

Integrovaná správa výzkumných dat

Governance, důkazy a rozhodovací rámce

Ondřej Ryška

O této knize

Tato monografie se zabývá strukturním problémem v governance výzkumných dat. Normativní a procedurální vrstva, která výzkumná data upravuje, narůstala nabalováním a její nástroje — FAIR, modely životního cyklu, plány správy dat, metadatové standardy, specifikace provenience, referenční modely dlouhodobého uchování, právo ochrany osobních údajů, certifikace repozitářů a reforma hodnocení výzkumu — jsou sice každý vnitřně soudržné, žádný z nich však nebyl navržen tak, aby se skládal s ostatními. Chybí spojovací vrstva, která podmínky konkrétního projektu nebo instituce převádí na obhajitelné rozhodnutí, a s ní i doklad o tom, že rozhodnutí bylo provedeno.

Kniha navrhuje, operacionalizuje a posuzuje integrovaný rámec institucionální správy výzkumných dat. Rámec vymezuje dimenze, v nichž se rozhodnutí o správě dat přijímají, proměnné, jimiž je každá dimenze naplněna, vztahy mezi nimi a rozhodovací pravidla, podle nichž se stanoví priorita mezi konkurujícími si požadavky. Je odvozen ze zmapovaného korpusu rámců, standardů, institucionálních politik a literatury pokrývajících dvacet tři domén praxe a je posuzován vůči institucionálním a projektovým prostředím, v nichž konkurující si požadavky skutečně vznikají.

Komu je kniha určena

Kniha je adresována badatelům v oblasti správy výzkumných dat a otevřené vědy, data stewardům a profesionálům podpory výzkumu, osobám odpovědným za výzkumnou politiku na univerzitách a ve výzkumných ústavech a doktorandům pracujícím na governance digitálních výzkumných objektů. Předpokládá obeznámenost se slovníkem správy výzkumných dat, nikoli však s kterýmkoli konkrétním rámcem.

Co kniha nedělá

Není příručkou a nepředepisuje jednotlivým institucím postupy. Tam, kde z výsledků plynou praktické důsledky, jsou odvozeny v části VI a označeny jako odvození, nikoli předloženy jako instrukce. Rámec explicitně vymezuje hranice vlastní přenositelnosti: oborové konvence, právní režim, velikost

instituce a dostupná infrastruktura mění to, co může rámec tohoto druhu vůbec určit, a kapitola 14 uvádí podmínky, za nichž jeho určení slábnou.

Poznámka k důkaznímu statusu textu

Tvrzení v této knize patří ke třem druhům a každé z nich je na místě svého výskytu označeno. Převzaté fakty nesou citaci svého primárního zdroje. Syntézy provedené touto studií jsou jako takové označeny ve větě, která je uvádí. Původní výsledky této studie jsou uvedeny jako výsledky a nesou svůj důkaz.

Tam, kde argument vyžaduje empirický výsledek, který výchozí výzkumný program dosud neposkytl, nese text na příslušném místě vsazenou poznámku uvozenou návěstím **CHYBĚJÍCÍ DOKLAD**, jež pojmenovává, co má být změřeno a jaké tvrzení by to podepřelo. Úplný soupis těchto poznámek podává příloha A. Citační značky přítomné v dřívějších verzích rukopisu byly vyřešeny proti ověřeným zdrojům a bibliografie toto ověření zaznamenává. Důkazní značky jsou záměrné. Zaznamenávají hranici mezi tím, co rámec tvrdí, a tím, co dosud doložil, a mají být vyřešeny před předložením rukopisu k recenznímu řízení.

Stav rukopisu

Tento dokument je pracovní verzí. Následující tabulka zaznamenává stav jednotlivých kapitol.

Tabulka 0-1. Stav rozpracovanosti podle kapitol.

Část	Kapitola	Stav
I	1. Fragmentace v governance výzkumných dat	napsáno
I	2. Data stewardship jako institucionální způsobilost	napsáno
I	3. FAIR, CARE a TRUST: kritické čtení	napsáno
II	4. Výzkumný design, korpus a analytická metoda	napsáno
III	5. Governance, role a odpovědnost za výsledek	napsáno
III	6. Životní cyklus výzkumných dat a provenience	napsáno
III	7. Plánování jako závazek: plány správy dat	napsáno
III	8. Popis, kvalita a podmínky opětovného použití	napsáno
III	9. Ochrana, právní režim a řízený přístup	napsáno
III	10. Reprodukovatelnost a výpočetní integrita	napsáno
III	11. Dlouhodobé uchování, publikování a uznání	napsáno
III	12. Operacionalizace rámce	napsáno
IV	13. Aplikace na institucionální prostředí	napsáno
IV	14. Robustnost, citlivost a hranice platnosti	napsáno
V	15. Kritická diskuse	napsáno
VI	16. Důsledky pro instituce a infrastruktury	napsáno
VI	17. Závěr	napsáno

Všech sedmnáct kapitol je napsáno a závěrečné části nyní zahrnují konsolidovanou a ověřenou bibliografii, glosář, osm obrázků, protokol externí recenze (příloha B) a rejstřík práv (příloha C). Příloha A indexuje každou dosud nevyřešenou důkazní a citační značku v rukopisu a tvoří výzkumný program, který musí být dokončen před předložením. Všechny citační

značky byly vyřešeny proti ověřeným zdrojům a bibliografie obsahuje 59 záznamů. Nadále zbývá: empirický program přílohy A, vlastní externí recenze a rejstřík, který se sestavuje až při produkci.

