve zdrojích, které umožňují jejich snadné vyhledávání a přístup. Registrace metadat v prohledávatelných repozitářích a agregátorech dat zvyšuje jejich dohledatelnost a umožňuje jejich integraci do širších výzkumných ekosystémů.</p> <p>Metadata by měla být indexována v:</p> <ul style="list-style-type: none"> • Národním metadatovém adresáři (např. NMA). • Mezinárodních agregátorech (např. OpenAIRE, DataCite Commons, Google Dataset Search). • Oborových vyhledávačích (např. PubMed, Europe PMC, PANGAEA). • Obecných vyhledávacích službách (např. Google, Bing).
Doporučení CDIF	Discovery profil zdůrazňuje potřebu publikovat metadata ve standardizovaných formátech (např. Schema.org nebo DCAT), aby byla strojově čitelná a indexovatelná vyhledávacími službami a agregátory. Tím se zvyšuje jejich dohledatelnost pro výzkumníky i automatizované nástroje.
Požadavky CTS	Tento princip úzce souvisí s požadavky CoreTrustSeal R8 (Uložená metadata) a R13 (Možnost vyhledávání a citování). CoreTrustSeal R8 požaduje, aby metadata byla strojově čitelná a indexovaná v odpovídajících metadatových systémech. Požadavek R13 zdůrazňuje, že metadata musí být registrována ve vyhledávačích a uživatelé musí mít možnost citovat data prostřednictvím perzistentních identifikátorů.
Naplnění v NRP	Základní metadata datových sad publikovaných v repozitářích zapojených do NRP jsou sklížena a indexována v NMA (primárně prostřednictvím OAI-PMH v profilu CCMM), aby byla vyhledatelná z jednoho místa a zároveň připravená pro další sdílení. Pro webovou a katalogovou indexaci jsou metadata zpřístupňována také ve strojově čitelných reprezentacích (např. Schema.org/JSON-LD a exportní profil typu DCAT-AP); tím je umožněna integrace do mezinárodních agregátorů (např. OpenAIRE), přičemž integrace do DataCite Commons je podmíněna DOI registrovaným u DataCite.
Požadavky na NRP	<ul style="list-style-type: none"> • Automatická indexace metadat: Metadata musí být automaticky indexována v NMA a zpřístupněna pro vyhledávání. • Podpora rozhraní a exportních profilů pro indexaci: NRP/NMA musí podporovat OAI-PMH pro sklizeň metadat a současně poskytovat metadata i ve strojově čitelných exportních profilech pro indexaci (např. Schema.org/JSON-LD pro webové vyhledávače a DCAT-AP pro katalogové/integrační scénáře), dle implementačních možností jednotlivých platforem. • Registrace v mezinárodních agregátorech: Metadata by měla být pravidelně synchronizována s mezinárodními databázemi (OpenAIRE, DataCite Commons). • Optimalizace pro vyhledávače: Platformy podporované v NRP by měly být co nejlépe optimalizovány pro indexaci běžnými internetovými vyhledávači.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Poskytování metadat přes OAI-PMH: Repozitáře musí pro účely automatické sklizeň zpřístupnit metadata ve dvou klíčových formátech: <ul style="list-style-type: none"> ▫ CCMM (Czech Core Metadata Model) pro zajištění vyhledatelnosti v Národním metadatovém adresáři (NMA).

	<ul style="list-style-type: none"> ▫ Dublin Core pro splnění základních podmínek užívání protokolu OAI-PMH. • Podpora indexace a exportu: Repozitář musí vedle OAI-PMH zajistit strojově čitelnou reprezentaci metadat pro indexaci (minimálně Schema.org/JSON-LD na landing page nebo ekvivalentní export). Strojové API (např. REST) je doporučené rozšíření pro pokročilé integrace, není však minimální podmínkou pro splnění F4. • Validace metadat: Před uložením musí být metadata validována, aby byla zajištěna jejich úplnost a správnost pro efektivní indexaci.
<p>Požadavky na FIP</p>	<ul style="list-style-type: none"> • Definování minimální úrovně metadat pro indexaci: FIP by měl specifikovat, jaké atributy musí metadata obsahovat, aby byla úspěšně indexována v relevantních mezinárodních vyhledávačích. • Doporučení pro kompatibilitu se standardy: FIP by měl obsahovat pokyny pro správné mapování metadat na standardizované formáty používané ve vyhledávacích systémech.

Princip	A.1 (meta)data are retrievable by their identifier using a standardized communications protocol
Princip (CZ)	A1: (Meta)data jsou dostupná pomocí identifikátoru prostřednictvím standardizovaného komunikačního protokolu
Popis	Cílem této zásady je zajistit, aby data i metadata byla dostupná prostřednictvím jednoznačného identifikátoru (např. DOI, Handle ²⁶) za použití standardizovaného komunikačního protokolu, který je široce používán a podporován. Protokol by měl umožnit snadný a spolehlivý přístup k datům bez nutnosti specializovaného softwaru. Základním komunikačním protokolem je obvykle HTTP/HTTPS a pro sklizení metadat specificky určený OAI-PMH.
Doporučení CDIF	Relevantní CDIF doporučení: žádná
Požadavky CTS	Tento princip úzce souvisí s požadavkem CoreTrustSeal R15 (Technická infrastruktura). R15 požaduje, aby repozitáře disponovaly odpovídající technickou infrastrukturou, která zajistí spolehlivou správu a přístup k datům.
Naplnění v NRP	Data a metadata uložená v NRP jsou přístupná prostřednictvím perzistentních identifikátorů (PID) za použití standardních komunikačních protokolů. Pro zajištění kompatibility s mezinárodními standardy NRP podporuje protokoly HTTPS pro přímý přístup a OAI-PMH pro sklizeň metadat externími systémy, jako je OpenAIRE.
Požadavky na NRP	<ul style="list-style-type: none"> • Podpora protokolů: NRP musí podporovat standardní a široce používané komunikační protokoly (HTTPS, OAI-PMH). V odůvodněných případech může být rozšířeno o další protokoly, např. SPARQL pro dotazování nad metadaty. • Standardizace přístupu k metadatům: Metadata každé datové sady musí obsahovat informaci o tom, jakým způsobem jsou dostupná pro strojové zpracování. Doporučenou metodou je signposting, které pomocí HTTP hlaviček odkazuje na různé reprezentace metadat (např. ve formátu JSON-LD, RDF/XML). Odkaz na OAI-PMH endpoint repozitáře, kde jsou metadata dostupná, by měl být součástí záznamu repozitáře v Národním katalogu repozitářů (NKR).
Požadavky na repozitáře	<ul style="list-style-type: none"> • Kompatibilita s podporovanými protokoly: Všechny repozitáře musí implementovat alespoň jeden standardizovaný protokol pro přístup k datům a metadatům.
Požadavky na FIP	<ul style="list-style-type: none"> • Definování standardů pro přístup: FIP musí obsahovat seznam podporovaných protokolů. • Oborově specifické protokoly: Pokud určité obory vyžadují specifické protokoly, měly by být jasně uvedeny v příslušných FIP.

²⁶ <https://www.handle.net/>

Princip	A1.1 the protocol is open, free, and universally implementable
Princip (CZ)	A1.1 Protokol je otevřený, zdarma k dispozici a univerzálně použitelný
Popis	Tento princip zdůrazňuje, že komunikační protokoly používané pro přístup k datům a metadatům musí být otevřené, bezplatné a univerzálně implementovatelné. Cílem je zajistit, aby k datům mohl přistupovat kdokoli bez nutnosti specializovaného softwaru nebo placených licencí. Mezi doporučené protokoly patří zejména HTTPS a OAI-PMH.
Doporučení CDIF	Relevantní CDIF doporučení: žádná
Požadavky CTS	Viz A1.
Naplnění v NRP	NRP podporuje otevřené protokoly, jako je HTTPS pro přímý přístup k datům a OAI-PMH pro sklizení metadat externími systémy.
Požadavky na NRP	Implementace otevřených protokolů: NRP musí zajistit, že veškerý přístup k datům a metadatům probíhá přes otevřené, bezplatné a univerzálně podporované komunikační protokoly. Minimálním požadavkem pro všechny repozitáře je podpora HTTPS pro přímý přístup a OAI-PMH pro sklizeň metadat. Podpora dalších protokolů (např. SPARQL) se řídí oborovými potřebami definovanými v příslušných FIP.
Požadavky na repozitáře	Použití otevřených protokolů: Všechny repozitáře musí používat otevřené a široce přijímané protokoly pro zpřístupnění dat a metadat, minimálně však HTTPS a OAI-PMH.
Požadavky na FIP	Definice otevřených protokolů pro oborové využití: Každý FIP by měl specifikovat, které protokoly jsou doporučené pro daný obor.

Princip	A1.2 the protocol allows for an authentication and authorization procedure, where necessary
Princip (CZ)	A1.2: Protokol v případě potřeby umožňuje postup ověřování a autorizace
Popis	Tato zásada zohledňuje to, že přístup k některým datům může být omezený, a že citlivé nebo chráněné informace musí být dostupné pouze oprávněným uživatelům. Protokoly musí umožnit bezpečné ověření identity uživatelů (autentizace) a řízení jejich přístupových práv (autorizace). Tyto postupy zajišťují, že data mohou být přístupná, aniž by byla ohrožena jejich bezpečnost. Mezi doporučené autentizační a autorizační protokoly patří HTTP Basic Auth ²⁷ , OAuth ²⁸ , SAML ²⁹ nebo LDAP ³⁰ . Při implementaci těchto protokolů je důležité vyvážit potřebu zabezpečení a snadnost přístupu pro koncové uživatele.
Doporučení CDIF	Relevantní CDIF doporučení: žádná
Požadavky Co-reTrustSeal	Viz A1.
Naplnění v NRP	NRP zajišťuje ověřování a autorizaci přístupu k datům prostřednictvím klíčové aktivity AAI, zaměřené na bezpečný a standardizovaný přístup k výzkumným datům. Cílem je, aby všechny repozitáře zapojené do NRP měly možnost implementovat vhodné mechanismy ověřování identity uživatelů a jejich autorizace, zejména v případě citlivých nebo omezeně dostupných dat. Systém řízení přístupu bude integrován s existujícími autentizačními službami, což zajistí kompatibilitu s národními i mezinárodními standardy a identitními federacemi (např. eduID.cz), a umožní jednotné ověřování uživatelů napříč repozitáři. Tím bude možné efektivně spravovat přístupová práva a zajistit bezpečné nakládání s daty v souladu s principy FAIR a legislativními požadavky.
Požadavky na NRP	<ul style="list-style-type: none"> Podpora autentizačních a autorizačních protokolů: NRP musí zajistit, aby repozitáře (případně i samotná úložiště) mohla zavést nezbytné autentizační a autorizační služby, pokud to bude povaha dat vyžadovat.
Požadavky na repozitáře	<ul style="list-style-type: none"> Implementace autentizačních služeb: Repozitáře musí umožňovat přístup k datům prostřednictvím autorizačních a autentizačních protokolů (např. HTTP Basic Auth, OAuth, SAML) tam, kde je to nezbytné. Definice přístupových práv: Správci repozitářů musí nastavit jasná pravidla pro řízení přístupu a rozlišovat mezi různými úrovněmi oprávnění (veřejná data, data před zveřejněním, citlivá data, atd).
Požadavky na FIP	<ul style="list-style-type: none"> Doporučení ohledně bezpečnosti: FIP by měl zahrnovat doporučení pro bezpečný přístup k datům prostřednictvím autentizačních protokolů. Specifikace přístupových pravidel: FIP by měl obsahovat doporučení pro implementaci autentizace a autorizace v konkrétním oboru, zejména u dat, která podléhají etickým, bezpečnostním nebo právním omezením.

