


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	D5.1.1
Zodpovědná osoba	Ilona Trtíková
Perzistentní identifikátor	https://doi.org/10.71495/5dv6-v116
Licence dokumentu	 https://creativecommons.org/licenses/by-sa/4.0/deed.en

Historie verzí

Verze	Datum	Popis
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	David Novák	0000-0001-6592-6291	novak@arup.cas.cz
KNAV	Tomáš Čížek		cizek@knave.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
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

Obsah

1	Úvod.....	7
1.1	Mezinárodní standardy.....	7
2	Architektura implementace FAIR v NRP/NDI.....	9
2.1	Technická infrastruktura pro FAIRifikaci.....	9
2.2	Prostředky podpory FAIR datové správy v NRP/NDI.....	10
2.2.1	Integrace metadatových standardů	10
2.2.2	Podpora správy licencí	10
2.2.3	Řízení přístupu k datům a podpora pro citlivá data.....	10
2.2.4	Metadata z přístrojů	11
2.3	Znázornění implementace FAIR v prostředí NRP/NDI.....	11
2.3.1	Přehledové konceptuální schéma implementace FAIR v NRP/NDI	11
2.3.2	Schéma klíčových vztahů z pohledu repozitáře	13
3	FAIR Principy a jejich výklad v kontextu NRP/NDI.....	14
3.1	Tabulka implementace FAIR principů v NRP	15
4	Služby podporující implementaci FAIR v rámci oborů.....	33
4.1	FAIR Implementation Profiles.....	33
4.2	FIP Wizard	33
4.3	Integrované řešení pro plánování správy dat s podporou exekuce.....	34
5	Nástroje a modely pro posouzení naplnění FAIR principů	36
5.1	Sebehodnotící modely a nástroje	36

5.2	Automatizované hodnocení FAIR.....	37
6	Závěr.....	40
	Příloha 1 – hodnocení dat v repozitářích nástrojem F-UJI.....	43
	Repozitář ASEP.....	43
	LINDAT/CLARIAH-CZ.....	44
	Český sociálněvědní datový archiv (ČSDA).....	45
	Příloha č. 2 – popis uzlů schémat.....	47

Shrnutí

Tato analýza se zaměřuje na implementaci FAIR principů (Findable, Accessible, Interoperable, Reusable)¹ v rámci Národní repozitářové platformy (NRP)² a Národní datové infrastruktury (NDI). Cílem je poskytnout správcům repozitářů, vědcům, data stewardům a dalším uživatelům přehled klíčových požadavků a doporučení, která umožní efektivní správu výzkumných dat v souladu s mezinárodními standardy a nejlepší praxí.

Hlavním tématem, které analýza řeší, je zajištění správného nastavení repozitářů a souvisejících služeb tak, aby data byla FAIR. Naplnění FAIR principů totiž nespočívá jen na uživateli nahrávajících datové soubory, ale především v infrastruktuře a pravidlech repozitářů. Analýza proto poskytuje přehled konkrétních požadavků na NRP/NDI a jednotlivé repozitáře a uvádí jejich souvislosti s mezinárodními certifikacemi, jako je CoreTrustSeal (CTS)³, a s doporučeními Cross-Domain Interoperability Framework (CDIF)⁴.

Dokument popisuje architekturu NRP/NDI, nástroje na hodnocení FAIR dat a význam FAIR Implementation Profiles (FIP)⁵, které pomáhají oborově specifické implementaci FAIR. Součástí je několik dílčích výstupů pracovní skupiny KA5.1, vypracovaná schémata a tabulka FAIR principů, na nichž se podíleli všichni její členové.

Hlavním výstupem analýzy je systematický přehled požadavků na FAIR data a doporučených postupů pro správce repozitářů, který usnadní jejich aplikaci v praxi. Správná implementace těchto zásad je nezbytnou podmínkou pro to, aby výzkumná data v České republice byla nejen v souladu s FAIR principy, ale také snadněji využitelná v širším mezinárodním kontextu, čímž podpoří otevřenou vědu a efektivní sdílení dat.

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

² <https://www.eosc.cz/projekty/narodni-repozitarova-platforma-pro-vyzkumna-data-os-i-nrp/nrp>

³ <https://www.coretrustseal.org/>

⁴ <https://worldfair-project.eu/cross-domain-interoperability-framework/>

⁵ <https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/>

1 Úvod

V rámci implementace iniciativy EOSC⁶ 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)**⁷. Platforma má podpořit uchovávání dat, která budou snadno vyhledatelná, přístupná a opakovaně použitelná, tedy v souladu s principy FAIR⁸. 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. V souladu s mezinárodně uznávanou dobrou praxí vzniká v rámci NRP a NDI řada služeb a systémů k dosažení těchto cílů.

Naplnění FAIR principů nespočívá pouze na koncových uživateliích nahrávajících datové soubory, ale především v nastavení repozitářů, které tato data uchovávají. Proto se analýza zaměřuje zejména na správce repozitářů, kteří mají klíčovou roli při nastavování technických a organizačních opatření zajišťujících kompatibilitu s FAIR zásadami. Dokument je však určen nejen pro ně, ale také pro vědce, data stewardy a další uživatele služeb NRP/NDI, kteří se podílejí na správě a využívání výzkumných dat.

V rámci NRP i NDI jsou nebo budou nabízeny služby, které podporují naplňování FAIR principů. V následujícím textu jsou mapovány nároky FAIR principů a uváděny do celého kontextu prostředí NRD/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 uvedeny požadavky na NRP, které ze standardů vyplývají.

Dokument zároveň představuje plánované aktivity v rámci NRP, které jsou v současnosti předmětem implementace a mají pomoci uživatelům NRP/NDI systematicky plnit požadavky 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. Celkově 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

Sladění pravidel aplikovaných v českém prostředí s mezinárodním kontextem rozšiřuje možnosti využívání výzkumných FAIR dat a jejich sdílení vědeckou komunitou. Při analýze a v budoucí implementaci jsou proto zohledňovány výstupy řady iniciativ. Mezi východiska tak patří zejména iniciativy:

GO FAIR⁹ – mezinárodní iniciativa zaměřená na podporu principů FAIR. Iniciativa byla založena v roce 2017 jako společná aktivita Nizozemska, Německa a Francie s cílem

⁶ <https://www.eosc.cz/>

⁷ https://opjak.cz/vyzvy/vyzva-c-02_23_014-open-science-i/

⁸ Principy FAIR <http://www.go-fair.org/fair-principles/>

⁹ <https://www.go-fair.org/go-fair-initiative/>

podporovat otevřenou vědu a implementaci FAIR principů. Koordinaci zajišťuje GO FAIR International Support and Coordination Office (GFISCO), které sídlí v Göttingenu (Německo). V současnosti Evropská komise podporuje GO FAIR v rámci iniciativy EOSC. Na tvorbě standardů a vývoji FAIR nástrojů spolupracuje vědecká komunita, univerzity i výzkumné ústavy.

FAIRsFAIR (Fostering FAIR Data Practices in Europe)¹⁰ - 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)¹¹ - projekt cílí na zlepšení interoperability a opětovné použitelnosti FAIR dat napříč různými vědními obory a institucemi. CDIF poskytuje konkrétní doporučení a standardy pro naplňování FAIR principů v několika oblastech. 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.

WorldFAIR¹² - 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)¹³ ve spolupráci s RDA (Research Data Alliance)¹⁴.