Č Á S T I

Problém a jeho teoretické zakotvení

Kapitola 1. Fragmentace v governance výzkumných dat

1.1 Regulovaná oblast bez řídicího rámce

Výzkumná data upravuje hustá a stále se rozšiřující vrstva principů, standardů, právních nástrojů a podmínek poskytovatelů finanční podpory. Principy FAIR popisují cílové vlastnosti digitálních výzkumných objektů (Wilkinson et al. 2016), modely životního cyklu popisují fáze, jimiž data procházejí, plány správy dat zaznamenávají záměry, metadatové standardy předepisují popis a modely provenience zaznamenávají odvození. Dlouhodobému uchování se věnuje referenční model OAIS a certifikace repozitářů (Consultative Committee for Space Data Systems 2012; CoreTrustSeal 2022), právním a etickým podmínkám právo ochrany osobních údajů a etické posouzení (European Parliament and Council 2016) a publikování dat kritéria výběru repozitáře spolu s infrastrukturami persistentních identifikátorů (Science Europe 2021). Každý z těchto nástrojů je kompetentní v oblasti, pro niž byl navržen.

Obtíž, jíž se tato kniha zabývá, není nedostatečnost některého jednotlivého nástroje. Spočívá v tom, že nástroje nejsou propojeny do systému, který by podmínky daného projektu nebo instituce převedl na obhajitelné rozhodnutí. Instituce, která je přijme všechny současně, získává soubor částečně se překrývajících povinností adresovaných různým objektům, držení různých vlastníky a podepřených důkazy nestejné síly. Tabulka 1-1 toto částečné pokrytí zachycuje a pro každou oblast rozhodování uvádí, co dominantní nástroj ponechává neurčené.

Tabulka 1-1. Částečné pokrytí normativní a procedurální vrstvy upravující výzkumná data, s rozhodnutími, která jednotlivé dominantní nástroje ponechávají neurčená.

Oblast rozhodování	Dominantní nástroj	Adresovaný objekt	Ponecháno neurčené
Principy popisu a opětovného použití	FAIR	datová sada a její metadata	které vlastnosti platí pro kterou třídu dat a za jakou cenu
Plánování	plán správy dat	záměry projektu	zda je plán naplněn a kdo to ověřuje

Oblast rozhodování	Dominantní nástroj	Adresovaný objekt	Ponecháno neurčené
Životní cyklus	modely fází a toků	datové objekty v čase	autoritativní kopie, práh verzování, konec životnosti
Popis	metadatové standardy a profily	vyhledatelnost a sémantika	hranice mezi jádrovým a doménovým profilem
Provenience	PROV a specifikace balíčkování	odvození objektů	granularita úměrná riziku
Kvalita	rámce zajišťování kvality	vhodnost k použití	přijatelná zbytková chyba pro stanovený účel
Ochrana	právo ochrany osobních údajů, schémata klasifikace dat	osoby a důvěrnost	režim přístupu k datům, která nejsou ani otevřená, ani uzavřená
Dlouhodobé uchování	OAIS, certifikační požadavky	archivovaná informace	doba uchování a předání kusteodie
Publikování	kritéria výběru repozitáře	uložení a identifikátory	kompromis mezi oborovou vhodností a certifikovanou důvěryhodností
Uznání	reforma hodnocení výzkumu	výsledky a role	jak úsilí vynaložené na správu dat vstupuje do institucionálního hodnocení

Vrstva popsaná v tabulce 1-1 narůstala nabalováním a její tvar tuto historii odráží. Každý nástroj byl formulován jako odpověď na odlišné selhání: data, která nebylo možné najít, plány, které nikdy nevznikly, archivy, jež ztratily schopnost interpretovat vlastní fondy, repozitáře, které skončily bez nástupnictví. Protože každý byl navržen proti jinému způsobu selhání, je každý vnitřně soudržný a žádný nebyl navržen tak, aby se skládal s ostatními. Absence spojovací vrstvy je tedy předvídatelným důsledkem toho, jak se obor vyvíjel, spíše než pochybením kterékoli z jeho komunit.

Vzorec patrný napříč tabulkou je systematický, nikoli nahodilý. Nástroje jsou nejsilnější tam, kde lze objekt popsat izolovaně, a nejslabší tam, kde se rozhoduje za konkurujících si omezení. Rozhodnutí druhého druhu tvoří běžný obsah institucionální i projektové praxe. Mezera se proto nachází právě tam, kde je praxe nejhustší.

1.2 Proč je fragmentace výzkumným problémem

První reakce chápe situaci jako mezeru v implementaci, napravitelnou lepším školením, více nástroji a přísnějšími kontrolami souladu. Tato reakce je nedostatečná ze tří důvodů, které tato kniha považuje za konstitutivní pro daný problém. Nástroje jsou nesouměřitelné, nesou důkazy nestejné síly a dostávají se do konfliktů, jež nelze vyřešit zevnitř žádného z nich. Každý z důvodů je dále rozveden.

Nesouměřitelnost vzniká proto, že nástroje se vztahují k různým objektům. FAIR se vypovídá o datové sadě, TRUST o repozitáři, plán správy dat o projektu, právo ochrany osobních údajů o operaci zpracování a reforma hodnocení o kariéře. Soud, že datová sada je FAIR, tedy neříká nic o tom, zda ji instituce, která ji drží, dokáže dlouhodobě uchovat, a certifikovaný repozitář nečiní uloženou datovou sadu vhodnou k opětovnému použití. Skládání napříč těmito objekty není definováno žádným z rámců, které je konstituují.