²⁷ <https://www.rfc-editor.org/info/rfc7617>

²⁸ <https://oauth.net/2/>

²⁹ <https://www.oasis-open.org/standard/saml/>

³⁰ <https://ldap.com/>

Princip	A2. metadata are accessible, even when the data are no longer available
Princip (CZ)	A2: Metadata jsou přístupná, i když data již nejsou k dispozici
Popis	Tato zásada vyžaduje, aby metadata zůstala přístupná i v případě, že samotná data již nejsou k dispozici. Důvody pro ztrátu dat mohou zahrnovat vypršení časového limitu pro jejich uchování, změnu podmínek přístupu nebo smazání dat na žádost uživatele. Přesto je důležité, aby metadata byla nadále veřejně dostupná jako trvalý záznam o existenci datové sady, její provenienci a vlastnostech. Jednou z doporučených praktik (viz doporučení DataCite) je použití tzv. tombstone pages (náhrobků), které uchovávají základní informace o odstraněné datové sadě včetně důvodu její nedostupnosti.
Doporučení CDIF	CDIF všeobecně doporučuje ukládání metadat na landing pages nebo do samostatných souborů. Díky tomu metadata zůstávají dostupná i v případě, že původní data byla odstraněna nebo přesunuta, což pomáhá zajistit dlouhodobou dohledatelnost vědeckých výstupů.
Požadavky CTS	Tento princip úzce souvisí s požadavky CoreTrustSeal R8 (Uložená metadata) a zejména R10 (Plán uchování dat). Požadavek CoreTrustSeal R10 vyžaduje, aby repozitáře měly jasnou strategii pro dlouhodobou správu metadat i dat.
Naplnění v NRP	Repozitáře budou mít povinnost zajistit dlouhodobou dostupnost metadat, i v případě, že samotná data již nebudou k dispozici. NRP vyžaduje, aby metadata byla veřejně přístupná a propojená s perzistentním identifikátorem (PID), který povede na odpovídající záznam v repozitáři nebo na tombstone page. Minimální záznam metadat bude navíc uložen v Národním metadatovém adresáři (NMA).
Požadavky na NRP	NRP bude zajišťovat uchování a zveřejnění minimálního záznamu metadat v NMA, a to i v případě, že související data již nejsou dostupná.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Uchování metadat: Repozitáře musí zajistit, že metadata budou uchovávána i po odstranění příslušných datových sad. V rámci repozitáře musí být zachován úplný záznam metadat a musí být řádně namapován na minimální záznam v NMA. • Implementace tombstone pages: Zajistit, že každá datová sada má po odstranění přístupnou stránku obsahující PID, citaci a důvod nedostupnosti. • Přesměrování PID na tombstone page: Repozitáře musí implementovat systém přesměrování PID na stránku s informacemi o odstraněných datech.
Požadavky na FIP	<ul style="list-style-type: none"> • Specifikace minimálních metadat: FIP by měl jasně definovat, která metadata musí zůstat veřejně dostupná i po ztrátě dat.

Princip	I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
Princip (CZ)	I1.: (Meta)data používají pro reprezentaci znalostí jazyk, který je formální, dostupný, sdílený a široce aplikovatelný
Popis	Tato princip požaduje, aby metadata i data byla reprezentována sémantickým jazykem, který je formální, přístupný, sdílený a široce použitelný. To umožní jejich snadnou integraci a zpracování napříč různými systémy a obory. Cílem je používat jazyky, které mají formálně definovanou syntaxi a sémantiku a umožňují strojové zpracování. Je důležité rozlišovat mezi datovým modelem či frameworkem (např. RDF ³¹), jeho syntaxí/serializací (např. JSON-LD, Turtle, XML ³²) a ontologickými jazyky (např. OWL ³³), které na těchto technologiích staví. Vedle technologií sémantického webu lze tento princip naplnit i použitím jiných zavedených a v komunitě široce používaných formálních jazyků, jako jsou standardizované XML-based formáty.
Doporučení CDIF	CDIF profil Controlled Vocabularies doporučuje SKOS ³⁴ pro řízené slovníky a OWL pro sémantickou interoperabilitu. Tyto technologie umožňují propojení dat s jinými zdroji v rámci Linked Open Data a zlepšují automatizovanou analýzu dat. CDIF profil ³⁵ definuje standardy pro popis univerzálních konceptů, jako je čas, geografie a měrné jednotky, což je klíčové pro sjednocení různorodých datových struktur.
Požadavky CTS	Tento princip souvisí s požadavky CoreTrustSeal R14 (Podpora opětovného použití dat) a R15 (Technická infrastruktura). R14 CoreTrustSeal požaduje, aby repozitáře zajistily, že metadata a data budou podrobně popsána a znovu použitelná v průběhu času. R15 pak požaduje, aby repozitáře měly adekvátní technickou infrastrukturu, která umožní správné zpracování metadat v různých systémech.
Naplnění v NRP	Metadata a data mají být poskytována v sémanticky interoperabilních formátech. Doporučené jsou především frameworky umožňující propojení a strojovou interpretaci (např. RDF s různými serializacemi jako Turtle, JSON-LD, XML; nebo OWL pro ontologické modelování). Zároveň je přípustné využívat i jiné formáty, pokud jsou dostatečně standardizované a umožňují jednoznačnou výměnu a interpretaci dat (např. některé XML-based workflow používané v oborových komunitách). Rozlišujeme zde rámec (framework, např. RDF) a jeho konkrétní serializace (např. JSON-LD). Cílem není omezit se na jedinou technologii, ale podpořit interoperabilitu a opakovatelné zpracování dat v prostředí NRP/NDI.
Požadavky na NRP	<ul style="list-style-type: none"> • Používání standardizovaných jazyků pro reprezentaci znalostí: Metadata musí být publikována v sémantických jazycích, které umožňují modelování vztahů a propojení dat (např. RDF, OWL). • Použití strojově čitelných formátů: Metadata musí být dostupná v datových formátech umožňujících správnou serializaci a přenos metadat (např. JSON-LD, RDF/XML, Turtle, XML).

³¹ <https://www.w3.org/RDF/>

³² <https://www.w3.org/TR/REC-xml/>

³³ <https://www.w3.org/OWL/>

³⁴ <https://www.w3.org/2004/02/skos/>

³⁵ <https://zenodo.org/records/11236871> 6.7. Universals: Time, Geography, and Units of Measurement

Požadavky na repozitáře	<ul style="list-style-type: none"> • Používání standardizovaných jazyků pro reprezentaci znalostí: Metadata musí být publikována v sémantických jazycích, které umožňují modelování vztahů a propojení dat (např. RDF, OWL). • Použití strojově čitelných formátů: Metadata musí být dostupná v datových formátech umožňujících správnou serializaci a přenos metadat (např. JSON-LD, RDF/XML, Turtle, XML).
Požadavky na FIP	<p>Definování sémantiky a syntaxe: FIP musí specifikovat:</p> <ul style="list-style-type: none"> • Sémantické zdroje: Doporučené ontologie a řízené slovníky pro popis dat a vztahů mezi nimi (např. SKOS, OWL, PROV-O, SIO). • Syntax a serializaci: Preferované formáty pro strojově čitelnou reprezentaci metadat (např. JSON-LD, RDF/XML, Turtle).

Princip	I2. (meta)data use vocabularies that follow FAIR principles
Princip (CZ)	I2.: (Meta)data používají slovníky, které se řídí zásadami FAIR
Popis	<p>Tento FAIR princip požaduje, aby metadata a data používala slovníky a ontologie, které splňují FAIR principy. To znamená, že slovníky musí být dobře zdokumentované, veřejně přístupné a propojené s globálně jedinečnými perzistentními identifikátory (PID), pokud jsou k dispozici.</p> <p>Aby byly slovníky FAIR, měly by splňovat následující požadavky:</p> <ul style="list-style-type: none"> • Ideálně by měly být jednoznačně identifikovatelné prostřednictvím PID (F1), avšak v praxi je často používán URI/URL adresa vedoucí na strojově čitelnou verzi slovníku. • Musí být veřejně dostupné prostřednictvím standardních protokolů (A1). • Musí používat formální, přístupný a široce použitelný jazyk (I1). <p>Důležitou součástí FAIR slovníků jsou globální autoritní registry, které poskytují jedinečné identifikátory pro vědce (ORCID)³⁶, organizace (ROR)³⁷ a další entity. Pro identifikaci konceptů a pojmů se využívají zejména perzistentní identifikátory ve slovnících a znalostních bázích, jako je např. Wikidata nebo doménové řízené slovníky publikované ve formátu SKOS. Tyto registry umožňují lepší propojení metadat, zvýšení jejich dohledatelnosti a sémantickou interoperabilitu. Katalog FAIR kompatibilních slovníků a standardů nabízí stránky FAIRsharing.org.³⁸ Příklady FAIR slovníků:</p> <ul style="list-style-type: none"> • DCMI Metadata Term³⁹ (obecná metadata) • Schema.org⁴⁰ (webová metadata) • Gene Ontology⁴¹ (GO) (biomedicína) • MeSH⁴² (Medical Subject Headings) (medicína) • Wikidata⁴³ (globálně propojená znalostní databáze) • ROR (Research Organization Registry) (identifikátory výzkumných organizací) • ORCID (identifikátory výzkumníků)
Doporučení CDIF	Profil Controlled Vocabularies podporuje použití FAIR slovníků a doporučuje standard XKOS ⁴⁴ (rozšíření standardu SKOS o funkcionality potřebné pro statistická a ekonomická data). Tímto bude zajištěno, že metadata budou konzistentně strukturována a snadno interpretovatelná v různých vědeckých disciplínách.
Požadavky CTS	Tento princip souvisí s požadavkem CoreTrustSeal R14 (Podpora opětovného použití dat). R14 CoreTrustSeal vyžaduje, aby repozitáře zajistily, že metadata a data budou interpretovatelná a znovu použitelná v průběhu času.