¹⁰ <https://doi.org/10.3030/831558>

¹¹ <https://worldfair-project.eu/cross-domain-interoperability-framework/>

¹² <https://worldfair-project.eu/the-project/>

¹³ <https://codata.org/>

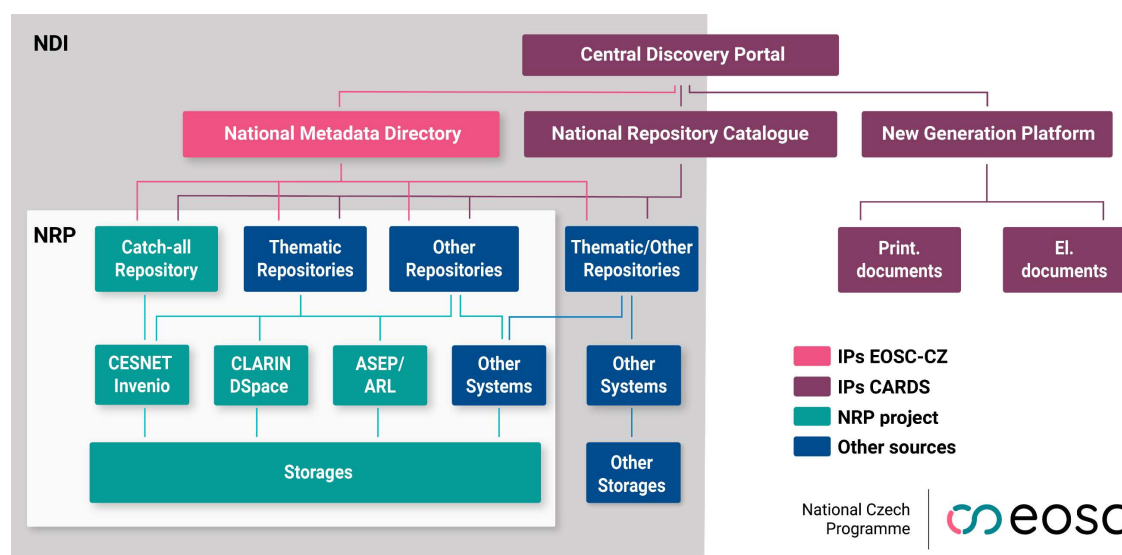
¹⁴ <https://www.rd-alliance.org/>

2 Architektura implementace FAIR v NRP/NDI

Následující kapitola představuje technickou infrastrukturu NRP/NDI se zaměřením na FAIR principy. Rovněž vysvětluje schémata, která přehledně zobrazují vztahy mezi jednotlivými uživateli NDI a dostupnými či plánovanými službami a nástroji.

2.1 Technická infrastruktura pro FAIRifikaci

Architektura implementace FAIR v rámci EOSC-CZ je postavena na infrastruktuře NRP/NDI. Klíčovými komponentami jsou federované repozitáře, které umožňují uložení a vyhledávání výzkumných dat s podporou persistentních identifikátorů (PID) a standardizovaných metadatových schémat. Na schématu (obr. 1) je přehledně ukázáno, z jakých částí se skládá plánované prostředí, a z kterých projektů jsou tyto části financovány. (Základní obsah projektu NRP – výzva Open Science I, @2025)



Obr. 1 - Ekosystém služeb e-infrastruktury

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)¹⁵ bude centrálním bodem pro ukládání a vyhledávání metadat o výzkumných datech nashromážděných v rámci projektů řešených na českých výzkumných institucích, z různých zdrojů v České republice i zahraničí. Hlavním cílem je zajistit lepší dostupnost dat a umožnit efektivní vyhledávání výzkumných datových sad napříč disciplínami. Metadata ze zapojených repozitářů budou do NMA sklížena automaticky pomocí protokolu OAI-PMH, přičemž bude uplatněn jednotný metadatový standard, který zajistí interoperabilitu mezi různými systémy.

Národní katalog repozitářů (NKR)

¹⁵ NMA - <https://nma.eosc.cz/>

NKR (National Repository Catalogue) bude poskytovat přehled o jednotlivých repozitářích zapojených do NRP/NDI. Kromě základních informací o repozitářích bude obsahovat podrobnosti o použitých řízených slovnících, ontologiích a metadatových schématech, 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 PID bude hrát klíčovou roli při správě a podpoře využívání perzistentních identifikátorů (PID) v České republice. Díky účasti v konsorciu DataCite budou mít české instituce možnost přidělovat DOI a další podporované PID, což usnadní citování, sledování a dlouhodobou dostupnost dat.

2.2 Prostředky podpory FAIR datové správy v NRP/NDI

Následující text představuje výčet oblastí, které jsou řešeny pomocí vyvíjených nástrojů, případně napomohou při FAIRifikaci dat uživatelům a správcům repozitářů v rámci technické infrastruktury NRP/NDI.

2.2.1 Integrace metadatových standardů

Jedním z hlavních pilířů NRP/NDI bude zajištění automatizované agregace metadat ze všech zapojených repozitářů, což podpoří vyhledatelnost, interoperabilitu a opětovné využití dat. Klíčovým prvkem této integrace bude základní metadatový model, který zajistí jednotné jádro pro strukturování metadat.

Správci repozitářů a odborné komunity budou mít k dispozici Dataspecer¹⁶, nástroj pro správu metadatových profilů. Nástroj umožňuje efektivní správu metadatových schémat (obecně datových specifikací) určených pro popis dat v jednotlivých repozitářích. To zajistí ucelený přehled o dostupných schématech v jednotném formátu, umožní reflektovat změny dílčích schémat, dále rozšiřovat konkrétní schémata pro podobory či specifická oborová zaměření, ale také snadno mapovat schémata mezi sebou za účelem zvýšení interoperability.

2.2.2 Podpora správy licencí

NRP poskytne nástroje a metodickou podporu pro výběr a přidělování licencí k datovým sadám. To pomůže vědcům a správcům dat efektivně nastavit podmínky pro přístup, sdílení a využití výzkumných dat. Kromě toho bude vytvořen interaktivní průvodce, který uživatele provede výběrem vhodné licence. Tato služba bude rovněž integrována do repozitářových platforem, což umožní automatizované nastavení licencí při publikaci datových souborů.

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 granulární nastavení oprávnění na úrovni datových souborů nebo jednotlivých položek.

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.

¹⁶ Dataspecer - <https://dataspecer.com/>

2.2.4 Metadata z přístrojů

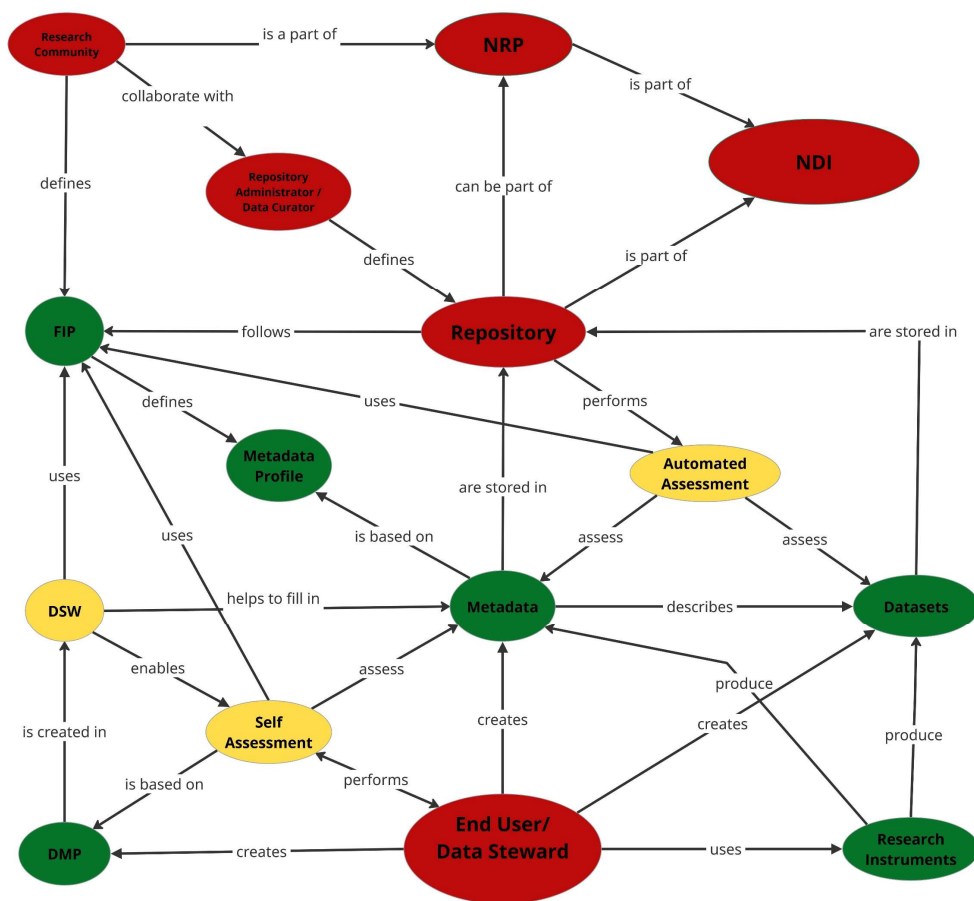
Součástí budoucího rozvoje infrastruktury NRP/NDI je také podpora automatizovaného sběru dat a metadat z laboratorních a jiných výzkumných přístrojů. To umožní snadnější správu a popis experimentálních dat ihned v okamžiku jejich vzniku. Metadata extrahovaná z přístrojů mohou být následně automaticky synchronizována s repozitářovými platformami a katalogy, což zjednoduší procesy FAIR správy dat.