Nestejná důkazní síla problém dále prohlubuje. Právní požadavek, ratifikovaný standard, recenzovaná srovnávací studie a nastupující komunitní praxe se v institucionálních politikách běžně citují vedle sebe, jako by byly téhož druhu. Revize, o niž se zde analyzovaný korpus opírá, oddělila zdroje do tří úrovní důkazní síly a toto oddělení uplatnila konzistentně napříč všemi 23 doménami. Srovnání uváděné v této knize naznačuje, že většina vodítek relevantních pro rozhodování v oboru spočívá na kazuistických důkazech a konsenzuálních dokumentech, nikoli na srovnávacím měření, což omezuje, jak silně lze formulovat kterýkoli rámec na nich vybudovaný.

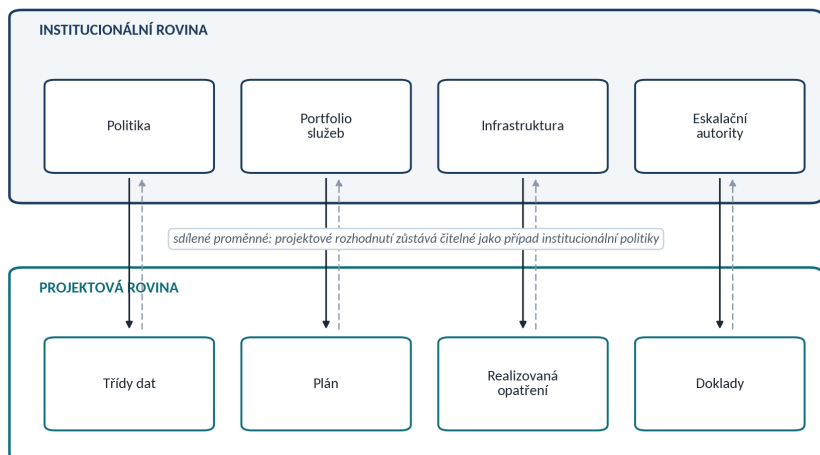
Konflikt je třetím a rozhodujícím důvodem. Otevřená licence maximalizuje opětovné použití a může současně porušit komunitní protokol, retenční pravidlo a povinnost výmazu mohou obojí dopadat na tytéž záznamy, oborová přiměřenost a certifikovaná důvěryhodnost mohou ukazovat k různým repozitářům. Takové případy nejsou selháními implementace, protože každý z konkurujících si požadavků je uplatněn správně. Vyřešit je lze pouze vnějším pravidlem, které přiděluje prioritu, pojmenovává autoritu oprávněnou rozhodnout a zaznamenává zbytkové riziko, jež rozhodnutí ponechává.

Čtvrtá úvaha se týká měření, nikoli rozhodování. Běžně užívané indikátory počítají dokumenty a činnosti: napsané plány, poskytnuté konzultace,

uskutečněná školení. Takové počty jsou necitlivé k tomu, zda byla opatření popsaná v plánu provedena, a nedokážou odlišit instituci s fungující eskalací od instituce s organogramem. Absence měření na úrovni výsledku není samostatným nedostatkem, nýbrž toutéž fragmentací nahlíženou ze strany posuzování.

Nedostatečnost implementačního čtení lze doložit i ze strany nabídky. Strukturovaný přehled omezený pouze na vědy o živé přírodě analyzoval více než 300 odlišných nástrojů pro data stewardship (Aksenova et al. 2024) a registr FAIRsharing eviduje tisíce standardů, databází a politik napříč obory (Sansone et al. 2019), což naznačuje spíše nasycení nabídky než její nedostatek. Přijetí dalšího nástroje digitalizuje nejasný postup, aniž by se vyřešilo, kdo rozhoduje, za jakých podmínek a s jakým dokladem o dokončení. Vzácná není instrumentace, nýbrž rozhodovací struktura, do níž lze instrumentaci umístit.

Podobu, jakou může řešení mít, určuje ještě jedna vlastnost problému. Rozhodnutí o výzkumných datech se přijímají na dvou úrovních, které se obvykle zkoumají odděleně: instituce nastavuje politiku, infrastrukturu a minimální požadavky, zatímco projekt přijímá konkrétní rozhodnutí o vlastních datech. Rámec fungující pouze na jedné úrovni produkuje buď politiku, kterou nemůže uplatnit žádný projekt, nebo projektovou praxi, kterou instituce nedokáže posoudit. Rámec rozvinutý v části III je proto vázán vyjádřit obě úrovně týmiž proměnnými tak, aby projektové rozhodnutí zůstalo čitelné jako instance institucionální politiky.



Obr. 1-1. Dvouúrovňová rozhodovací struktura zobrazující institucionální vrstvu (politika, portfolio služeb, infrastruktura, eskalační autority) a projektovou vrstvu (třídy dat, plán, realizovaná opatření, důkazy) se sdílenými proměnnými, které je spojují; podklad: syntéza 23 domén uvedená v části II. Vlastní zpracování autora.

CHYBĚJÍCÍ DOKLAD

Audit institucionálních politik pro výzkumná data napříč vymezenou populací výzkumných organizací, zaznamenávající pro každou z deseti oblastí rozhodování v tabulce 1-1, který nástroj je citován, zda je pojmenován vlastník rozhodnutí a zda je uvedena eskalační cesta — na podporu tvrzení, že pokrytí je částečné a nepropojené, nikoli pouze nerovnoměrné

CHYBĚJÍCÍ DOKLAD

Párové srovnání obsahu plánu správy dat při podání s realizovanou praxí při ukončení projektu na vzorku projektů, umožňující stanovit rozdíl mezi dokumentovým souladem a realizovanou praxí

1.3 Výzkumná otázka

Předchozí úvahy umísťují problém na úroveň návrhu rámce, nikoli instruktaže. Chybí zobrazení podmínek případu na rozhodnutí, které lze obhájit, spolu s kritérii, podle nichž lze rozhodnutí a jeho realizaci posoudit. Kniha klade jednu ústřední otázku a odvozuje z ní své části.