³⁶ <https://info.orcid.org/what-is-orcid/>

³⁷ <https://ror.org/>

³⁸ <https://fairsharing.org>

³⁹ <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

⁴⁰ <https://www.w3.org/community/schemaorg/>

⁴¹ <https://geneontology.org/>

⁴² <https://www.nlm.nih.gov/mesh/meshhome.html>

⁴³ https://www.wikidata.org/wiki/Wikidata:Main_Page

⁴⁴ https://ddialliance.org/xkos_v1.0

Naplnění v NRP	<p>Připravovaný Národní katalog repozitářů (NKR) bude sloužit jako nástroj pro evidenci repozitářů, metadatových schémat a řízených slovníků používaných v NDI.</p> <p>Pro účely interoperability je třeba zajistit komplementární vývoj bohatých metadatových schémat pro jednotlivé vědní obory – využití řízených slovníků, standardizovaných ontologií, tezaurů a taxonomií všude, kde to bude možné. Vhodné je proto požadovat po repozitářích vzniklých v rámci NRP rozšíření metadat nezbytných pro vyhledávání a citování o oborově specifická, bohatá metadata, která budou lépe sloužit potřebám odlišných výzkumných komunit. Rozsah metadat škáluje s přínosem pro discovery/reuse; repozitář nemá vyžadovat oborová metadata, která nevyužije (ponechat na FIP/komunitě).</p>
Požadavky na NRP	<ul style="list-style-type: none"> • Dokumentace slovníků v NKR: Používané slovníky a ontologie by měly být evidovány v NKR.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Metadata mají využívat sdílené a dobře dokumentované řízené slovníky a ontologie, které jsou dostupné prostřednictvím perzistentních identifikátorů (např. URI) a publikovány v otevřených repozitářích (např. LOV, BioPortal, OLS). • Repozitáře musí podporovat slovníky a ontologie doporučené komunitou, včetně těch uvedených v příslušném FIP. Pokud žádný vhodný standard neexistuje, může komunita definovat vlastní sémantický model, který bude zdokumentován a sdílen prostřednictvím FIP. • Napojení metadat na autoritní registry: Metadata by měla obsahovat odkazy na ORCID (vědci), ROR (organizace) pro zajištění jedinečné identifikace entit. • Možnost rozšíření o komunitní slovníky: Repozitáře by měly umožnit přidání oborově specifických slovníků podle potřeb vědeckých komunit.
Požadavky na FIP	<ul style="list-style-type: none"> • Specifikace doporučených slovníků a ontologií pro metadata i data: Každý FIP by měl definovat, jaké slovníky a ontologie jsou preferované v daném oboru, včetně pokynů pro jejich správné používání.

Princip	I3. (meta)data include qualified references to other (meta)data
Princip (CZ)	I3.: (Meta)data obsahují odkazy na další (meta)data
Popis	<p>Tento princip požaduje, aby metadata obsahovala kvalifikované odkazy na další související metadata a datové sady. Propojená data umožňují lepší kontextualizaci, integraci a znovupoužitelnost v širších výzkumných a informačních ekosystémech.</p> <p>Kvalifikované odkazy znamenají, že:</p> <ul style="list-style-type: none"> • Vztahy mezi metadaty a datovými sadami jsou explicitně definovány (např. vztah k původní datové sadě, odvozené verzi, doprovodným publikacím). • Používají se standardizované vztahy definované v metadatových schématech (např. DataCite, PROV-O⁴⁵, DCAT-AP). • Odkazy jsou strojově čitelné, což umožňuje automatizovanou analýzu a propojení. <p>Důležitou součástí kvalifikovaných odkazů jsou propojení na autority jako ORCID (vědci), ROR (organizace) a Wikidata (koncepty). Tyto odkazy poskytují globálně unikátní identifikátory, které umožňují efektivní propojení metadat a usnadňují jejich dohledatelnost.</p> <p>Příklad vztahů mezi metadaty ve schématu DataCite:</p> <ul style="list-style-type: none"> • isDerivedFrom – označuje datovou sadu, ze které jsou data odvozena. • isCitedBy – odkazuje na publikaci, která datovou sadu cituje. • hasPart / isPartOf – propojuje datové sady, které jsou součástí většího celku. • affiliation – propojuje datovou sadu s organizací prostřednictvím ROR identifikátoru. • creator – propojuje datovou sadu s autorem prostřednictvím ORCID identifikátoru.
Doporučení CDIF	Data Integration profil doporučuje využití DDI-CDI ⁴⁶ pro popis vztahů mezi logickou strukturou dat a jejich fyzickou reprezentací. To umožňuje jasně definovat vztahy mezi různými datovými sadami a zajistit, že uživatelé mohou data správně interpretovat a propojit s jinými relevantními zdroji.
Požadavky CTS	Tento princip souvisí s požadavkem CoreTrustSeal R14 (Podpora opětovného použití dat) a s požadavkem R11 (Zajištění kvality dat). R11 vyžaduje, aby koncoví uživatelé měli k dispozici dostatek informací pro hodnocení kvality. R14 se soustředí na opakované použití dat v průběhu času a zajištění toho, aby byla k dispozici vhodná metadata, která podporují porozumění a používání dat.
Naplnění v NRP	Je nezbytné, aby datové sady a související metadata obsahovaly kvalifikované odkazy na další relevantní metadata nebo zdroje dat. To zahrnuje odkazy na externí ontologie, datové sady, autoritní registry (ORCID, ROR) nebo jiná úložiště, která poskytují další kontext nebo informace, případně verze (meta)dat.
Požadavky na NRP	<ul style="list-style-type: none"> • Zajištění podpory pro kvalifikované odkazy: Metadata musí obsahovat pole umožňující propojení s jinými metadaty a datovými sadami.

⁴⁵ <https://www.w3.org/TR/prov-o/>

⁴⁶ https://ddialliance.org/ddi-cdi_v1.0

	<ul style="list-style-type: none"> • Použití standardizovaných ontologií pro popis vztahů: Metadata musí používat široce uznávané ontologie pro definování vztahů mezi entitami (např. PROV-O, DataCite, DCAT-AP). • Podpora autoritních identifikátorů: Metadata musí umožňovat propojení na ORCID, ROR.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Zahrnutí odkazů na související data: Metadata musí obsahovat odkazy na původní datovou sadu, publikace a další relevantní zdroje. • Validace odkazů na externí metadata: Repozitáře by měly zajistit, že odkazy v metadatach vedou na dostupné a správné zdroje. • Podpora verzování a sledování změn: Každá datová sada by měla být propojena se svými předchozími verzemi a odvozenými datovými sadami.
Požadavky na FIP	<ul style="list-style-type: none"> • Definování pravidel pro propojení metadat: FIP by měl specifikovat, jaké vztahy mezi metadaty jsou povinné a jaké doporučené pro daný obor. • Doporučení pro využití mezinárodních ontologií: Každý obor by měl definovat preferované slovníky a modely pro popis vztahů mezi datovými sadami.

Princip	R1. meta(data) are richly described with a plurality of accurate and relevant attributes
Princip (CZ)	R1. meta(data) jsou bohatě popsána množstvím přesných a relevantních atributů
Popis	Tento princip klade důraz na podrobný a přesný popis datových sad prostřednictvím metadat. Metadata by měla obsahovat dostatečně detailní informace, které umožní potenciálním uživatelům pochopit kontext dat, jejich původ (provenienci), použité metody a licenční podmínky pro opětovné použití. Bohatá metadata zvyšují srozumitelnost dat a usnadňují jejich znovupoužitelnost v různých kontextech. Pro plnění tohoto principu je důležité používat standardizované metadatové schéma, zahrnout jasné informace o autorech, datu vzniku, metodách sběru dat, jejich verzi a použitých formátech, standardech nebo řízených slovnících.
Doporučení CDIF	Relevantní CDIF doporučení: žádná
Požadavky CTS	Tento princip úzce souvisí s požadavky CoreTrustSeal R8 (Uložená metadata), R11 (Zajištění kvality dat) a R14 (Podpora opětovného použití dat). R8 vyžaduje, aby metadata byla dostatečně podrobná, aby umožnila dlouhodobé uchování a správnou interpretaci dat. R11 požaduje, aby repozitáře měly jasné definované procesy pro validaci kvality metadat, čímž se zajistí jejich konzistence a správnost.
Naplnění v NRP	V rámci NRP/NDI se pro naplnění tohoto principu rozlišují dvě úrovně metadat: <ul style="list-style-type: none"> • Discovery profil (CCMM): Jádro metadatových prvků sklizené do NMA, které slouží primárně k vyhledání a identifikaci datové sady. • Reusability profil (Oborová metadata): Bohatý soubor metadat, který zůstává v repozitáři a je definován oborovými standardy (např. v rámci FIP). Tento profil obsahuje detailní informace nezbytné pro znovupoužití dat, jako jsou podrobná provenience, popis metodologie a kontext vzniku dat. <p>NMA slouží jako rozcestník, zatímco plnou znovupoužitelnost zajišťují bohatá metadata dostupná na landing page v repozitáři.</p>
Požadavky na NRP	<ul style="list-style-type: none"> • Zajistit, aby minimální metadatové schéma obsahovalo všechny klíčové atributy potřebné pro opakované použití dat.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Implementace minimálního metadatového schématu: Všechna základní metadata musí být v souladu s požadavky NMA. • Podpora oborově specifických metadat: Repozitáře by měly umožňovat přidání oborově specifických metadatových prvků definovaných v příslušných FIP. • Verzování metadat: Doporučuje se podporovat verzování metadat. V systémech, které plné verzování metadat neumožňují, by mělo být alespoň zajištěno, že nové verze datové sady obdrží nový PID (DOI) a budou explicitně propojeny s předchozí verzí pomocí relačních vlastností v metadatach (např. isNewVersionOf ve schématu DataCite).
Požadavky na FIP	<ul style="list-style-type: none"> • Definování povinných metadatových atributů. • Popis používaných slovníků a standardů: FIP by měl obsahovat seznam doporučených slovníků a standardů, které zvýší interoperabilitu dat.