2.3 Znárodnění implementace FAIR v prostředí NRP/NDI

Výzkumníci, data stewardi a správci repozitářů musí spolupracovat na zajištění kvality a interoperability dat, přičemž využívají specializované nástroje, standardy a automatizované procesy. Tato kapitola nabízí ve formě dvojice schémat koncepční přehled předpokládaného způsobu ukládání vědeckých dat v ekosystému NRP/NDI s důrazem na FAIR principy. Klíčové vztahy mezi aktéry v rámci NDI jsou představeny v **Přehledovém konceptuálním schématu** ([obr. 2](#)), které ukazuje propojení mezi uživateli, datovými výstupy, jejich metadaty, dílčími nástroji NDI a odbornými komunitami. Naproti tomu **Schéma klíčových vztahů z pohledu repozitáře** ([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 považujeme za zásadní pro ujasnění, v jakých fázích workflow a v jakých částech datové infrastruktury je potřebné zohledňovat doporučení pro zvýšení FAIRness ukládaných datových souborů. V Příloze je uveden podrobný popis jednotlivých uzlů ve schématech (viz [Příloha 2](#)). Ve schématech jsou entity vystupující v pozici aktérů vyznačeny červeně. Služby/nástroje žlutě, zelené je pak označení výstupů a produktů těchto aktérů a jejich vztahů.

2.3.1 Přehledové konceptuální schéma implementace FAIR v NRP/NDI

Výzkumník nebo data steward (**End User/Data steward**) vytváří a spravuje výzkumná data (**Datasets**) a jejich metadatový popis (**Metadata**), která ukládá do repozitáře. V některých oborech mohou být výzkumná data a metadata generována výzkumnými přístroji (**Research Instruments**), jež zajišťují automatizovaný přenos 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.



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

Pro efektivní správu dat využívá koncový uživatel plán datové správy (**DMP**), který je v ideálním případě vytvořen v odpovídajícím nástroji umožňujícím strojové zpracování, jako je například Data Stewardship Wizard (**DSW**)¹⁷. 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 podílí 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 souborů (**Metadata Profile**). Tyto profily zajišťují soulad s FAIR principy a definují vhodné ontologie, sémantické slovníky, formátové standardy a další kritéria nezbytná pro správnou správu a interpretaci dat. Repozitář (**Repository**) slouží k ukládání, správě a publikaci citovatelných digitálních objektů v souladu s FAIR principy. Kromě toho zajišťuje technické, personální a procesní podmínky pro dlouhodobou archivaci dat.

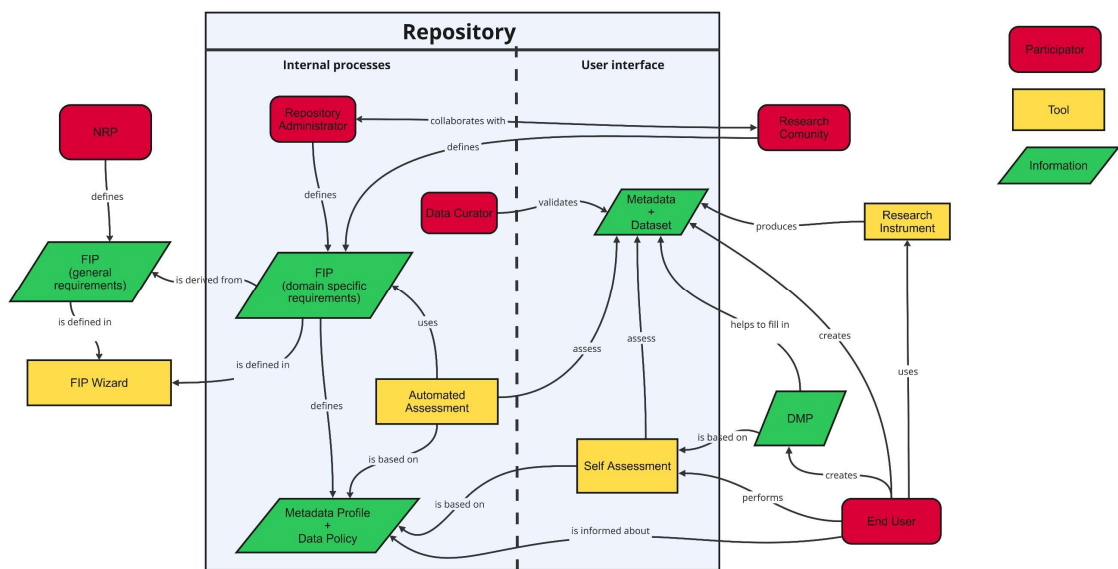
Administrátor repozitáře (**Repository Administrator**) zajišťuje technické a provozní aspekty repozitáře. Integruje požadavky odborné komunity na využití FIP, čímž naplňuje odpovídající komunitní standardy. Uplatnění FIP a DMP podporuje nasazení automatizovaných nástrojů

¹⁷ <https://ds-wizard.org/>

pro hodnocení FAIRness datových souborů (**Automated Assessment**) v repozitářích. Administrátor repozitáře úzce spolupracuje s datovým kurátorem (**Data Curator**), který stanovuje pravidla pro ukládání dat a správu metadat. Kurátor rovněž provádí kontrolu kvality datových souborů 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**).

2.3.2 Schéma klíčových vztahů z pohledu repozitáře

V tomto schématu vztahů existujících uvnitř a vně repozitáře jsou zobrazeny hlavní prvky, které mají vliv na naplňování FAIR principů, a které jsou definovány repozitářem. Uživatel (**End User**) vyplňuje metadata a nahrává dataset (**Metadata, Dataset**), vždy je ale ve svých rozhodnutích omezen administrátorem konkrétního repozitáře (**Repository Administrator**), zvoleným metadatovým profilem, politikou repozitáře (**Metadata Profile, Data Policy**) a oborovým FAIR implementation profile (**FIP**). V praktické rovině je práce uživatele řízena automatickými kontrolami polí a hodnot, které je mu umožněno vyplnit, již přednastavenými systémy ontologií a řízených slovníků, povinnými metadatovými poli (**Metadata Profile, Data Policy**) a také nastavením kurátorského procesu prováděného datovým kurátorem (**Data Curator**). Na úrovni uživatelského rozhraní repozitáře by měl být dostupný jednoduchý sebehodnotící nástroj (**Self Assessment**) a automatizované hodnocení (**Automated Assessment**) pro snadnou indikaci ne/souladu nahrávaného datasetu s pravidly a parametry repozitáře. Repozitářem deklarovaná pravidla a podmínky by měly být otevřeně dostupné ve strojově čitelném formátu (skrže FIP nebo jiným vhodným způsobem) umožňujícím přenos do dalších systémů pomocí API (např. pro účely tvorby **DMP**).



Obrázek 3: Schéma klíčových vztahů z pohledu repozitáře

3 FAIR Principy a jejich výklad v kontextu NRP/NDI

Následující tabulka poskytuje systematický přehled FAIR principů, jak jsou interpretovány a aplikovány v rámci NRP a NDI. Tabulka vychází z oficiálního dokumentu FAIRsFAIR Data Object Assessment Metrics (v0.5) (Devaraju et al., 2022)¹⁸ a dále jsou zohledněny i závěry z 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.

Struktura tabulky a vysvětlení jednotlivých sloupců

Sloupec	Vysvětlení
Princip (CZ)	Český název daného FAIR principu, který shrnuje jeho klíčový 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 NRP, které musí být splněny pro 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 a řízení datové infrastruktury a podpora FAIR principů na úrovni repozitářů. CTS doplňuje FAIR principy tím, že se nezaměřuje pouze na technické vlastnosti dat, ale také na správu, udržitelnost a důvěryhodnost repozitářů, v nichž jsou data uložena. Sleduje tak technické a organizační podmínky, které jsou nutné pro naplnění FAIR správy dat. 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 CoreTrustSeal¹⁹.
- Doporučení Cross-Domain Interoperability Framework (CDIF)²⁰ 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. V tabulce jsou jednotlivým FAIR principům přiřazena relevantní doporučení CDIF, která pomáhají zajistit, že metadata a data budou interoperabilní v širším výzkumném ekosystému.