Za jakých podmínek lze částečné normativní, procedurální a technické požadavky upravující výzkumná data integrovat do jednoho rámce proměn-

ných, rozhodovacích pravidel a kritérií posouzení, který přináší obhajitelná rozhodnutí na institucionální i projektové úrovni?

Šest podřazených otázek tento problém rozděluje a odpovídá jednotlivým částem knihy. Zaprvé, jak je způsobilost produkující rozhodnutí o datech uvnitř instituce uspořádána a které vlastnosti tohoto uspořádání jsou relevantní pro rozhodování, nikoli kosmetické? Zadruhé, co principiální rámce oboru určují a v kterém bodě přestávají určovat cokoli? Zatřetí, které proměnné popisují datovou situaci s dostatečnou rozlišovací silou, aby poháněly rozhodovací pravidlo? Začtvrté, jakou podobu rozhodovacího pravidla ospravedlňují důkazy té síly, která je v tomto oboru skutečně k dispozici? Zapáté, podle jakých kritérií lze posoudit implementaci, nikoli dokumentaci? Zašesté, jak daleko se přenáší rámec odvozený z jednoho národního a institucionálního prostředí a co značí hranici, za níž se již nepřenáší?

1.4 Materiál, metoda a nárokový přínos

Materiálem knihy je korpus 23 domén praxe výzkumných dat, z nichž každá je zdokumentována svým stabilním konsenzem, svými změnami v období 2024–2026, svými otevřenými otázkami a svou národní implementační vrstvou. Korpus byl sestaven jako srovnávací čtení právních nástrojů, ratifikovaných standardů, infrastrukturních metodik, recenzovaných přehledů a institucionálních kazuistických důkazů, přičemž stav proměnlivých zdrojů byl zaznamenán k pevnému datu ověření. Analýza postupuje tak, že pro každou doménu vytěžuje skutečně přijímaná rozhodnutí, proměnné, které je podmiňují, a zbytkové riziko, jež každá varianta ponechává. Část II tento design vykládá v úplnosti, včetně kritérií, podle nichž se tvrzení připouští jako převzatý fakt, jako syntéza této studie, nebo jako původní výsledek.

Nárokový přínos je trojí a je uveden zde, aby jej bylo možné testovat proti následujícím kapitolám. První složkou je návrh integrovaného rámce, v němž jsou rozhodnutí napříč všemi 23 doménami vyjádřena sdíleným soubojem proměnných. Druhou je operacionalizace tohoto návrhu jako rozhodovacích pravidel s explicitními podmínkami, prioritním uspořádáním mezi konfliktními požadavky, pojmenovanou rozhodovací autoritou a zaznamenaným zbytkovým rizikem, spolu s kritérii posouzení, která se týkají realizace, nikoli dokumentace. Třetí je validace proti doménovým profilům a integro-

vanému případu, doprovázená explicitním vymezením hranic, v nichž je rámec přenositelný.

Napříč celou knihou jsou konzistentně odlišovány tři druhy tvrzení a toto odlišení je součástí přínosu, nikoli redakční konvencí. Převzaté fakty jsou připisány svému primárnímu zdroji. Syntézy této studie jsou jako takové označeny a jsou dohledatelné ke srovnání napříč 23 doménami. Původní výsledky jsou označeny, a tam, kde důkazy, jež by je podepřely, dosud nebyly shromážděny, je požadavek explicitně uveden namísto toho, aby byl naplněn ilustrací. Hustota takových formulací v následujících kapitolách je záměrným rysem argumentu vybudovaného na oboru, jehož srovnávací důkazní základna zůstává tenká.

1.5 Vymezení

Knihy není příručkou a nenabízí instrukce. Neposuzuje ani nesestavuje pořadí softwarových nástrojů, jejichž krajina se mění rychleji, než jakákoli publikovaná analýza dokáže sledovat. Nepředstavuje právní poradenství: právní otázky jsou pojednány jako třídy rozhodnutí s identifikovatelnou autoritou, nikoli jako odpovědi platné v určité jurisdikci v určitém čase. Není systematickým přehledem provedeným podle PRISMA a nenavrhuje národní politiku, ačkoli česká implementační vrstva slouží v celém textu jako test přenositelnosti, zejména zavedení digitální datové kolekce jako uznaného druhu výsledku a související národní metadatový adresář (Úřad vlády České republiky 2025; EOSC CZ 2025).

1.6 Struktura argumentu

Část I ustavuje problém a jeho zasazení. Obhazuje tvrzení, že fragmentace je výzkumným problémem, definuje data stewardship jako institucionální způsobilost, nikoli jako pracovní místo nebo servisní pult, a kriticky čte normativní vrstvu FAIR, CARE a TRUST, aby přesně stanovila, co ponechává neurčené.

Část II vykládá výzkumný design. Popisuje korpus 23 domén, odstupňování důkazní síly, postup, jímž byly rozhodovací proměnné vytěženy z dokumentované praxe, a pravidla, podle nichž byly váženy konfliktní zdroje. Uvádí rovněž meze designu, zejména převahu kazuistických důkazů nad srovnávacím měřením v tomto oboru.