Princip	R1.1. (meta)data are released with a clear and accessible data usage license
Princip (CZ)	R1.1. (meta)data jsou zveřejněna s jasnou a přístupnou licencí na používání dat
Popis	Tento princip zdůrazňuje, že každá datová sada a její metadata mají mít jasně definovanou a snadno dostupnou licenci, která specifikuje podmínky použití. Licenční podmínky musí být strojově čitelné a srozumitelné jak pro lidi, tak pro automatizované systémy. Transparentnost licenčních podmínek je klíčová pro umožnění opakovaného použití dat bez právních nebo etických nejasností. Mezi často používané licence patří Creative Commons (např. CC BY 4.0), Open Data Commons nebo specifické oborové licence. Jasně definovaná licence by měla být součástí metadat a měla by být propojena s datovou sadou. Klíčová je také strojová čitelnost licence. V nejjednodušší formě to znamená uvedení licence prostřednictvím jejího standardizovaného URI (např. https://creativecommons.org/licenses/by/4.0/), jak to vyžaduje např. schéma DataCite (vlastnost rightsURI). Pokročilejší reprezentace mohou využívat formální jazyky jako CC REL, ODRL nebo SPDX.
Doporučení CDIF	Data Access profil doporučuje použití ODRL ⁴⁷ pro popis licencí a podmínek použití. Tímto způsobem mohou být licence snadno interpretovány automatizovanými systémy, což usnadňuje dodržování licenčních podmínek při opětovném použití dat.
Požadavky CTS	Tento princip souvisí s požadavkem CoreTrustSeal R2 (Licence). R2 vyžaduje, aby repozitáře spravovaly všechny platné licence, které se vztahují na přístup k datům a jejich používání.
Naplnění v NRP	<p>Na úrovni NRP tuto problematiku řeší Klíčová aktivita “Podpora práce s licencemi”, která se zaměřuje na sjednocení přístupu k licencování dat a jejich správu v metadatech. Tato problematika je řešena ve spolupráci NRP s projektem IPs CARDS. Konkrétní implementace licencí na úrovni repozitářů bude definována prostřednictvím FIP. NRP vyžaduje, aby každá datová sada měla přiřazenou licenci, která je součástí metadatového záznamu (název licence, URI). Metadata obsahující licenční informace jsou dostupná prostřednictvím NMA.</p> <p>Pro zajištění strojové čitelnosti licenčních podmínek je doporučeno používat standardizované formáty, jako jsou:</p> <ul style="list-style-type: none"> • Creative Commons Rights Expression Language (CC REL)⁴⁸ • Dublin Core (DC Terms) • DCAT-AP • JSON-LD • ODRL (Open Digital Rights Language) • SPDX (Software Package Data Exchange)⁴⁹, zejména pokud se metadata vztahují k softwarovým komponentám nebo open-source datovým sadám. <p>SPDX může být použit pro standardizované označování licencí v metadatech repozitářů a je široce podporován v softwarových ekosystémech, zatímco ODRL a</p>

⁴⁷ <https://www.w3.org/ns/odrl/2/ODRL20.html>

⁴⁸ https://wiki.creativecommons.org/wiki/CC_REL

⁴⁹ <https://spdx.dev/>

	CC REL umožňují detailnější správu podmínek použití datových sad v širším kontextu.
Požadavky na NRP	Vytvořit a spravovat centrální katalog doporučených licencí (v rámci KA4.2) a poskytnout metodickou podporu pro jejich výběr. K dispozici bude nástroj Public Licence Selector, jehož vývoj a implementaci řeší Klíčová aktivita "Podpora práce s licencemi".
Požadavky na repozitáře	Repozitáře by měly umožňovat výběr z doporučených licencí, zajistit, aby u každé datové sady byla licence uvedena, a validovat její přítomnost a strojově čitelný formát (minimálně jako URI).
Požadavky na FIP	Doporučit sadu licencí, které jsou v dané komunitě považovány za vhodné a přijatelné, a specifikovat formáty pro strojově čitelné licence.

Princip	R1.2. (meta)data are associated with detailed provenance
Princip (CZ)	R1.2. (meta) data jsou spjata s podrobnou proveniencí
Popis	<p>Tento princip vyžaduje, aby metadata obsahovala podrobné informace o provenienci dat, včetně informací o jejich vzniku, transformacích, předchozích verzích a souvisejících datových sadách. Metadata by měla obsahovat podrobné informace o původu a historii dat, včetně:</p> <ul style="list-style-type: none"> • Tvůrců a přispěvatelů, kteří se na vzniku dat podíleli. • Použitých metodologií a nástrojů, které byly využity při sběru, zpracování a analýze dat. • Verzování dat, které umožňuje sledování změn a aktualizací datové sady. • Odkazů na zdrojová data, pokud byla data transformována nebo odvozena z jiných datových sad.
Doporučení CDIF	Data Integration profil se zaměřuje na propojení dat s jejich původem a transformacemi pomocí DDI-CDI. Tento přístup umožňuje sledovat historii datových sad a zajistit jejich důvěryhodnost a reprodukovatelnost.
Požadavky CTS	<p>Tento princip úzce souvisí s požadavky CoreTrustSeal R7 (Integrita a autenticita dat) a R14 (Podpora opětovného použití dat). R7 vyžaduje, aby repozitáře zajišťovaly integritu a autenticitu dat, což zahrnuje i uchování informací o provenienci. R14 požaduje, aby repozitáře zajistily, že metadata budou dostatečně podrobná, aby umožnila opakované použití dat. Důležitou součástí tohoto požadavku je zachování informací o provenienci, včetně citací a odkazů na zdroje dat. FAIR R1.2 a CoreTrustSeal R7, R14 jsou tedy úzce propojené, protože shodně požadují uchování provenienčních informací pro zvýšení důvěryhodnosti a opětovné využitelnosti dat.</p>
Naplnění v NRP	<p>Implementace podrobného sledování provenience je strategickým cílem. V současné fázi se implementace zaměřuje na zajištění základních provenienčních informací v metadatach:</p> <ul style="list-style-type: none"> • Jednoznačná identifikace tvůrců a přispěvatelů (pomocí ORCID). • Popis klíčových kroků zpracování dat (formou textového popisu metodologie). • Propojení na zdrojová data, pokud je to relevantní (pomocí PID). <p>Do budoucna se počítá s podporou strukturované, strojově čitelné provenience s využitím standardizovaných sémantických modelů, jako je W3C PROV Ontology (PROV-O). Konkrétní požadavky budou definovány v návaznosti na výstupy klíčových aktivit zaměřených na automatizaci sběru dat a ověřování souladu s definovanými pravidly a standardy (compliance).</p>
Požadavky na NRP	<ul style="list-style-type: none"> • Standardizace metadat o provenienci: Zajistit, že metadatové schéma v NRP bude obsahovat informace o původu dat, včetně autorů, metodologie a historie změn, a to minimálně v rozsahu požadavků CCMM. • Podpora verzování: Implementovat mechanismy pro sledování změn v datových sadách a jejich metadatach. • Propojení s PID a publikacemi: Umožnit propojení datových sad s jejich odpovídajícími vědeckými publikacemi a projekty prostřednictvím DOI, ORCID a dalších perzistentních identifikátorů.

Požadavky na repozitáře	<ul style="list-style-type: none"> • Povinné zahrnutí provenienčních informací: Repozitáře musí zajistit, že metadata obsahují základní informace o původu dat, jejich verzi a historii změn, a to minimálně v rozsahu požadavků CCMM. • Podpora nástrojů pro sledování verzí: Umožnit uchování různých verzí datové sady a metadat, aby bylo možné dohledat jejich historii.
Požadavky na FIP	<ul style="list-style-type: none"> • Definování minimálního rozsahu provenienčních informací: Každý FIP by měl specifikovat, jaké informace o provenienci jsou pro daný obor povinné a doporučené. • Doporučené standardy pro popis provenience: FIP by měl obsahovat doporučení pro využití existujících standardů, jako jsou PROV-O nebo DataCite Metadata Schema.

Princip	R1.3. (meta)data meet domain-relevant community standards
Princip (CZ)	R1.3. (meta)data splňují komunitní standardy relevantní pro danou oblast
Popis	<p>Tento princip zdůrazňuje nutnost dodržování komunitně uznávaných standardů pro metadata a datové formáty. Cílem je zajistit, aby data byla konzistentní, srozumitelná a snadno interoperabilní v rámci konkrétní vědní oblasti. Použití komunitních standardů umožňuje lepší sdílení a znovuvyžití dat jak v národním, tak mezinárodním kontextu. Vědecké komunity si často vyvíjejí vlastní standardy pro metadata a formáty dat, které odpovídají specifickým potřebám jejich oborů. Například:</p> <ul style="list-style-type: none"> • Biologie: Darwin Core pro biologická data. • Biomedicína: MIAME (Minimum Information About a Microarray Experiment)⁵⁰ • Sociální vědy: DDI-C (Data Documentation Initiative - Codebook)⁵¹ • Fyzika: HDF5 (Hierarchical Data Format)⁵² • Humanitní vědy: CIDOC-CRM (CIDOC Conceptual Reference Model)⁵³
Doporučení CDIF	Controlled Vocabularies profil doporučuje použití SKOS a OWL pro integraci komunitních standardů do metadat. To zajišťuje, že data budou kompatibilní s mezinárodními výzkumnými infrastrukturami a umožní snadnější spolupráci mezi vědními obory.
Požadavky CTS	Tento princip opět úzce souvisí s požadavkem CoreTrustSeal R14 (Podpora opětovného použití dat). R14 vyžaduje, aby metadata byla kompletní a podrobná, aby umožnila opětovné použití dat v různých výzkumných komunitách.
Naplnění v NRP	V rámci NRP je řešena implementace komunitně uznávaných standardů pro metadata a datové formáty na základě spolupráce s projektem IPs CARDS. NRP by mělo podporovat ukládání dat přednostně ve vhodném oborovém repozitáři. Tato problematika je adresována také zavedením FIP profilů, které budou napomáhat k tomu, že data uložená v rámci repozitářů v NRP budou splňovat příslušné standardy specifické pro danou doménu.
Požadavky na NRP	<ul style="list-style-type: none"> • Podpora komunitních standardů: Zajistit, aby repozitáře podporovaly metadata a datové standardy odpovídající jednotlivým vědním oborům, především na základě FIP profilů. • Validace standardů: Zavést mechanismy pro podporu (částečně automatického sebehodnocení a validaci metadat vůči komunitním standardům (např. pomocí FAIR-Checker⁵⁴, F-UJI⁵⁵). • Doporučení pro nové obory: Pro vědní oblasti, kde komunitní standardy dosud nejsou plně definovány, podpořit vývoj a přijetí nových standardů, ve spolupráci s projektem OSII.