¹⁸ <https://zenodo.org/records/6461229>

¹⁹ Aktuální požadavky CoreTrustSeal (2023–2025) <https://www.coretrustseal.org/why-certification/requirements/>

²⁰ <https://worldfair-project.eu/cross-domain-interoperability-framework/>

Další informace o CDIF a jeho implementaci v rámci EOSC jsou dostupné na webu EOSC²¹ případně v souhrnném přehledu CDIF doporučení²². (Cross-Domain Interoperability Framework, @2022-25; GREGORY, Arofan, 2024)

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 a v repozitářích a k identifikaci oblastí pro zlepšení, podpoře 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, 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.

3.1 Tabulka implementace FAIR principů v NRP

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 každý datový objekt a jeho metadata mají být jednoznačně identifikovatelné prostřednictvím globálně jedinečného a trvalého identifikátoru (PID). 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. • Handle System – alternativní 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 (URL a URN).</p>
Doporučení CDIF	Relevantní CDIF doporučení: žádná
Požadavky CTS	CoreTrustSeal R12 vyžaduje, aby repozitáře umožnily uživatelům vyhledávat digitální objekty 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/>

²² <https://worldfair-project.eu/cross-domain-interoperability-framework/>

Naplnění v NRP	Každá datová sada uložená v NRP bude mít přidělený perzistentní identifikátor (např. DOI nebo Handle). Metadata budou obsahovat jednoznačný odkaz na PID, který umožní jeho použití pro vyhledávání a přístup k datům. Přiřazování PID bude probíhat prostřednictvím Národního konsorcia DOI vedeného NTK (přidělován prostřednictvím registrační agentury DataCite) nebo alternativních služeb podporovaných v rámci NRP.
Požadavky na NRP	<ul style="list-style-type: none"> • Zajištění PID pro všechny datové sady: Každý dataset v NRP musí mít přiřazený perzistentní identifikátor, který bude součástí metadat. • Podpora různých typů PID: Zajistit kompatibilitu s běžně používanými systémy PID (DOI, Handle, ARK) a umožnit jejich mapování na mezinárodní identifikační schémata. • Integrace PID do metadatového schématu: PID musí být standardní součástí metadatového záznamu v NMA a být propojen s odpovídajícími repozitářovými záznamy. • Zajištění dlouhodobé dostupnosti PID: PID musí být trvale dostupný a spravovaný tak, aby nedocházelo k jeho zneplatnění nebo ztrátě.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Přiřazování PID při ukládání datových sad: Repozitář musí zajistit, že každé nové datové sadě bude přidělen perzistentní identifikátor, buď prostřednictvím DOI, nebo jiného akceptovaného systému PID. Repozitáře spravované v rámci NRP musí každému záznamu před jeho publikováním přiřadit PID pomocí služeb poskytovaných NRP. Pokud je přiřazován DOI, je možné využít služeb Národního konsorcia DOI. Repozitáře mimo platformu NRP, ale zapojené do NDI, musí rovněž zajistit přiřazení PID, a to buď prostřednictvím služeb NRP, nebo poskytovatelů třetích stran, kteří podporují všeobecně uznávaná schémata PID. • Zajištění správného propojení PID s metadaty: PID musí být vždy součástí metadatového záznamu a propojen s datovým objektem. • Podpora správy a aktualizace PID: Repozitáře musí zajistit, že PID budou spravovány tak, aby odpovídaly aktuálnímu umístění a stavu datových sad.
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.</p> <p>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, jejich strukturu, vztahy k jiným data setům a podmínky použití. Mezi klíčové prvky metadat patří:</p>

	<ul style="list-style-type: none"> • Popisná metadata – název, autor, datum vzniku, přidělený perzistentní 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>Minimální požadavky na metadata pro zařazení do Národního adresáře metadat (NMA) jsou primárně definovány Pracovní skupinou EOSC-CZ Metadata. V NRP jsou vyvíjeny související nástroje pro správu metadatových profilů. Kromě toho se metadata zabývá několik projektů napříč ekosystémem EOSC-CZ. Klíčovou iniciativou je vytvoření NMA v rámci projektu IPs CARDS. NMA bude fungovat jako agregátor a bude využívat jednotné schéma pro shromažďování vybraných metadat. Potřeba bohatých metadat bude řešena v rámci projektu OSII, kde budou vytvořeny oborově specifické metadatové profily, které umožní lepší vyhledávání, interoperabilitu a opakované použití dat tak, aby vyhovovala potřebám konkrétních oborů.</p> <p>Pro NRP byly definovány minimální požadavky pomocí Core metadata schématu NMA s tím, že jde o implementaci schématu metadat DataCite doplněnou o prvky z DCAT-AP²³ a celkově navázanou na tuto obecnější ontologii. Doménově agnostické vlastnosti schématu DataCite umožňují dostatečně flexibilní a konzistentní identifikaci dat pro účely citování a vyhledávání. 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.</p>
Požadavky na NRP	<ul style="list-style-type: none"> • Definování sady minimálních požadavků na metadata. • 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 používaný při agregaci do NMA.

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

	<ul style="list-style-type: none"> Podpora oborově specifických metadat: Umožnit vědeckým komunitám definovat oborové standardy metadat v rámci FIP jako součást implementace repozitáře.
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 atributů relevantních pro konkrétní vědeckou komunitu.
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. To může být buďto PID datové sady, anebo, v případě že PID odkazuje na landing page s metadaty, vede pak odtud URL odkaz na datový objekt. K propojení metadat a dat je pak potřeba využít signposting ²⁴ – metodu využívající standardní HTTP hlavičky, které umožňují automatické a strojově čitelné navigování mezi daty, metadaty a dalšími souvisejícími informacemi.
Doporučení CDIF	Všeobecné doporučení CDIF stanovuje, že metadata by měla být uložena v JSON-LD ²⁵ 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 srozumitelně a trvale propojena s jejich metadaty.
Naplnění v NRP	NRP bude zajišťovat podporu PID, které mohou být přiděleny jak metadatům, tak přímo datovým objektům. NRP bude po repozitářích požadovat jasnou metodiku tvorby odkazu na datový objekt a standardní signposting umožňující strojovou zpracovatelnost.
Požadavky na NRP	<ul style="list-style-type: none"> Zajištění jednoznačného propojení metadat a datových sad: Je potřeba definovat, jak má být úložiště dat z metadat odkazováno, zda odkazem na datovou sadu prostřednictvím adresy URL s využitím signpostingu, nebo jiným způsobem. Podpora standardizovaných identifikačních polí: Metadata musí obsahovat jedinečné identifikátory distribucí dat (datových souborů), které umožňují jednoznačnou identifikaci a zpřístupnění konkrétní distribuce.
Požadavky na repozitáře	<ul style="list-style-type: none"> Zahrnutí identifikátoru dat do metadat: Repozitáře musí definovat, jakým způsobem jsou propojena metadata s datovými objekty, které popisují, a kontrolovat, že záznamy v repozitáři tyto požadavky splňují. Implementace signpostingu: Repozitáře by měly zajistit, že v na landing page datové sady je k dispozici signposting pro jednoznačné a strojově čitelné propojení metadat s daty pomocí URL.

²⁴ Metoda zpřístupnění metadat a souvisejících zdrojů pomocí HTTP hlaviček a odkazů <https://signposting.org/>