Část III předkládá integrovaný rámec. Definuje proměnné, rozhodovací pravidla spojující stavy proměnných s variantami, prioritní uspořádání uplatňované při konfliktu požadavků, přidělení rozhodovací autority a eskalace a kritéria posouzení odlišující realizovaná opatření od dokumentovaných záměrů.

Část IV rámec aplikuje a validuje. Doménové profily pro technická, přírodovědná, společenskovední a humanitní data testují, zda si sdílené proměnné uchovávají rozlišovací sílu za oborové variability, a integrovaný případ sleduje jeden projekt celou sadou rozhodnutí. Národní implementační vrstva slouží jako druhé validační prostředí.

Část V výsledky diskutuje. Zkoumá alternativní vysvětlení pozorované konvergence institucionální praxe, uvádí hranice přenositelnosti, identifikuje empirickou práci, kterou rámec vyžaduje, ale dosud jí nedisponuje, a specifikuje, jaké důkazy by falzifikovaly jeho ústřední tvrzení.

Následující kapitola se ujímá první podřazené otázky. Dříve než rámec může přidělit rozhodnutí autoritám, musí být způsobilost, která tyto autority drží, charakterizována jako předmět analýzy, a nikoli předpokládána jako organizační fakt.

Kapitola 2. Data stewardship jako institucionální způsobilost

2.1 Jednotka analýzy

Data stewardship bývá často ztotožňován s pracovním místem, oddělením nebo servisním pultem a přítomnost kteréhokoli z těchto tří se pak čte jako doklad, že instituce svá výzkumná data spravuje. Tato kapitola obhazuje jinou jednotku analýzy. Data stewardship se zde pojímá jako institucionální způsobilost: jako rozdělení rozhodovacích práv, kompetencí a služebních závazků mezi několik útvarů, v němž je jmenovaný data steward jedním prvkem mezi mnoha. Obhajované tvrzení zní, že o tom, zda jsou rozhodnutí o výzkumných datech přijímána, zaznamenávána a prováděna, rozhodují vlastnosti tohoto rozdělení, nikoli organizační forma, do níž je zabaleno.

Vztah k ústřední výzkumné otázce je přímý. Rámec, který přiděluje rozhodnutí autoritám, předpokládá, že autority existují, jsou identifikovatelné a přežijí odchod osob, jež je právě zastávají. Kdyby organizační forma kolísala bez důsledků, mohl by ji rámec pojednat jako šum; kdyby byla jedna forma prokazatelně optimální, mohl by ji rámec zafixovat jako konstantu. Analýza v této kapitole naznačuje, že neplatí ani jedno, a že organizační forma proto vstupuje do rámce jako proměnná, jejíž stavy nesou různá zbytková rizika.

2.2 Od nahodilého poradenství k řízenému portfoliu služeb

První doložený posun v analyzovaném materiálu vede od příležitostného poradenství k popsanému portfoliu služeb. V poradenském režimu je podpora poskytována na vyžádání, její obsah závisí na tom, koho se kdo zeptá, a žádný záznam nespojuje dotaz s jeho vyřešením. V portfoliovém režimu je každá služba popsána svými vstupy, výstupy, vlastníkem, očekávanou dobou vyřešení a bodem, v němž předává věc jiné autoritě. Rozdíl není administrativním zabalením: druhý režim činí podpůrnou funkci auditovatelnou a auditovatelnou funkci lze posoudit podle výsledku, nikoli podle objemu.

S tímto posunem je spřažen druhý, od samostatně stojící podpory k integraci se systémy, v nichž se výzkum skutečně administruje. Plány správy dat, informační systémy o výzkumu, správa grantů, poskytování úložné kapacity, repozitáře a etické schvalování jsou stále častěji propojeny tak, že rozhodnutí zaznamenané v jednom systému se stává viditelným v ostatních. Integrace tohoto druhu mění povahu důkazů dostupných o datové praxi instituce, protože realizovaná opatření zanechávají stopy nezávislé na sebesouzení. Srovnání napříč 23 doménami uváděné v této knize naznačuje, že právě toto spřažení, a nikoli existence katalogu služeb jako dokumentu, odlišuje instituce schopné doložit implementaci od institucí, které dokážou popsat pouze záměry.

2.3 Organizační modely a jejich kompromisy

Napříč institucionálními kazuistickými důkazy se opakují čtyři organizační modely. Centralizovaný model soustřeďuje metodiku, přezkum a poskytování služeb do jednoho útvaru, distribuovaný model umísťuje podporu do fakult nebo ústavů, federovaný model kombinuje ústřední metodické a infrastrukturní jádro s data stewardy ukotvenými v doménách a projektově vnořený model připojuje data stewarda ke konkrétnímu projektovému týmu. Tabulka 2-1 uvádí přínos, omezení a vhodný kontext každého z nich spolu se silou důkazů, jež toto přiřazení podpiřejí.

Tabulka 2-1. Organizační modely data stewardshipu s jejich hlavním kompromisem, kontextem, do něž je kazuistické důkazy umísťují, a silou těchto důkazů.