⁵⁰ <https://pubmed.ncbi.nlm.nih.gov/11726920/>

⁵¹ <https://ddialliance.org/ddi-codebook>

⁵² <https://www.hdfgroup.org/solutions/hdf5/>

⁵³ <https://cidoc-crm.org/>

⁵⁴ <https://fair-checker.france-bioinformatique.fr/>

⁵⁵ <https://www.f-uji.net/>

Požadavky na repozitáře	<ul style="list-style-type: none"> • Spolupráce s oborovou komunitou: V rámci FIP definovat metadatové schéma odpovídající standardům daného oboru. • Validace metadat: Repozitáře musí umožnit validaci metadat oproti definovaným komunitním schématům. • Možnost rozšíření o specifické standardy: Repozitáře by měly umožňovat implementaci nově definovaných či aktualizovaných oborových standardů podle potřeb vědeckých komunit.
Požadavky na FIP	<ul style="list-style-type: none"> • Definování komunitních standardů pro jednotlivé obory: Každý FIP by měl obsahovat seznam doporučených metadatových a datových standardů relevantních pro danou vědeckou oblast. • Specifikace minimálních požadavků na metadata: FIP by měl definovat povinné a doporučené atributy metadat.

4 Služby podporující implementaci FAIR v rámci oborů

4.1 FAIR Implementation Profiles

FAIR principy záměrně nestanovují ani nedoporučují žádné konkrétní technologie, nástroje, standardy nebo postupy – jsou to pouze obecná pravidla, která mají podpořit maximální využití stávajících a osvědčených řešení a usměrnit vývoj nových tam, kde v současnosti chybí. Cesta k FAIR datům musí být z podstaty věci do značné míry uchopena na úrovni jednotlivých komunit, které potřebují efektivní metody a nástroje k dosažení optimálního naplnění požadavků FAIR. K tomuto směřuje iniciativa FAIR Implementation Profile (FIP)⁵⁶ – Profil implementace FAIR, což je metodika vyvinutá v rámci iniciativy GO FAIR⁵⁷. Díky ní mají vědecké a výzkumné komunity možnost vyjádřit svůj výběr standardů, technologií, nástrojů a postupů, tzv. FAIR-Enabling Resources (FER)⁵⁸, zdroje umožňující FAIR, kterými naplňují FAIR principy. Tím se vytváří globální znalostní báze, ze které mohou čerpat, jak samotní členové dané komunity, tak i další komunity, které se mohou inspirovat při tvorbě vlastních FIP.

Na rozdíl od jiných přístupů zaměřených na vytváření pokynů a doporučení pro FAIRifikaci dat FIP nabízí i plně strojově zpracovatelné výstupy, např. ve formě nano-publikací. Ty umožňují automatizovanou podporu a propojení s nástroji, a zároveň zajišťují sledovatelnost a důvěryhodnost v globálně distribuovaném prostředí založeném na digitálně podepsaném původu dat.

FIP určují například jaké formáty a struktury dat a metadat jsou vhodné, jak podrobné mají být, a jaké další nástroje a zdroje jsou potřeba pro jejich správnou interpretaci. Tyto profily jsou vytvářeny odbornými skupinami, a mohou popisovat jak současné postupy, tak i cíle pro zlepšení správy dat. FIP pomáhají rozhodnout, jaké nástroje a zdroje podporující zásady FAIR by měly být použity, jako například řízené slovníky, ontologie, standardy pro metadata nebo persistentní identifikátory. FIP mohou být aplikovány na různých úrovních, ať už je to jednotlivá datová sada, kolekce dat nebo celý repozitář. Slouží jako vzor pro správu dat, zefektivňuje přípravu plánu správy dat a jeho zlepšování.

Oborovou specifičnost je v NDI zamýšleno zohlednit právě pomocí FIP, které budou definovat dobrou praxi oborově specifických repozitářů pro popis a parametry dat do nich ukládaných. Podobně bylo postupováno i ve významných celosvětových projektech, např. v projektu WorldFAIR⁵⁹, kde bylo vytvořeno 11 modelových FIP z rozsáhlých případových studií rozličných disciplín. Závěry projektu WorldFAIR zdůrazňují význam FIP a práce budou pokračovat v navazujícím projektu WorldFAIR+.

4.2 FIP Wizard

FIP Wizard (FAIR Implementation Profile Wizard)⁶⁰ je interaktivní webový nástroj určený k usnadnění tvorby a správy FIP. Uživatelům umožňuje systematicky definovat, dokumentovat

⁵⁶ <https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/>

⁵⁷ <https://www.go-fair.org/go-fair-initiative/>

⁵⁸ <https://w3id.org/fair/fip/terms/FAIR-Enabling-Resource>

⁵⁹ <https://worldfair-project.eu/the-project/>

⁶⁰ <https://catalogue.fair-impact.eu/node/212>

a sdílet své FIP. Nástroj je založen na open-source projektu Data Stewardship Wizard⁶¹, nabízí tedy podobně pohodlné rozhraní pro tvorbu a správu FIP v podobě dynamických inteligentních dotazníků. FIP Wizard informace z profilů umožňuje odeslat do distribuovaného systému nanopublikací⁶², což je strojově zpracovatelné tvrzení opatřené metadaty. Tím je zajištěna transparence a trasovatelnost (provenance). Uložené FAIR-Enabling Resources je následně možné používat dalšími komunitami pro sestavování FIP, což významně přispívá ke konvergenci a využívání vhodných prostředků pro zajištění FAIR.

V rámci NDI bude FIP Wizard využíván pro tvorbu FIP, jak bylo zmíněno výše. Bude spravována vlastní instance, která bude technicky nezávislá na globální síti nanopublikací z důvodu zajištění plné technické kontroly na národní úrovni, nicméně tyto globální nanopublikace bude možné využívat a bude možné do nich i přispívat.

4.3 Integrované řešení pro plánování správy dat

FIP představuje zdroj informací, které mohou významně usnadnit přípravu kvalitního plánu správy dat (DMP) – výzkumníci se mohou opřít o kvalitní, prověřená a dobře použitelná řešení pro zajišťování FAIR principů, která jsou navíc komunitně zaměřena, a použít je do svého plánu správy dat. K tomu bude na technické úrovni sloužit vyvinuté propojení mezi FIP Wizard a Data Stewardship Wizard⁶³, které zajistí přenos vhodných FER z FIP do DMP, jak je popsáno v sekci [2.3.1](#) a [2.3.2](#).

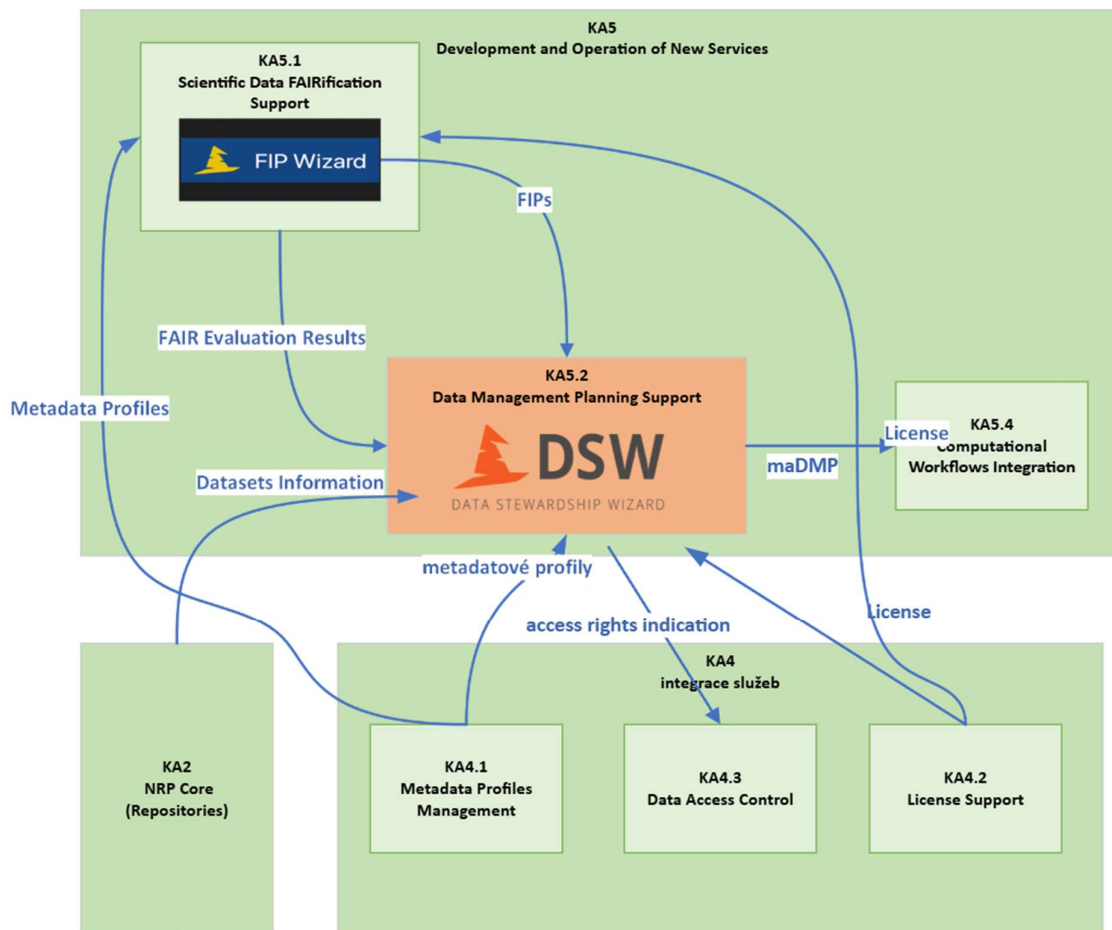
Samotný Data Stewardship Wizard poté představuje nejen nástroj pro plánování správy dat, ale může sloužit též ve všech částech životního cyklu dat jako “kokpit pilota”. To je umožněno strojovou zpracovatelností plánů dat (machine-actionable DMP, maDMP⁶⁴) a rozsáhlými možnostmi technických integrací skrze API. Konkrétní možnosti integrací jsou v tuto chvíli předmětem dalších technických analýz. Kromě popsané integrace s FIP Wizardem je na ideové úrovni možné uvažovat o integracích znázorněných ve schématu níže ([obr. 4](#)).