²⁵ <https://json-ld.org/>

Požadavky na FIP	<ul style="list-style-type: none"> • Specifikace, jakým způsobem má být identifikátor dat zahrnut do metadat: FIP by měl v rámci metadat definovat standardizované pole pro identifikátor dat. • Doporučení pro podporu signpostingu: FIP by měl uvádět doporučený způsob implementace signpostingu v daném oboru.
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, Google Dataset Search). • Oborových vyhledávačích (např. PubMed, Europe PMC, PANGAEA) • Obecných vyhledávacích službách (např. Google, Bing).
Doporučení CDIF	Discovery profil zdůrazňuje potřebu registrace metadat v systémech vyhledávání, jako jsou Schema.org a DCAT. Tímto způsobem se metadata stávají dohledatelnými jak pro výzkumníky, tak pro automatizované indexační nástroje, což zvyšuje jejich dostupnost.
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 všech uložených datových sad budou automaticky registrována v Národním metadatovém adresáři (NMA), čímž se zajistí centralizovaný přístup a vyhledatelnost v rámci ekosystému EOSC-CZ. Metadata budou dostupná prostřednictvím OAI-PMH nebo jiných protokolů pro harvesting, což umožní jejich indexaci v OpenAIRE, DataCite a dalších mezinárodních službách.
Požadavky na NRP	<ul style="list-style-type: none"> • Automatická indexace metadat: Metadata musí být automaticky indexována v NMA a zpřístupněna pro vyhledávání. • Podpora protokolů pro harvesting: NRP musí podporovat OAI-PMH, Schema.org, DCAT-AP a další relevantní standardy pro indexaci. • Registrace v mezinárodních agregátorech: Metadata by měla být pravidelně synchronizována s mezinárodními databázemi (OpenAIRE, DataCite). • 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"> • Zajištění mapování a indexování v NMA: Metadata repozitáře musí být ve strojově čitelném a standardizovaném formátu a musí být mapovatelná na minimální standard indexovaný v NMA. • Zajištění dostupnosti metadat pro indexaci: Základní i podrobná oborová metadata musí být ve formátech umožňujících efektivní automatickou indexaci vyhledávači a v souladu s odpovídajícími standardy (např. Dublin Core, DataCite, Schema.org). • Podpora exportu metadat: Repozitáře musí umožnit přístup k metadatům prostřednictvím OAI-PMH²⁶, případně též REST API. • Validace metadat před registrací: Metadata by měla být pravidelně kontrolována na správnost a úplnost, aby byla dobře indexovatelná.
Požadavky na FIP	<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 ²⁷) za použití standardizovaného komunikačního protokolu, který je široce používaný a podporovaný. 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"> • Definice podporovaných protokolů: Zajistit podporu běžně používaných protokolů, jako jsou HTTPS, OAI-PMH, případně rozšířit o protokoly SFTP nebo SPARQL pro specifické případy. • Uvedení odkazu na API: S využitím signpostingu do metadat uvést odkaz na API sloužící pro indexaci metadat datové sady pomocí konkrétního

²⁶ <https://www.openarchives.org/pmh/>

²⁷ <https://www.handle.net/>

	protokolu. Adresu použitého API uvést též v metadatech indexovaných v NMA.
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. • Uvedení odkazu na API: Komunikační API endpoint musí být dostupný v obdobném rozsahu jako pro repozitáře NRP.
Požadavky na FIP	<ul style="list-style-type: none"> • Definování standardů pro přístup: FIP musí obsahovat požadavky na podporované protokoly. • Oborově specifické protokoly: Pokud určité obory vyžadují specifické protokoly, měly by být jasně uvedeny v příslušných FIP.
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. Otevřené protokoly podporují interoperabilitu a dlouhodobou dostupnost dat, čímž naplňují základní požadavky FAIR principů. 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. Podpora těchto protokolů zajišťuje, že data a metadata uložená v NRP mohou být přístupná i v budoucnosti bez závislosti na proprietárních technologiích.
Požadavky na NRP	<ul style="list-style-type: none"> • Implementace otevřených protokolů: Zajistit, že všechny komunikační protokoly používané v rámci NRP jsou otevřené, bezplatné a univerzálně podporované. • Podpora více protokolů: Zajistit kompatibilitu s různými protokoly, zejména HTTPS a OAI-PMH, a dále např. SFTP²⁸ či SPARQL²⁹, podle potřeb jednotlivých oborů.
Požadavky na repozitáře	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • Definice otevřených protokolů pro oborové využití: Každý FIP by měl specifikovat, které protokoly jsou doporučené pro daný obor.

²⁸ <https://www.ssh.com/academy/ssh/sftp-ssh-file-transfer-protocol>

²⁹ <https://www.w3.org/TR/sparql12-query/>

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 ³⁰ , OAuth, SAML ³¹ nebo LDAP ³² . Při implementaci těchto protokolů je důležité vyvážit potřebu zabezpečení a snadnost přístupu pro koncové uživatele.
Doporučení CDIF	Relevantní CDIF doporučení: žádná
Požadavky CoreTrustSeal	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 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 bezpečnostních 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, chráněná data, citlivá data).
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 nebo právním omezením.
Princip	A2. metadata are accessible, even when the data are no longer available

³⁰ HYPERLINK "<https://www.rfc-editor.org/info/rfc7617>"<https://www.rfc-editor.org/info/rfc7617>

³¹ <https://www.oasis-open.org/standard/saml/>

³² <https://ldap.com/>

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 od Datacite ³³) 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	Minimální záznam metadat bude uložen v Národním metadatovém adresáři (NMA), který zajistí jejich dlouhodobou dostupnost, i když samotná data již nebudou k dispozici. NRP vyžaduje, že metadata musí být veřejně přístupná a vždy propojena s perzistentním identifikátorem (PID), který povede na odpovídající záznam v repozitáři nebo na tombstone page.
Požadavky na NRP	<ul style="list-style-type: none"> NRP musí zajistit, aby byla metadata vždy zveřejněna a uchována jak v rámci úložiště, tak 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 objektů. V rámci úložiště 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 být vždy 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

³³ <https://support.datacite.org/docs/tombstone-pages>

Popis	<p>Tato princip požaduje, aby metadata i data byla reprezentována formálním, přístupným, sdíleným a široce použitelným jazykem, což umožní jejich snadnou integraci a zpracování napříč různými systémy a obory.</p> <p>Mezi tyto jazyky patří:</p> <ul style="list-style-type: none"> • RDF³⁴ (Resource Description Framework) – standardní model pro reprezentaci propojených dat. • OWL³⁵ (Web Ontology Language) – jazyk pro tvorbu ontologií, používaný v sémantickém webu. • JSON-LD³⁶ (JavaScript Object Notation for Linked Data) – zjednodušený formát pro reprezentaci sémantických dat, kompatibilní s webovými technologiemi.
Doporučení CDIF	<p>Controlled Vocabularies profil doporučuje SKOS³⁷ 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.</p> <p>Universals profil definuje standardy pro popis univerzálních konceptů, jako je čas, geografie a měrné jednotky, což je klíčové pro sjednocení různorodých datových struktur.</p>
Požadavky CTS	<p>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 interpreto-vatelná a znovu použitelná v průběhu času. R15 pak požaduje, aby repozitáře měly adekvátní technickou infrastrukturu, která umožní správné zpracování metadat v různých systémech.</p>
Naplnění v NRP	<p>Metadata musí být serializována v strojově čitelných formátech, jako je RDF, JSON-LD nebo XML³⁸.</p>
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, JSON-LD). • Použití strojově čitelných formátů: Metadata musí být dostupná v datových formátech umožňujících správnou serializaci a přenos metadat (např. JSON-LD, RDF/XML, Turtle³⁹, XML).
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, JSON-LD). • 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	<ul style="list-style-type: none"> • Definování vhodného jazyka pro reprezentaci znalostí: FIP by měl doporučit formáty a standardy nejvhodnější pro daný obor (např. RDF, JSON-LD).

³⁴ <https://www.w3.org/RDF/>

³⁵ <https://www.w3.org/OWL/>

³⁶ <https://json-ld.org/>

³⁷ <https://www.w3.org/2004/02/skos/>

³⁸ <https://www.w3.org/TR/REC-xml/>

³⁹ Terse RDF Triple Language - <https://www.w3.org/TR/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 používána 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)⁴⁰, organizace (ROR)⁴¹ a koncepty (Wikidata⁴², ISNI). Tyto registry umožňují lepší propojení metadat, zvýšení jejich dohledatelnosti a sémantickou interoperabilitu. Katalog FAIR kompatibilních slovníků a standardů nabízí stránky FAIRsharing.org.</p> <p>Příklady FAIR slovníků:</p> <ul style="list-style-type: none"> • Dublin Core⁴³ (obecná metadata) • Schema.org⁴⁴ (webová metadata) • Gene Ontology⁴⁵ (GO) (biomedicína) • MeSH⁴⁶ (Medical Subject Headings) (medicína) • Wikidata⁴⁷ (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 ⁴⁸ (rozšíření standardu SKOS ⁴⁹ o funkcionality potřebné pro statistická a ekonomická data). Tímto bude zajištěno, že metadata budou konzistentně strukturována a snadno interpretovatelná v různých vědeckých disciplínách.