Model	Hlavní přínos	Hlavní omezení	Kontext v kazuistických důkazech	Síla důkazů
Centralizovaný	jednotná kvalita a kontrola	odstup od oborové praxe; kapacitní úzké hrdlo	menší nebo metodologicky homogenní instituce; služba v počátcích	institucionální případové studie
Distribuovaný	oborová relevance	nekonzistentní pravidla; duplicita	velké fakulty se zralou lokální podporou	institucionální případové studie

Model	Hlavní přínos	Hlavní omezení	Kontext v kazuistických důkazech	Síla důkazů
Federovaný	standardizace spojená s doménovou expertizou	koordinační náklady; sporné linie financování	výzkumně intenzivní multidisciplinární univerzity	evropské rámce a praxe
Projektově vnořený	blízkost datům a týmu	ztráta znalostí při skončení projektu	velké nebo vysoce rizikové projekty	kazuistické důkazy

Federovaný model je v evropské případové literatuře nejčastěji uváděným výsledkem a udávaný důvod je konzistentní: čistě centrální uspořádání se vzdaluje oborové praxi, zatímco čistě lokální produkuje nevyrovnanou kvalitu a duplicitní úsilí. Tato konvergence je však slabším důkazem, než jak se jeví. Spočívá na případových studiích a institucionálních zprávách, přičemž srovnávacích studií měřících dlouhodobý účinek různých organizačních uspořádání na výsledky v oblasti výzkumných dat zůstává málo. Konvergence se zde proto čte jako pravděpodobná institucionální odpověď na koordinační problém, nikoli jako doložené optimum.

CHYBĚJÍCÍ DOKLAD

Srovnání více institucí s konstantní velikostí instituce a oborovým složením, měřící výsledkové indikátory (doba do vyřešení, úplnost metadat při uložení, podíl projektů s provedeným plánem dlouhodobého uchování) napříč centralizovaným, distribuovaným a federovaným uspořádáním, aby se otestovalo, zda organizační model predikuje výsledek po zohlednění zdrojů

Jakékoli obecné doporučení omezuje ještě jedna úvaha, s níž se v této knize zachází jako se zjištěním, nikoli jako s výhradou. Modely se neliší kvalitou, nýbrž podmínkami, za nichž se jejich omezení stávají snesitelnými. Malá instituce nemůže udržet plně specializovaný tým a přesto potřebuje identifikované interní i externí kompetence; výzkumně intenzivní univerzita specializaci udržet dokáže a platí za ni koordinací. Rámec v části III proto model nevybírání, nýbrž požaduje, aby model, který je zaveden, byl doprovázen kompenzujícími opatřeními vůči svému známému omezení a pojmenovaným vlastníkem této kompenzace.

2.4 Odpovědnost za výsledek napříč rolemi

Trvalou slabinou analyzovaného materiálu není absence rolí, nýbrž implicitnost odpovědnosti za výsledek — tedy odpovědnosti za to, že věc dopadne, na rozdíl od odpovědnosti za provedení konkrétního úkonu. Odpovědnost za bezpečnost, metadata, kvalitu, smlouvy a dlouhodobé uchování bývá často pokládána za sdílenou a sdílená odpovědnost, která není rozložena na přiřazení, je fakticky bez vlastníka. Nástroj sloužící k jejímu rozložení, tedy přiřazení odpovědností napříč pozicemi provádějící, odpovědné za výsledek, konzultované a informované, se v této knize pojímá jako důkazní prostředek, nikoli jako manažerská formalita, protože převádí organizační schéma na soubor testovatelných tvrzení o tom, kdo o čem rozhodl.

Toto rozložení má hranici, která se opakuje v každé doméně korpusu. Data steward strukturuje problém, srovnává varianty, koordinuje vstupy specialistů, dokumentuje rozhodnutí a ověřuje, že navazující kroky existují. Právní závěry, etické schválení, bezpečnostní výjimky, akceptace vysokého zbytkového rizika, podpis smlouvy a alokace významných zdrojů náleží hlavnímu řešiteli, právnímu útvaru, pověřenci pro ochranu osobních údajů, etické komisi, bezpečnosti informačních technologií nebo vedení instituce. Splnutí poradní a rozhodovací role je nejzávažnějším způsobem selhání identifikovaným v analyzovaném materiálu, protože produkuje rozhodnutí, která nejsou ani autorizovaná, ani dohledatelná.

Pro rozhodování jsou relevantní další dvě vlastnosti odpovědnosti za výsledek. První je kontinuita: odpovědnost navázaná na osobu, a nikoli na roli, nepřechází obměnou personálu ani ukončení projektu a otázka, kdo přebírá kustodii dat, rozhodnutí o přístupu, klíče a dokumentaci po skončení projektu, je v korpusu zodpovídána podstatně řidčeji, než je kladena. Druhou je sdílená funkce: jak se instituce opírají o federované infrastruktury, externě zajišťovaná úložiště a vnější repozitáře, stává se doložitelnost odpovědnosti za funkce vykonávané jinde podmínkou důvěryhodnosti, nikoli smluvním detailem (CoreTrustSeal 2022).

CHYBĚJÍCÍ DOKLAD

Retrospektivní sledování vzorku ukončených projektů s cílem zjistit, zda pojmenovaná role držela kustodii dat a rozhodnutí o přístupu 12 měsíců po skončení projektu, na podporu tvrzení, že vazbou omezující celý systém je kontinuita odpovědnosti za výsledek, nikoli její počáteční přidělení

2.5 Kompetence, uznání a nabídka způsobilosti

Kompetenční rámce se posunuly od generických popisů k diferencovaným profilům. Evropská doporučení ke kompetencím pro data stewardship uvádějí, že role není jednotná a že kompetence se liší podle institucionálního, oborového, technického a projektového kontextu (EOSC Association 2024). Od té doby bylo publikováno základní kurikulum pro nastupující data stewardy (Skills4EOSC 2025) a následovala národní kurikula organizovaná kolem kompetenčních oblastí a kariérních očekávání, z nichž nejrozpracovanější instancí je nizozemská cestovní mapa k národní implementaci FAIR data stewardshipu (Jetten et al. 2021). Pokročilá praxe je naproti tomu v korpusu charakterizována řízením rizik, smluvním workflow, auditem, prostředím s řízeným přístupem a doménovými profily, z nichž žádné vstupní kurikula nepokrývají.