⁶¹ <https://ds-wizard.org/>

⁶² <https://nanopub.net/>

⁶³ <https://ds-wizard.org/>

⁶⁴ <https://ds-wizard.org/machine-actionability>



Obrázek 4: Schéma integrace FIP Wizard v rámci NRP

5 Nástroje a modely pro posouzení naplnění FAIR principů

Nástroje pro hodnocení FAIR je v nejobecnějším smyslu možné rozdělit podle způsobu hodnocení do dvou kategorií – na sebehodnotící testy vyplňované uživatelem a na automatizované nástroje kontroly datových sad a metadat. Zásadní rozdíl mezi těmito přístupy spočívá v načasování: sebehodnotící nástroje lze použít v jakékoli fázi životního cyklu dat, dokonce i před jejich vznikem (např. při tvorbě DMP). Naopak automatizované nástroje vyžadují již existující, publikovaný metadatový záznam a hodnotí jeho technickou implementaci v konkrétním repozitáři. Existují i nástroje, které oba přístupy kombinují. Případně lze nástroje dělit podle způsobu vzniku FAIR politiky, vůči které je konkrétní use case porovnáván (komunitně definované, expertně definované/zajištěné autoritou instituce).

Tyto nástroje jsou užitečné pro ověření, zda jsou data skutečně dohledatelná, přístupná, interoperabilní a znovupoužitelná. Systematicky měří atributy datových sad podle objektivních kritérií, identifikují oblasti, kde data nebo jejich úložiště nesplňují požadované standardy (např. chybějící metadata, nevhodné licence, neadekvátní struktura datových sad), a poskytují doporučení k nápravě těchto nedostatků. Mnoho aspektů FAIR metrik je v praxi zajištěno podobou a nastavením metadatového popisu použitého repozitáře. Z toho plyne, že FAIR hodnotící nástroje jsou pro FAIRifikaci datových sad ukládaných v NRP/NDI pouze doplňkem k jasným pravidlům ukládání dat, nastaveným správcem repozitářů.

5.1 Sebehodnotící modely a nástroje

Metodologický rámec pro hodnocení FAIR poskytuje **FAIR Data Maturity Model (FDMM)**⁶⁵. Není to nástroj sám o sobě, ale slouží ke sjednocení popisu toho, jak dobře data splňují principy FAIR. Tento model definuje metriky a úrovně „zralosti“ dat, což umožňuje organizacím systematicky posuzovat kvalitu a připravenost jejich dat pro sdílení a opětovné využití. FDMM je koncipován jako standardizovaný přístup, který mohou využívat různé nástroje a metodiky k hodnocení FAIR úrovně dat. Model byl vyvinut organizací Research Data Alliance (RDA).

Sebeevaluační test nabízí uživateli možnost odpovídat na předem připravený set otázek s uzavřenou množinou možných odpovědí, kdy jednotlivé možnosti jsou předem expertně ohodnoceny, a to v jaké míře přispívají k naplnění principů FAIR. Touto variantou nástrojů je možné pokrývat i fáze před samotným vznikem datové sady. Následující nástroje jsou příklady sebehodnotících testů, první dva jsou určeny zejména jednotlivcům, třetí je zaměřen na správce repozitářů.

FAIR Self-Assessment Tool⁶⁶ je jednoduchý nástroj navržený pro jednotlivce a organizace, který umožňuje provést základní hodnocení, jak data splňují principy FAIR. Tento nástroj je velmi užitečný pro jednoduché a rychlé hodnocení datových sad a poskytuje první krok k lepší správě a sdílení dat podle FAIR principů. Nástroj vytvořený australským Australian Research Data Commons (ARDC)⁶⁷ je dostupný na GitHub pod otevřenou licencí. Konkrétní hodnocení zvolených odpovědí je postaveno na srovnání s interpretací „FAIR Data Principles“

⁶⁵ <https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/solution/rda-fair>

⁶⁶ <https://github.com/au-research/FAIR-Data-Assessment-Tool>

⁶⁷ <https://ardc.edu.au/>

definovaných skupinou „FORCE11“⁶⁸ – tedy komunitou vědců, knihovníků, archivářů, vydavatelů a poskytovatelů podpory. Nástroj má i český překlad⁶⁹, který zajistil kolektiv knihovny VUT.

FAIR Aware⁷⁰ byl vyvinut organizací DANS (Data Archiving and Networked Services)⁷¹ v rámci iniciativy FAIRsFAIR. Je to sebehodnotící nástroj určený pro výzkumníky, kteří chtějí lépe pochopit FAIR principy a jejich implementaci ve své práci. Uživatel prochází sérií otázek zaměřených na jeho přístup k FAIR datům a na znalosti potřebné pro jejich správné publikování. Poskytuje personalizovanou zpětnou vazbu a doporučení ke zlepšení. Slouží jako vzdělávací nástroj pro zvýšení povědomí o FAIR principech. Neprovádí přímé technické hodnocení datových sad.

Compliance Assessment Toolkit (CAT)⁷² je praktický nástroj, který umožňuje organizacím a výzkumným týmům hodnotit, nakolik jejich data splňují FAIR principy. CAT poskytuje konkrétní testy a kontrolní seznamy, které pomáhají identifikovat nedostatky v metadatech, přístupnosti a interoperabilitě dat. Kromě hodnocení nabízí také doporučení k nápravě a zlepšení datové správy. Tento nástroj byl vyvinut v rámci projektu FAIRsFAIR a je zaměřen především na praktickou aplikaci FAIR principů v datových repozitářích a výzkumných organizacích.

5.2 Automatizované hodnocení FAIR

Hodnocení FAIR principů v této kategorii probíhá automaticky, a to zejména prostřednictvím kontroly metadat, použitých ontologií a kontrolovaných slovníků. V některých případech nástroje porovnávají hodnocenou datovou sadu s komunitně definovanými schémata. Tyto nástroje pracují výhradně s již popsanými a uloženými datovými sadami, což znamená, že hodnocení je fixováno na konkrétní fázi životního cyklu dat.

5.2.1 Klíčová role repozitáře v automatizovaném hodnocení

Je zásadní si uvědomit, že výsledky automatizovaných testů neodráží primárně práci autora dat, ale technickou připravenost a nastavení samotného repozitáře. Skóre je do značné míry ovlivněno tím, jaká metadatová schémata správce repozitáře zpřístupnil a jak jsou metadata strojově dostupná. Výsledky by proto měly sloužit především jako zpětná vazba pro správce repozitářů, nikoliv jako absolutní měřítko kvality konkrétní datové sady.

5.2.2 Příklady nástrojů pro automatické hodnocení

Následující příklady nástrojů lze považovat za nejčastěji používané při automatickém hodnocení.

FAIR-Checker⁷³ - nástroj vytvořený v rámci francouzského nodu projektu ELIXIR (IFB - French Institute for Bioinformatics) je dostupný pod otevřenou licencí na portálu GitHub. Z webových stránek je extrahována anotace datové sady, která tvoří minimální znalostní graf. Tento znalostní graf je následně doplněn o data z již existujících znalostních grafů, jako jsou DataCite PID Graph, OpenAIRE Graph a Wikidata. Výsledný graf je testován, zda použité prvky odpovídají standardizovaným Linked Open Vocabularies (LOV), Ontology Lookup Service (OLS) nebo Bioportal. Nakonec je ověřena kvalita metadat vůči komunitním profilům, jako jsou Bioschemas. Výstupem je paprskový graf s procentuálně vyjádřeným skóre ve čtyřech

⁶⁸ <https://force11.org/>

⁶⁹ <https://github.com/skupaBUT/FAIR-Data-Assessment-Tool-CZ>

⁷⁰ <https://fairaware.dans.knaw.nl/>

⁷¹ <https://dans.knaw.nl/en/about/>

⁷² <https://faircore4eosc.eu/eosc-core-components/compliance-assessment-toolkit-cat>

⁷³ <https://fair-checker.france-bioinformatique.fr/>

kategoriích (F, A, I, R), tabulka bodového ohodnocení jednotlivých testů (0, 1, 2) a textová doporučení pro zlepšení v případě nižšího skóre. Důležité je, že stoprocentní FAIR compliance není podmíněna dosažením plného bodového výsledku.

F-UJI⁷⁴ - nástroj vyvinutý v rámci projektu FAIRsFAIR. Hodnotí již uložené datové sady na základě agregovaných metadat získaných z landing pages, PID služeb (např. DataCite) a dalších zdrojů, jako je re3data⁷⁵. Testuje 16 ze 17 FAIR metrik (vše kromě A2 principu – Metadata jsou přístupná, i když data již nejsou k dispozici). Výstupem je procentuální skóre FAIR compliance. Tabulka bodového hodnocení pro čtyři FAIR kategorie na základě bodového hodnocení přiřazuje každé kategorii úroveň plnění FAIR principů:

- Incomplete (neúplné)
- Initial (počáteční)
- Moderate (střední)
- Advanced (pokročilé)

F-UJI poskytuje podrobnější analytický pohled na FAIR compliance než FAIR-Checker. Je vhodný zejména pro organizace a správce repozitářů, kteří chtějí získat přehled o slabinách svého metadatového popisu a jeho strojového zpřístupnění. V [Příloze 1](#) je k dispozici hodnocení konkrétních repozitářů, které ilustruje, jak pracovat s výsledky hodnocení nástrojem F-UJI a následnými doporučeními.

FAIR Evaluator⁷⁶ - nástroj vyvinutý v rámci iniciativy GO FAIR, zaměřený na automatické hodnocení FAIR compliance pomocí sady FAIR Metrics definovaných FAIR Metrics Group. Ověřuje metadata digitálních objektů na základě jejich identifikátorů (např. DOI, URL) a provádí testy strojové čitelnosti a přístupnosti. Na rozdíl od jiných nástrojů se soustředí výhradně na automatizované testování, přičemž výsledky jsou prezentovány jako skóre pro jednotlivé FAIR kategorie. Umožňuje také manuální úpravy hodnocení a zpětnou vazbu. Tento nástroj je vhodný zejména pro výzkumníky a vývojáře, kteří chtějí systematicky hodnotit FAIR vlastnosti svých datových zdrojů.

OpenAIRE Validator⁷⁷ - vytvořen v rámci projektu OpenAIRE. Hodnotí naplňování FAIR principů v rámci jednotlivých repozitářů. Vyhodnocuje nastavení metadatových standardů podle doporučení Research Data Alliance (RDA). Na rozdíl od F-UJI a FAIR-Checker se zaměřuje na hodnocení celého repozitáře, jeho politiku a strukturu, nikoli na hodnocení jednotlivých datových záznamů. Výstupem je report o souladu repozitáře s FAIR principy a doporučení ke zlepšení metadatových standardů v repozitáři. Tento nástroj je užitečný především pro správce repozitářů, kteří chtějí vylepšit FAIR kompatibilitu své infrastruktury.