⁴⁰ <https://info.orcid.org/what-is-orcid/>

⁴¹ <https://ror.org/>

⁴² https://www.wikidata.org/wiki/Wikidata:Main_Page

⁴³ Dublin Core Metadata Initiative - metadata terms <https://www.dublincore.org/>

⁴⁴ Schemas for structured data on the Internet - <https://www.w3.org/community/schemaorg/>

⁴⁵ <https://geneontology.org/>

⁴⁶ <https://www.nlm.nih.gov/mesh/meshhome.html>

⁴⁷ https://www.wikidata.org/wiki/Wikidata:Main_Page

⁴⁸ https://ddialliance.org/xkos_v1.0

⁴⁹ <https://www.w3.org/2004/02/skos/>

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.
Naplnění v NRP	V rámci EOSC-CZ na vývoj a správu slovníků dohlíží projekt IP CARDS ⁵⁰ , konkrétně prostřednictvím Centra repozitářů a správy metadat, které bude sloužit jako nástroj pro evidenci repozitářů, metadatových schémat a řízených slovníků používaných v NDI a potažmo v NRP.
Požadavky na NRP	<ul style="list-style-type: none"> • Prosazování FAIR slovníků a ontologií: NRP by mělo podporovat použití slovníků, které jsou zdokumentované, přístupné a propojené s globálně jedinečnými a trvalými identifikátory. • Dokumentace slovníků v NKR: Používané slovníky a ontologie by měly být evidovány v NKR a kde je to možné, obsahovat napojení na autority jako ORCID, ROR a Wikidata.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Použití standardizovaných slovníků: Metadata v repozitářích musí používat řízené slovníky nebo ontologie, které splňují FAIR principy. • Napojení metadat na autoritní registry: Metadata by měla obsahovat odkazy na ORCID (vědci), ROR (organizace) a Wikidata (koncepty) 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: Každý FIP by měl definovat, jaké slovníky a ontologie jsou preferované v daném oboru. • Doporučení pro správnou implementaci slovníků: FIP by měl obsahovat pokyny, jak zajistit správné používání slovníků v metadatech.
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 objekty jsou explicitně definovány (např. vztah k původnímu datasetu, odvozené verzi, doprovodným publikacím). • Používají se standardizované vztahy definované v metadatových schématech (např. DataCite, PROV-O⁵¹, 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</p>

⁵⁰ Czech Academic and Research Discovery Services - <https://www.techlib.cz/cs/84666-czech-academic-and-research-discovery-services-cards>

⁵¹ <https://www.w3.org/TR/prov-o/>

	<p>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 dataset, ze kterého jsou data odvozena. • isCitedBy – odkazuje na publikaci, která dataset cituje. • hasPart / isPartOf – propojuje datové sady, které jsou součástí většího celku. • affiliation – propojuje dataset s organizací prostřednictvím ROR identifikátoru. • creator – propojuje dataset s autorem prostřednictvím ORCID identifikátoru.
Doporučení CDIF	Data Integration profil doporučuje využití DDI-CDI ⁵² 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	<p>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, Wikidata) nebo jiná úložiště, která poskytují další kontext nebo informace, případně verze (meta)dat.</p> <p>NRP zajistí, aby tyto odkazy byly řádně formátovány, pokud možno používaly trvalé identifikátory (PID) a aby byla jasně popsána povaha vztahů mezi různými prvky metadat.</p>
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. • Použití standardizovaných ontologií pro popis vztahů: Metadata musí používat mezinárodně uznávané ontologie pro definování vztahů mezi entitami (např. PROV-O, DataCite, DCAT-AP). • Podpora autoritních identifikátorů: Metadata musí umožňovat propojení na ORCID, ROR a Wikidata, čímž se zvýší propojenost a reprodukovatelnost výzkumných dat.
Požadavky na repozitáře	<ul style="list-style-type: none"> • Zahrnutí odkazů na související data: Metadata musí obsahovat odkazy na původní dataset, 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: Každý dataset by měl být propojen 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.

⁵² https://ddialliance.org/ddi-cdi_v1.0

	<ul style="list-style-type: none"> 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 dosažení této metriky je důležité používat standardizované metadatové schéma, zahrnout jasné informace o autorech, datu vzniku, verzi dat, použitém formátu a metodách sběru dat. Klíčová je také přítomnost strojově čitelné licence a popis použitých standardů nebo řízených slovníků.
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. R14 požaduje dostatek informací v metadatech, aby byla data dlouhodobě pochopitelná a bylo jasné definováno, za jakých podmínek je lze znovu použít.
Naplnění v NRP	NRP řeší tento požadavek prostřednictvím iniciativy IPs CARDS. NRP vyžaduje, aby všechna metadata bylo možné namapovat na minimální metadatový záznam uložený v NMA. Metadata musí obsahovat informace o původu dat, autorství, použité metodologii a licenčních podmínkách. Pro podporu opětovného použití dat jsou doporučovány standardizované ontologie a řízené slovníky.
Požadavky na NRP	<ul style="list-style-type: none"> Zajistit, aby minimální metadatové schéma obsahovalo všechny klíčové atributy potřebné pro opakované použití dat. Implementovat systém, který umožní přidání licence v strojově čitelné podobě.
Požadavky na repozitáře	<ul style="list-style-type: none"> Implementace minimálního metadatového schématu: Všechna metadata musí být v souladu s požadavky NMA a obsahovat informace o původu dat, licencích a metodologii. Podpora oborově specifických metadat: Repozitáře by měly umožňovat přidání oborově specifických metadatových prvků definovaných v příslušných FIP. Verzování metadat: Zajistit uchování historie metadat a verzování, které umožní sledování změn a vývoje datových sad.
Požadavky na FIP	<ul style="list-style-type: none"> Definování povinných metadatových atributů: FIP musí obsahovat seznam povinných a doporučených atributů pro daný obor. Doporučení pro licenční standardy: FIP by měl specifikovat vhodné licenční podmínky pro daný typ dat a způsoby jejich implementace.

	<ul style="list-style-type: none"> • 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 datovým objektem.
Doporučení CDIF	Data Access profil doporučuje použití ODRL ⁵³ 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. Obecně je potřeba řešit tuto problematiku na základě spolupráce NRP s IPs CARDS, podrobněji pak na úrovni repozitářů prostřednictvím FIP. NRP vyžaduje, aby každá datová sada měla přiřazenou licenci, která je součástí metadatového záznamu. Metadata obsahující licenční informace jsou dostupná prostřednictvím NMA.</p> <p>Pro zajištění strojově čitelnosti licenčních podmínek je doporučeno používat standardizované formáty, jako jsou:</p> <ul style="list-style-type: none"> • Creative Commons Rights Expression Language (CC REL)⁵⁴ • Dublin Core (DC Terms) • DCAT-AP • JSON-LD • ODRL (Open Digital Rights Language) • SPDX (Software Package Data Exchange)⁵⁵, zejména pokud se metadata vztahují k softwarovým komponentám nebo open-source datovým souborům. <p>SPDX může být použit pro standardizované označování licencí v metadatech repozitářů a je široce podporován v softwarových ekosystémech, zatímco ODRL a CC REL umožňují detailnější správu podmínek použití datových sad v širším kontextu.</p>

⁵³ <https://www.w3.org/ns/odrl/2/ODRL20.html>

⁵⁴ https://wiki.creativecommons.org/wiki/CC_REL

⁵⁵ <https://spdx.dev/>

Požadavky na NRP	<ul style="list-style-type: none"> Podpora strojově čitelných licencí: Zajistit, že všechny repozitáře podporují formáty strojově čitelných licenčních podmínek. Centrální katalog licencí: Vytvořit seznam doporučených licencí a zajistit jejich snadné použití při publikaci dat.
Požadavky na repozitáře	<ul style="list-style-type: none"> Povinné uvedení licence: Každý metadatový záznam musí obsahovat informace o licenci datové sady. Důležitý je aspekt strojové zpracovatelnosti licenčních podmínek. Podpora standardizovaných licencí: Repozitáře by měly umožňovat výběr z doporučených licencí. Validace licenčních podmínek: Automatické kontroly by měly ověřovat přítomnost a správnost licenčních informací.
Požadavky na FIP	<ul style="list-style-type: none"> Doporučení vhodných licencí pro daný obor: FIP by měl specifikovat preferované licence, které odpovídají potřebám dané vědecké komunity. Specifikace formátů pro strojově čitelné licence: FIP by měl jasně uvádět, jakým způsobem mají být licenční informace uvedeny v metadatech.
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	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.
Naplnění v NRP	V rámci NRP bude povinnou součástí metadatového záznamu informace o provenienci. Metadata by měla obsahovat minimální úroveň informací o původu pro všechny soubory údajů. To zahrnuje základní podrobnosti, jako je tvůrce dat, kroky zpracování a stav publikace. Implementace podrobného sledování původu v rámci NPR je spojena s několika klíčovými aktivitami. Klíčová aktivita "Compliance" zajišťuje, aby informace o původu odpovídaly regulačním normám a standardům osvědčených postupů. Klíčová aktivita "Automatizace sběru dat a

	metadat” se zaměřuje se na automatizaci zachycování metadat o původu během fáze sběru a zpracování dat.
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. • Podpora verzování: Implementovat mechanismy pro sledování změn v datových sadách a jejich metadatech. • 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 informace o původu dat, jejich verzí a historii změn. • Podpora nástrojů pro sledování verzí: Umožnit uchovávání 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 (W3C Provenance Ontology) 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 znovuvyužití 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)⁵⁶ • Sociální vědy: DDI-C (Data Documentation Initiative - Codebook)⁵⁷ • Fyzika: HDF5 (Hierarchical Data Format)⁵⁸
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.