Nabídka kompetencí je také funkcí uznání. Reforma hodnocení výzkumu explicitně rozšiřuje uznání na kurátorství dat, vývoj výzkumného softwaru a technické příspěvky vedle publikací (CoARA 2022; DORA 2013). Důsledek pro institucionální způsobilost je strukturní, nikoli symbolický: bez kariérních drah, kompetenčních profilů a rozpočtových linií závisí způsobilost na dobrovolném úsilí osob, jejichž příspěvek je v hodnocení neviditelný. Důkazů o účinku reformy specificky na datové a softwarové role zůstává málo, a tvrzení předkládané zde proto zní, že uznání je nutnou podmínkou stabilní způsobilosti, nikoli že bylo prokázáno jako podmínka postačující.

Kompetenční literatura se sbíhá v dalším bodě, který kvalifikuje technické čtení této role. Přehledy kompetencí datového knihovnictví identifikují komunikaci, budování vztahů a schopnost pracovat po boku výzkumníků jako rozhodující vedle technických znalostí (Federer 2018). Institucionální změna v datové praxi postupuje konzultací, facilitací a advokací, což jsou právě ty kompetence, které kurikulární inventáře zachycují nejhůře. Tato asymetrie mezi tím, co se vyučuje, a tím, co se uvádí jako účinné, patří k otevřeným otázkám přenášeným do části V.

2.6 Posuzování způsobilosti

Problém měření identifikovaný v kapitole 1 zde nabývá specifické podoby. Počítání vytvořených plánů, poskytnutých konzultací nebo absolvovaných školení měří činnost podpůrné funkce a je necitlivé k jejímu účinku. Operační indikátory opakující se v analyzovaném materiálu jsou jiného

druhu: doba prvního kontaktu vzhledem k zahájení projektu, četnost incidentů, kvalita metadat při uložení, obnovitelnost dat při testu, doba do vyřízení žádosti o přístup a podíl projektů s provedeným plánem dlouhodobého uchování. Takové indikátory se hůře sbírají a hůře přisuzují, a proto je korpus zaznamenává jako atribut zralých služeb, nikoli služeb nových.

Toto rozlišení je pro rámec podstatné, protože fixuje, čím kritérium posouzení musí být. Kritérium formulované jako existence dokumentu je naplněno vytvořením dokumentu; kritérium formulované jako doložitelné provedení opatření je naplněno pouze provedením a zanechává stopu, kterou může třetí strana zkontrolovat. Část III proto formuluje kritéria posouzení na úrovni realizovaných opatření, přičemž dokumentový artefakt pojímá jako důkaz, nikoli jako posuzovaný objekt.

2.7 Interpretace a hranice platnosti

Dvě alternativní vysvětlení pozorované variability zasluhují uvedení. První připisuje rozmanitost organizačních modelů institucionální historii a lokální politice, nikoli jakékoli racionální adaptaci; v tom případě by volba modelu nenesla žádnou informaci o výkonu. Druhé ji připisuje pouze úrovni zdrojů; v tom případě se model jeví jako zástupný ukazatel financování. Ani jedno nelze na dostupných důkazech vyloučit a obě jsou slučitelná se stanoviskem zaujatým zde, protože rámec pojímá model jako proměnnou a požaduje kompenzaci jeho známého omezení, místo aby z formy usuzoval na výkon.

Hranicí platnosti této kapitoly je důkazní asymetrie mezi popisem a účinkem. Organizační uspořádání jsou v případové literatuře bohatě popsána a ve svých důsledcích tence změřena, což omezuje jakékoli tvrzení o optimalitě na tvrzení o podmínkách přiměřenosti. Z analýzy plyne, že institucionální způsobilost vstupuje do rámce prostřednictvím čtyř vlastností relevantních pro rozhodování: přidělených rozhodovacích práv, eskalačních cest, ujednání o kontinuitě a měření na úrovni výsledku.

Následující kapitola se obrací k druhé podřazené otázce a zkoumá, zda normativní vrstva, na niž se instituce při výkonu těchto rozhodovacích práv odvolávají, tato rozhodnutí vůbec určuje.

Kapitola 3. FAIR, CARE a TRUST: kritické čtení normativní vrstvy

3.1 Co se od normativní vrstvy očekává

Normativnímu slovníku výzkumných dat dominují tři soubory principů: FAIR pro vlastnosti digitálních objektů (Wilkinson et al. 2016), CARE pro kolektivní práva komunit, z nichž data pocházejí (Carroll et al. 2020), a TRUST pro kvality repozitářů, které je drží (Lin et al. 2020). Institucionální politiky se běžně dovolávají všech tří a toto dovolání se obvykle čte tak, že vyřešilo otázku, jak se má s daty zacházet. Tato kapitola tvrdí, že takové čtení je mylné. Uvedené tři soubory se vztahují k různým objektům, odpovídají na různé otázky a ani společně nedourčují rozhodnutí, jimž instituce a projekty skutečně čelí.