Další relevantní nástroje:

- **FAIR Assist**⁷⁸: Portál s užitečnými nástroji a průvodci pro FAIR hodnocení.
- **FAIR Implementation Framework**⁷⁹: Rámec pro implementaci FAIR principů v různých kontextech vytvořený v projektu FAIR-Impact.

⁷⁴ <https://www.f-uji.net/>

⁷⁵ <https://www.re3data.org/>

⁷⁶ <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/>

⁷⁷ https://catalogue.openaire.eu/service/openaire.metadata_validator/overview

⁷⁸ <https://fairassist.org/#!/>

⁷⁹ <https://fair-impact.eu/fair-implementation-framework>

- **FC4E-CAT-UI**⁸⁰: FAIR Compliance Assessment Toolkit – Nástroj pro podrobné hodnocení FAIR compliance organizací a jejich datových sad vytvořený v projektu FAIR-CORE4EOSC.

⁸⁰ <https://cat.argo.grnet.gr/>

6 Závěr

Implementace FAIR principů v rámci NRP/NDI představuje zásadní krok k podpoře sdílení a opětovného využívání výzkumných dat. FAIR data nejsou pouze konceptem, ale praktickým rámcem, který vyžaduje systematickou podporu ze strany správců repozitářů, poskytovatelů technických řešení a tvůrců datových politik.

Tato analýza poskytuje zásadní informace a doporučení pro všechny uživatele NDI, včetně vědců, data stewardů, správců repozitářů a dalších aktérů zapojených do správy a sdílení výzkumných dat. Slouží jako vodítko pro pochopení principů FAIR a jejich implementace v rámci NRP a NDI, což přispívá k lepší organizaci, dostupnosti a opětovnému využití dat. Pro vědce představuje návod, jak efektivně spravovat a publikovat data tak, aby byla maximálně FAIR a lépe využitelná v rámci mezinárodní vědecké spolupráce. Data stewardi mohou závěry analýzy využít k podpoře výzkumníků při správě dat a k zajištění správného popisu dat v souladu s komunitními standardy. Pro správce repozitářů pak zjištění analýzy poskytují praktická doporučení k nastavení technických a organizačních pravidel, která zajistí kompatibilitu repozitářů s mezinárodními standardy a usnadní integraci s širším výzkumným ekosystémem.

Technická architektura NRP/NDI poskytuje technické a organizační prostředí pro ukládání, správu a sdílení FAIR dat. Tato infrastruktura zahrnuje metadatový adresář NMA, národní katalog repozitářů, systémy správy perzistentních identifikátorů a dále služby, umožňující harmonizaci a vyhledávání výzkumných dat napříč disciplínami. Schémata znázorňující architekturu a klíčové vztahy v rámci NRP a NDI jsou důležitá pro pochopení, jakým způsobem jsou FAIR principy implementovány v celém ekosystému a jak mohou jednotliví aktéři efektivně spolupracovat na jejich naplnění.

Klíčovým nástrojem pro správce repozitářů je tabulka FAIR principů uvedená v této analýze. Tato tabulka poskytuje systematický přehled jednotlivých FAIR požadavků, jejich praktické aplikace v rámci NDI a zároveň představuje i související doporučení mezinárodních iniciativ CoreTrustSeal (CTS) a Cross-Domain Interoperability Framework (CDIF). Umožňuje správcům repozitářů nejen pochopit jednotlivé principy, ale také je aplikovat při nastavování vnitřních pravidel a technických požadavků pro budované repozitáře. Její využití v praxi napomůže tomu, aby výzkumná data v České republice odpovídala mezinárodním standardům a byla plně integrována do evropských a globálních infrastruktur pro otevřenou vědu.

Důležitou součástí podpory FAIR dat jsou také nástroje pro hodnocení jejich FAIRness, jako jsou automatizované testovací systémy nebo sebehodnotící modely. Tyto nástroje umožňují správcům repozitářů a uživatelům získat základní vodítka k posouzení toho, do jaké míry data odpovídají FAIR principům, a identifikovat oblasti pro zlepšení. Navzdory svým limitům jsou tyto nástroje cenné pro získání zpětné vazby. Je však klíčové, aby jejich výsledky nesloužily jako jediné kritérium kvality datové sady. Automatizované hodnocení totiž primárně odráží technické nastavení repozitáře – za které zodpovídá jeho správce – a nemůže postihnout specifický kontext dat. Kromě toho mnohdy i malá úprava může vést k významně odlišnému hodnocení, proto je s výsledky třeba nakládat nanejvýš obezřetně, jako s měkkým nástrojem pro zajištění zpětné vazby.

Zásadní službou v efektivní implementaci FAIR principů se mohou v prostředí NDI stát také FIP. Tyto profily umožňují oborově specifické nastavení FAIR principů a pomáhají komunitám definovat standardy, nástroje a osvědčené postupy odpovídající jejich specifickým potřebám. To ovšem znamená, že samotné vědecké komunity musí být přesvědčeny o významu využívání FIP a podporovat je.

Správná implementace principů popisovaných v analýze zajistí, že data uložená v NRP a dílčích repozitářích zapojených do NDI budou skutečně FAIR – snadno vyhledatelná, přístupná,

interoperabilní a znovu použitelná. To nejen podporuje efektivnější vědeckou spolupráci, ale také zvyšuje hodnotu výzkumných dat pro současné i budoucí generace vědců.

Reference

Cross-Domain Interoperability Framework (CDIF). Online. WorldFAIR, ©2022-24. Dostupné z: <https://worldfair-project.eu/cross-domain-interoperability-framework/>. [citováno 2024-10-01].

DEVARAJU, Anusuriya; HUBER, Robert; MOKRANE, Mustapha; HERTERICH, Patricia; CEPINSKAS, Linas et al. *FAIRsFAIR Data Object Assessment Metrics (0.5)*. Online. Zenodo, 2022. Licence CC BY 4.0. Dostupné z: <https://doi.org/10.5281/zenodo.6461229>. [citováno 2024-11-10].

FAIR Data Maturity Model Working Group. *FAIR Data Maturity Model: Specification and Guidelines (1.0)*. Online. Zenodo, 2020. Licence CC BY 4.0. Dostupné z: <https://zenodo.org/records/3909563#.ZEDzT-xBzRZ>. [citováno 2024-10-01].

GREGORY, Arofan; BELL, Darren; BRICKLEY, Dan; BUTTIGIEG, Pier Luigi; COX, Simon et al. *WorldFAIR (D2.3): Cross-Domain Interoperability Framework (CDIF): Report Synthesising Recommendations for Disciplines and Cross-Disciplinary Research Areas (1.0)*. Online. Zenodo, 2024. Licence CC BY 4.0. Dostupné z: <https://doi.org/10.5281/zenodo.11236871>. [citováno 2024-11-10].

WILKINSON, Mark D.; DUMONTIER, Michel; AALBERSBERG, IJsbrand Jan, APPLETON, Gabrielle; AXTON, Myles et al. *The FAIR Guiding Principles for scientific data management and stewardship*. Online. Scientific Data, vol. 3 (2016), article no. 160018. Dostupné z: <https://doi.org/10.1038/sdata.2016.18>. [citováno 2024-10-01].

Základní obsah projektu NRP – výzva Open Science I. Online. CESNET, ©2025. Dostupné z: https://www.eosc.cz/media/3713976/nrp_zakladni_obsah_projektu_2024.pdf. [citováno 2025-03-05].

Příloha 1 – hodnocení vybraných datasetů v repozitářích pomocí nástroje F-UJI

Postup hodnocení

Ze souboru repozitářů zapojených do platformy NRP byl vybrán vzorek metadatových záznamů, který byl následně analyzován pomocí nástroje F-UJI. Po zadání perzistentního identifikátoru (PID) – ve formátu DOI nebo Handle – nástroj automaticky vyhodnotil jednotlivé FAIR metriky.

F-UJI zároveň identifikoval chybějící údaje v metadatech a přiřadil každé datové sadě bodové ohodnocení, které vyjadřuje míru naplnění FAIR principů. Hodnocení vychází z metodiky FAIRsFAIR Data Object Assessment Metrics (verze 2022)⁸¹.

Výsledkem je zařazení datového záznamu do jedné ze čtyř úrovní FAIR souladu (FAIR Level): 0 – Incomplete, 1 – Initial, 2 – Moderate, 3 – Advanced.

Nástroj F-UJI neposkytuje přímá doporučení pro zlepšení FAIR úrovně, nicméně generuje podrobný výpis chybějících nebo nedostatečných prvků v metadatech, které mohou uživatelům napovědět, jakým směrem FAIRifikaci záznamu dále rozvíjet.

Na základě hodnocení jednotlivých datasetů nelze přímo vyvozovat celkovou úroveň FAIR implementace repozitáře. Výsledky spíše indikují, jakým způsobem jsou FAIR principy uplatňovány v praxi na úrovni konkrétních datových záznamů.

Pro komplexní posouzení repozitáře je nutné zohlednit také jeho technické možnosti, metadatové modely, politiky a integrační schopnosti v rámci širší infrastruktury.

Repozitář ASEP

ASEP je institucionální repozitář AV ČR⁸². Testování probíhalo v období leden až únor 2025.

Příklad hodnoceného záznamu

MACHÁČOVÁ, K. *Substantial uptake of nitrous oxide (N₂O) by shoots of mature European beech* [dataset]. Verze 2. Repozitář ASEP, 2025. Licence: CC BY 4.0. Dostupné z: <https://doi.org/10.57680/asep.0616897>

Metadatový záznam obsahuje metadata podle pravidel Repozitáře ASEP, je verzovaný, obsahuje informaci o navazujících zdrojích (publikovaný článek).