⁵⁶ <https://pubmed.ncbi.nlm.nih.gov/11726920/>

⁵⁷ <https://ddialliance.org/ddi-codebook>

⁵⁸ <https://www.hdfgroup.org/solutions/hdf5/>

Naplnění v NRP	<p>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.</p>
Požadavky na NRP	<ul style="list-style-type: none"> • Podpora komunitních standardů: Zajistit, aby repozitáře podporovaly metadatové 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-Checker⁵⁹, F-UJI⁶⁰). • 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.
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 schémátům. • 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.

⁵⁹ <https://fair-checker.france-bioinformatique.fr/>

⁶⁰ <https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>

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)⁶¹ – profil implementace FAIR, což je metodika vyvinutá v rámci iniciativy GO FAIR⁶². Díky ní mají vědecké a výzkumné komunity možnost vyjádřit svůj výběr standardů, technologií, nástrojů a postupů, tzv. FAIR-Enabling Resources (FER)⁶³, 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 celé úložiště dat (repozitář). Slouží jako vzor pro správu dat, zefektivňuje přípravu plánu správy dat a jeho zlepšování.

Oborovou specifičnost je v NRP 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 WorldFAIR⁶⁴, kde bylo vytvořeno 11 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)⁶⁵ je interaktivní webový nástroj určený k usnadnění tvorby a správy FIP. Uživatelům umožňuje systematicky definovat, dokumentovat

⁶¹ FIP - <https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/>

⁶² <https://www.go-fair.org/go-fair-initiative/>

⁶³ <https://w3id.org/fair/fip/terms/FAIR-Enabling-Resource>

⁶⁴ <https://worldfair-project.eu/the-project/>

⁶⁵ <https://catalogue.fair-impact.eu/node/212>

a sdílet své FIP. Nástroj je založen na open-source projektu Data Stewardship Wizard⁶⁶, 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 nanopublikací, 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 NRP bude FIP Wizard využíván pro tvorbu FIP, jak bylo zmíněno výše. Bude spravována vlastní instance, která bude technicky nezávislá na globální síti nanopublikací⁶⁷ 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 s podporou exekuce

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 Wizard⁶⁸, které zajistí přenos vhodných FER z FIP do DMP, jak je popsáno v sekci [2.3.1](#) a [2.3.2](#).

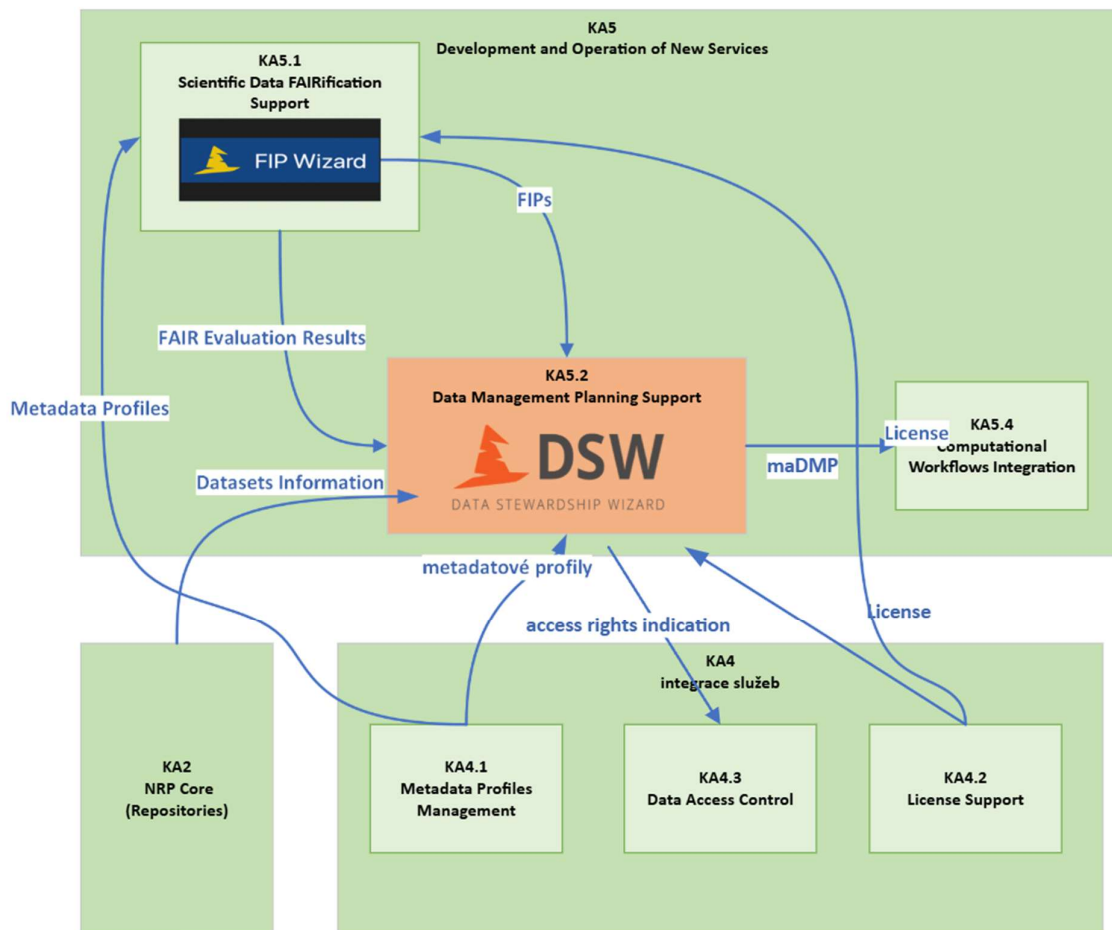
Samotný Data Stewardship Wizard poté představuje nejen nástroj pro plánování správy dat, ale může sloužit též ve všech částech životního cyklu dat jako “kokpit pilota”. To je umožněno strojovou zpracovatelností plánů dat (machine-actionable DMP, maDMP⁶⁹) a rozsáhlými možnostmi technických integrací skrze 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](#)).

⁶⁶ DSW - <https://ds-wizard.org/>

⁶⁷ <https://nanopub.net/>

⁶⁸ <https://ds-wizard.org/>

⁶⁹ Machine-actionable DMP - strojově čitelný DMP



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 digitálního objektu. 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 zásadní 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 souborů), a poskytují doporučení k nápravě těchto nedostatků.

Mnoho aspektů FAIR metrik je v praxi zajištěno podobou a nastavením metadatového popisu použitého repozitáře. Z logiky věci jsou sebehodnotící testy schopny uživatelům poskytnout zpětnou vazbu ještě před volbou vhodného repozitáře, depozicí dat, a dokonce i před jejich samotnou tvorbou. Automatizované testy jsou uplatnitelné až v pozdější fázi životního cyklu dat, protože vycházejí z již vytvořeného metadatového popisu v konkrétním repozitáři. Z toho plyne, že FAIR hodnotící nástroje jsou pro FAIRifikaci datových souborů ukládaných v NRP 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**⁷⁰ (**FDMM**). 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 datasetu. Následující nástroje jsou příklady sebehodnotících testů, první dva jsou určeny zejména jednotlivcům, třetí je zaměřen na správce repozitářů.

FAIR Self-Assessment Tool⁷¹ 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)⁷² je dostupný na GitHub pod otevřenou licencí. Konkrétní hodnocení

⁷⁰ <https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/solution/rda-fair>

⁷¹ <https://github.com/au-research/FAIR-Data-Assessment-Tool>

⁷² <https://ardc.edu.au/>

zvolených odpovědí je postaveno na srovnání s interpretací „FAIR Data Principles“ definovaných skupinou „FORCE11“⁷³ – tedy komunitou vědců, knihovníků, archivářů, vydavatelů a poskytovatelů podpory. Nástroj má i český překlad⁷⁴, který zajistil kolektiv knihovny VUT.