⁸¹ <https://doi.org/10.5281/zenodo.6461229>

⁸² <https://asep.lib.cas.cz/arl-cav/cs/index/>

Výsledek hodnocení

Celkový FAIR level: 2 - moderate (70%)



Komentované výsledky hodnocení

- Chybějící strojově čitelná deklarace typu zdroje:** Nástroj F-UJI očekává, že metadata budou obsahovat explicitní deklaraci, že se jedná o datovou sadu (např. pomocí schema.org/Dataset). Ačkoliv ASEP typ zdroje uvádí ('Image'), pro automatizované nástroje není tato informace dostatečně sémanticky explicitní. Doporučení: Mapovat interní typologii zdrojů na standardizované ontologie, jako je Schema.org.
- V metadatech chybí informace o formátu a o velikosti datové sady.** Informace o formátu a velikosti datové sady nejsou reprezentovány v metadatech. Při ukládání datových sad do repozitáře ASEP se doporučuje jednotlivé soubory zazipovat, zároveň by měly být v metadatech uvedeny i skutečné formáty souborů a jejich velikost.
- V metadatech chybí údaje o verzování a souvisejících publikacích.** Na landing page chybí metadata o souvisejících zdrojích a verzích datové sady – **hodnocení se liší podle použití handle nebo DOI:** datové záznamy s uloženou datovou sadou v ASEP mají přidělen handle a DOI, datové záznamy s odkazem na datovou sadu v jiném repozitáři mají přidělen handle. Rozdíl je patrný při zadání handle – hodnocení má výrazně nižší skóre. Rozdíl je způsoben tím, že na landing page chybí údaje v metadatech. Pokud nástroj hodnotí podle DOI, dohledává metadata v DataCite. Údaj o verzi datové sady chybí v metadatech na landing page i v DataCite.
- V metadatech chybí údaje o přístupnosti a podmínkách přístupu dat.** Chybí informace o přístupnosti dat (v ASEP – veřejně nepřístupný, přístupný pro ústav – omezeno na IP adresy pracoviště AV ČR, veřejně přístupný s embargem, veřejně přístupný) a podmínkách, za jakých je možné datovou sadu získat.
- V metadatech chybí informace o použitém slovníku nebo ontologii.** Repozitář ASEP je multioborový repozitář, používá číselník oborových skupin OECD – Frascati Manual 2015⁸³, který se používá se v ISVaVal.⁸⁴

⁸³ OECD (2015), Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris, <https://doi.org/10.1787/9789264239012-en>.

⁸⁴ <https://www.isvavai.cz>

LINDAT/CLARIAH-CZ

LINDAT/CLARIAH-CZ⁸⁵ je český uzel mezinárodních infrastruktur CLARIN a DARIAH.

Příklad hodnoceného záznamu

Rohacek, Jakub, 2024, Corpus from the Aozora Bunko Library, LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, <http://hdl.handle.net/11234/1-5682>.

Výsledek hodnocení

Celkový FAIR level: 2 – moderate (56%)

Summary:



Komentované výsledky hodnocení

Metadata neobsahují identifikátor dat, na která odkazují. Jelikož nástroj nenalezl identifikátor samotných dat, tak celá řada dalších testů vychází neúspěšně, a celkový bodový výsledek hodnocení je nízký. To ovšem neznamená, že požadované údaje nejsou k dispozici, pouze to, že je nástroj nebyl schopen nalézt.

Český sociálněvědní datový archiv (ČSDA)

ČSDA⁸⁶ je český uzel mezinárodní infrastruktury CESSDA a národní archiv pro data ze sociálních věd.

Příklad hodnoceného záznamu

Tabery, Paulína; Čadová, Naděžda; Vinopal, Jiří; Ďurdovič, Martin; Pilnáček, Matouš; Červenka, Jan; Tuček, Milan; Vitíková, Jana; Hájková, Helena; Plačková, Marie; Spurný, Martin; Kyselá, Monika; Weikertová, Štěpánka; Ježková, Karolína; Novotná, Lucie, 2025, "Naše společnost listopad 2024–červen (červen až srpen)", CSDA, V1⁸⁷

Výsledek hodnocení

Celkový FAIR level: 3 - advanced (83%)

⁸⁵ <https://lindat.cz/services/catalog/>

⁸⁶ <https://archivdv.soc.cas.cz>

⁸⁷ <https://doi.org/10.14473/CSDA/IUKDAJ>

Summary:



Komentované výsledky hodnocení

- Chybí URI jmenného prostoru (namespace) slovníku. ČSDA používá standardní kontrolované slovníky, zde může být chyba v metadatovém popisu.
- Chybí strojově čitelné licence pro datové sady, tato poznámka je validní, ČSDA používá vlastní licence/podmínky použití dat.
- Předchozí verze systému Dataverse vykazovala při stejném záznamu nižší FAIR skóre. Systém Dataverse byl přímo upraven v reakci na způsob automatického testování nástroje F-UJI.

Příloha č. 2 – popis uzlů schémat

Koncový uživatel

Koncový uživatel, vědec nebo data steward, spravuje vědecká data pomocí plánů datové správy (DMP) vytvořených v nástrojích jako Data Stewardship Wizard (DSW). Vytváří a spravuje metadata, ukládá datové sady a metadatové záznamy do repozitářů a hodnotí FAIR-ness dat pomocí sebehodnocení či automatizovaných nástrojů. Jako člen oborové komunity se podílí na definici komunitních standardů a FIP.

Repozitář

Repozitář je platforma pro ukládání, správu, publikaci citovatelných datových sad v souladu s FAIR principy, která zahrnuje technické, personální a procesní zajištění dlouhodobé archivace dat. Integruje FIP profily a automatizované nástroje pro hodnocení FAIR a ověřuje data a metadata během jejich ukládání. Nabízí robustní zabezpečení a kontrolu přístupu uživatelů.

Správce repozitáře

Správce repozitáře zajišťuje technické a provozní funkce repozitáře, včetně souladu s FAIR principy a požadavky komunity. Mezi jeho odpovědnosti patří správa infrastruktury, zabezpečení a dlouhodobá ochrana dat a školení uživatelů. Spolupracuje s komunitami na definování oborových FAIR standardů. Implementuje a kontroluje FIP profily aplikované v repozitáři a zavádí pravidla a nástroje pro kontrolu kompatibility datových sad se zásadami FAIR.

Datový kurátor

Datový kurátor stanovuje pravidla pro ukládání dat a správu metadat v repozitáři. Kontroluje datové sady a metadata z hlediska kvality a souladu nad rámec automatizovaných kontrol, schvaluje a zveřejňuje záznamy a harmonizuje metadata pro zajištění interoperability s dalšími systémy, zejména NMA. Podílí se také na definici oborových metadatových profilů a dohlíží na pravidla dlouhodobé archivace dat.

Oborová komunita

Oborová komunita stanovuje standardy pro správu dat a spolupracuje se správci repozitářů na definici a sladění FIP na národní i mezinárodní úrovni. Koordinuje tvorbu oborově specifických repozitářů (OSII) a podporuje vývoj metadatových a interoperabilních standardů v pracovních skupinách EOSC-CZ.

Národní repozitářová platforma (NRP)

Národní repozitářová platforma (NRP) je distribuovaný systém pro tvorbu a správu instancí repozitářů, budovaný zejména v rámci projektů NRP a IPs EOSC-CZ s podporou IPs CARDS. Podporuje vývoj FAIR-kompatibilních repozitářů poskytováním infrastruktury, nástrojů a služeb pro dlouhodobé ukládání a publikaci dat.

Národní datová infrastruktura (NDI)

Národní datová infrastruktura (NDI) zahrnuje služby, instituce a infrastruktury na národní úrovni, zapojené do správy výzkumných dat. Infrastruktura podporuje ukládání, sdílení a opětovné využití výzkumných výstupů v souladu s FAIR principy, a to i nezávisle na Národní repozitářové platformě (NRP).

Výzkumná data (datové sady)

Výzkumná data zahrnují veškeré informace vytvořené měřením, pozorováním, výpočty nebo sběrem za účelem dosažení, ověření nebo reprodukce výsledků výzkumu. Tyto datové sady tvoří základ vědecké analýzy a musí splňovat FAIR principy pro zajištění přístupnosti a opětovného využití.

Metadata (metadatový záznam)

Metadata jsou strukturovaná data popisující datové sady, a umožňují jejich vyhledávání, správu a opětovné využití. Jsou obvykle vytvářena vědci nebo data stewards, a mohou být uložena samostatně nebo spolu s popisovanými daty. Jsou klíčová pro zajištění souladu s FAIR principy.

Metadatové profily

Metadatové profily jsou strukturované soubory pravidel a standardů, které určují, jak mají být metadata spravována a organizována. Popisují metadatová pole, jejich použití, správu a interoperabilitu, formáty, standardy, obsah, vazby a hierarchie apod. Zajišťují konzistenci metadat a jejich interpretovatelnost a sladění s oborovými a mezinárodními standardy.

FIP Wizard

Nástroj pro snadnou tvorbu FAIR Implementation Profiles (FIP), které definují oborově specifické požadavky na bohaté metadatové popisy datových sad.

FIP (FAIR Implementation Profile)

FIP definují oborově specifické implementační rozhodnutí pro splnění FAIR principů. Zahrnují metadatová schémata, ontologie, sémantické slovníky a formáty potřebné pro správu a interpretaci dat. Vznikají v rámci mezinárodních aktivit, na národní úrovni především v rámci OSII.

DMP (Data Management Plan, plán správy dat)

Plán datové správy obsahuje informace o vzniku, ukládání a dostupnosti dat pro další využití. Vytváří se od počátku výzkumného projektu a pomáhá splnit požadavky na ukládání a sdílení dat v souladu s FAIR principy. DMP je možné vytvářet v nástroji Data Stewardship Wizard, kde je také možné přehledně sledovat míru naplnění FAIR principů pomocí self-assessmentu.

DSW (Data Stewardship Wizard)

DSW je nástroj pro plánování a správu dat, který usnadňuje tvorbu DMP a umožňuje automatické vytváření části metadat za účelem jejich uložení do repozitářů. Poskytuje také přehled o naplňování FAIR principů.

Self-assessment / sebehodnocení

Sebehodnocení umožňující uživatelům zkontrolovat FAIRness datové sady pomocí jednoduchého dotazníku. Hodnocení může kombinovat obecná pravidla NRP a doporučení FIP, není však omezeno pouze na FIP.

Automated assessment / automatizované hodnocení

Automatické hodnocení umožňuje kontrolu datových sad z hlediska souladu s FAIR principy pomocí konfigurovatelných nástrojů. Toto hodnocení kombinuje obecné požadavky NRP a specifické požadavky dané FIP. Výsledky hodnocení jsou generovány ve strojově čitelném formátu, například JSON, a mohou být buď privátní, nebo veřejné, v závislosti na potřebách uživatele nebo organizace. Tyto výsledky jsou verzované, což umožňuje sledování změn v čase a poskytuje přehled o pokroku v souladu s FAIR principy.

Výzkumné přístroje

Výzkumná (měřicí) zařízení a/nebo akviziční pipeline disponující vlastním perzistentním identifikátorem (PID) a umožňující odesílání dat a metadat přes API do úložiště a následně do repozitáře. Automaticky generovaná metadata mohou být obohacena o další položky koncovým uživatelem.