FAIR Aware⁷⁵ byl vyvinut organizací DANS (Data Archiving and Networked Services)⁷⁶ 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)⁷⁷ 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 hodnocený objekt nebo dataset s komunitně definovanými schémata.

Tyto nástroje pracují výhradně s již popsány a uloženými datovými sadami, což znamená, že hodnocení je fixováno na konkrétní fázi životního cyklu dat. Výsledky hodnocení tedy do značné míry odrážejí parametry repozitáře, v němž jsou data uložena, a volby správce repozitáře týkající se dostupných metadatových schémat. Původce dat, resp. uživatel repozitáře má pouze omezený vliv na výsledky hodnocení, což je třeba zohlednit zejména při jejich dalším využití. Výsledky automatizovaných hodnocení **zpravidla nelze zjednodušeně interpretovat jako měřítko kvality datasetu** a v žádném případě by nemělo být posuzováno bez hlubší znalosti širšího kontextu a detailních informací o datasetu samotném. Jedná se pouze o jedno z možných vodítek pro práci s datasetem.

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

FAIR-Checker⁷⁸ - nástroj vytvořený v rámci francouzského nodu projektu ELIXIR (IFB - French Institute for Bioinformatics). Dostupný pod otevřenou licencí na portálu GitHub. Z webových stránek je extrahována anotace datasetu, 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

⁷³ <https://force11.org/>

⁷⁴ <https://github.com/skupaBUT/FAIR-Data-Assessment-Tool-CZ>

⁷⁵ <https://fairaware.dans.knaw.nl/>

⁷⁶ <https://dans.knaw.nl/en/about/>

⁷⁷ <https://faircore4eosc.eu/eosc-core-components/compliance-assessment-toolkit-cat>

⁷⁸ <https://fair-checker.france-bioinformatique.fr/>

Bioportal. Nakonec je ověřena kvalita metadat vůči komunitním profilům, jako jsou Bioschemas. Výstupem je paprskový graf s procentuálně vyjádřeným skóre ve čtyřech 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-UJI⁷⁹ - 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 re3data⁸⁰. 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 slabínách 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 Evaluator⁸¹ - 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 Validator⁸² - 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**⁸³: Portál s užitečnými nástroji a průvodci pro FAIR hodnocení.

⁷⁹ <https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>

⁸⁰ Registry of Research Data Repositories - mezinárodní registr repozitářů
<https://doi.org/10.17616/R3D>

⁸¹ <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/>

⁸² <https://catalogue.openaire.eu/service/openaire.metadata.validator/overview>

⁸³ <https://fairassist.org/#!/>

- **FAIR Implementation Framework**⁸⁴: Rámec pro implementaci FAIR principů v různých kontextech vytvořený v projektu FAIR-Impact.
- **FC4E-CAT-UI**⁸⁵: FAIR Compliance Assessment Toolkit – Nástroj pro podrobné hodnocení FAIR compliance organizací a jejich datových sad vytvořený v projektu FAIRCORE4EOSC.

⁸⁴ <https://fair-impact.eu/fair-implementation-framework>

⁸⁵ <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, technických řešení a 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, 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. Jejich 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. I přes jejich možné nedostatky nebo omezení jsou dobrým nástrojem na zjištění zpětné vazby, tak jak je to na konkrétních příkladech ukázáno v [Příloze 1](#). Na druhé straně by nikdy neměly sloužit jako samostatné kritérium pro hodnocení kvality datových souborů, a to jak proto, že naplňování FAIR principů je z velké části odpovědností správců repozitářů, tak proto, že automatizované nástroje nemohou v dostatečné míře odrážet všechna dílčí specifika dat a metadat. 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 dat v repozitářích nástrojem 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 každému datasetu bodové ohodnocení, které vyjadřuje míru naplnění FAIR principů. Hodnocení vychází z metodiky FAIRsFAIR Data Object Assessment Metrics (verze 2022)⁸⁶.

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.

Repozitář ASEP

ASEP je institucionální repozitář AV ČR⁸⁷. 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. Repožitář ASEP, 2025. Licence: CC BY 4.0. Dostupné z: <https://doi.org/10.57680/asep.0616897>

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

Výsledek hodnocení

Celkový FAIR level: 2 - moderate (70%)



Komentované výsledky hodnocení

- V metadatech chybí základní definice digitálního objektu.**
F-UJI hlásí, že digitální objekt, který je předmětem hodnocení, se neidentifikuje jako

⁸⁶ <https://doi.org/10.5281/zenodo.6461229>

⁸⁷ <https://asep.lib.cas.cz/arl-cav/cs/index/>

dataset. V ASEP se zapisuje typ datasetu (dataset, text, image, audio/video, software), který se popisuje do metadat. V případě tohoto záznamu je uveden typ datasetu "Image". Pokud byl v jiných hodnocených metadatových záznamech uveden typ datasetu "dataset", nástroj chybu nehlásí. Chybí obecná reprezentace digitálního objektu v metadatech, typ datasetu je třeba uvést v metadatech s odpovídajícími atributy.

2. **V metadatech chybí informace o formátu a o velikosti datasetu.** Informace o formátu a velikosti datasetu nejsou reprezentovány v metadatech. Při ukládání datových souborů do repozitáře ASEP se doporučuje soubory zazipovat, zároveň by měly být v metadatech uvedeny i skutečné formáty souborů a jejich velikost.
3. **V metadatech chybí údaje o verzování a souvisejících publikacích.** Na landing page chybí metadata o souvisejících zdrojích a verzích datasetu – **hodnocení se liší podle použití handle nebo DOI:** datové záznamy s uloženým datasetem v ASEP mají přidělen handle a DOI, datové záznamy s odkazem na dataset 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 datasetu chybí v metadatech na landing page i v DataCite.
4. **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 dataset získat.
5. **V metadatech chybí informace o použitém slovníku nebo ontologii.** Repozitář ASEP je multioborový repozitář, používá číselník oborových skupin OECD – Frascati Manual 2015⁸⁸, který se používá se v ISVaVal.⁸⁹

LINDAT/CLARIAH-CZ

LINDAT/CLARIAH-CZ⁹⁰ 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%)

⁸⁸ 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>.

⁸⁹ <https://www.isvavai.cz>

⁹⁰ <https://lindat.cz/services/catalog/>

Summary:



Komentované výsledky hodnocení

Metadata neobsahují identifikátor dat, na která odkazují. Jelikož nástroj nenalezl identifikátor samotných dat, tak celá řada dalších testů v 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)

ČSDA⁹¹ je český uzel mezinárodní infrastruktury CESSDA a národní archiv pro data ze sociálních věd.

Příklad hodnoceného záznamu

Tabery, Paulína; Čadová, Naděžda; Vinopal, Jiří; Ďurďovič, 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, V1⁹²

Výsledek hodnocení

Celkový FAIR level: 3 - advanced (83%)

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 soubory, 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.

⁹¹ <https://archivdv.soc.cas.cz>

⁹² <https://doi.org/10.14473/CSDA/IUKDAJ>

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 datasey a metadatové záznamy do repozitářů a hodnotí FAIRovost 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 digitálních objektů 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 vkládání. Nabízí robustní zabezpečení a kontrolu přístupu uživatelů. Spolupracuje s komunitami na udržování definování oborových FAIR standardů.

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. 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 datasey 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 správcem 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 (datasety)

Výzkumná data zahrnují veškeré informace vytvořené měřením, pozorováním, výpočty nebo sběrem za účelem dosažení, ověření nebo reprodukce výsledků výzkumu. Tyto datasety

tvoří základ vědecké analýzy a musí splňovat FAIR principy pro zajištění přístupnosti a opětovného využití.

Metadata (metadatový záznam)

Metadata jsou strukturovaná data popisující datasety, a umožňují jejich vyhledávání, správu a opětovné využití. Jsou obvykle vytvářena vědci nebo datastewardsy, a mohou být uložena samostatně nebo spolu s popisovanými daty. Jsou klíčová pro 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 souborů.

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 datové správy)

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 umožňuje uživatelům zkontrolovat FAIRovost dat pomocí jednoduchého dotazníku založeného 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

Výzkumné přístroje 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.