Vztah k ústřední výzkumné otázce je přesný. Kdyby normativní vrstva institucionální rozhodnutí určovala, byl by integrovaný rámec nadbytečný a úkol by se zredukoval na zajištění souladu. Zde předkládaná analýza naznačuje, že vrstva určuje cílové vlastnosti výsledků, ponechává však nespecifikované podmínky, za nichž je každá vlastnost dosažitelná, prioritu mezi vlastnostmi v konfliktu i autoritu oprávněnou konflikt vyřešit. Tento zbytek je předmětem části III.

3.2 FAIR: strojová zpracovatelnost mylně čtená jako otevřenost

Principy FAIR stanovují, že digitální objekty mají být nalezitelné, přístupné, interoperabilní a opětovně použitelné, a jejich deklarovaný účel se týká schopnosti strojů data lokalizovat a pracovat s nimi s minimálním lidským zprostředkováním (Wilkinson et al. 2016). Přístupnost v tomto smyslu je vlastností protokolu a metadat, nikoli licence: principy požadují, aby podmínky přístupu byly zjištěitelné standardizovaným postupem a aby metadata zůstala dostupná i tam, kde data samotná dostupná nejsou. Autentizovaný a řízený přístup je s FAIR plně slučitelný. Časté ztotožňování FAIR s otevřenými daty je proto kategoriální chybou, a to chybou s důsledky, protože převádí specifikaci strojové zpracovatelnosti na politiku zveřejňování.

Důsledky jsou patrné na třech místech. Data, která nesou legitimní omezení, jsou pojímána jako selhání vůči principu, který neporušují, což odrazuje od popisu právě toho materiálu, u něž na dohledatelnosti záleží nejvíce. Naopak uložení pod otevřenou licencí je pojímáno jako naplnění FAIR i tam, kde je dokumentace pro jakékoli opětovné použití nedostatečná, což je selhání vůči principům pod zdáním souladu. Obě chyby plynou z toho, že se soubor vlastností nahradí stavem licence.

Praxe posuzování tutéž záměnu reprodukuje ve druhé podobě. Automatizované nástroje ověřují to, co je strojově testovatelné — přítomnost persistentního identifikátoru, získatelnost strukturovaných metadat, deklaraci licence, registraci v prohledávatelném indexu — a v tomto rozsahu jsou účinné. Nedokážou posoudit faktickou správnost dokumentace, přiměřenost ontologie, dostatečnost provenience pro deklarované opětovné použití ani legitimitu zveřejnění. Skóre odvozené z testovatelné podmnožiny pak funguje jako verdikt o celku a daná vlastnost je postupně předefinována na to, co nástroj měří (RDA FAIR Data Maturity Model Working Group 2020).

CHYBĚJÍCÍ DOKLAD

Analýza shody mezi automatizovanými skóre posouzení FAIR a nezávislým expertním auditem vhodnosti k použití na stratifikovaném vzorku datových sad, uvádějící podíl vysoce bodovaných datových sad shledaných nepoužitelnými pro jejich deklarovaný účel, na podporu tvrzení, že automatizovaná skóre neměří opětovnou použitelnost

3.3 CARE: otázky, které FAIR neklade

Principy CARE byly formulovány v rámci governance dat původních obyvatel a týkají se kolektivního prospěchu, oprávnění ke kontrole, odpovědnosti a etiky (Carroll et al. 2020). Jejich objektem není digitální záznam, nýbrž vztah mezi daty, lidmi, které popisují, a účely, k nimž se používají. Zatímco FAIR se ptá, zda lze objekt najít a opětovně použít, CARE se ptá, zda je opětovné použití legitimní, kdo drží oprávnění rozhodnout a jak je prospěch z opětovného použití rozdělen. Nejde o soupeřící výklady téže vlastnosti, nýbrž o odpovědi na otázky, které se neprotínají.

Komplementarita je skutečná, ale asymetrická způsobem, který bývá přehlížen. Soulad s FAIR lze posoudit na objektu samotném, kdežto soulad s CARE nelze posoudit bez účasti komunity, jež drží oprávnění. Posouzení podle CARE, které spočívá v opatření datové sady štítkem, přidání etického

prohlášení nebo citování principů v politice, proto není jejich implementací. Komunitní návštěví, jako jsou Traditional Knowledge Labels a Biocultural Labels, vyjadřují autoritu a očekávání komunity, nikoli právní oprávnění, a jejich uplatnění bez rozhodování komunity obrací jejich účel naruby (Local Contexts n.d.; Liggins, Hudson and Anderson 2021).

Druhá výhrada omezuje, jak daleko lze CARE generalizovat. Jeho ustanovení vycházejí ze specifické historie extrakce a vyvlastnění a přenesení jeho slovníku na jakoukoli datovou sadu týkající se jakékoli skupiny jej odpoutává od politického kontextu, v němž má oprávnění ke kontrole určitého držitele. Zde zaujaté stanovisko zní, že analytické otázky, které CARE klade — legitimita účelu, identifikovatelná autorita, rozdělení prospěchu —, platí široce, zatímco jeho normativní síla platí tam, kde existuje komunita s uznanou autoritou. S tímto rozlišením se v části III zachází jako se stavem proměnné, nikoli jako s věcí interpretace.

3.4 TRUST: posun objektu od dat ke kustodovi

Principy TRUST se týkají transparentnosti, odpovědnosti, orientace na uživatele, udržitelnosti a technologie a jejich objektem je repozitář (Lin et al. 2020). Posun objektu je podstatný, protože kvality, které pojmenovávají, nelze o datové sadě vůbec vypovídat: datová sada nemá model financování, plán nástupnictví ani personál. Certifikační schémata tyto kvality operacionalizují jako doložitelnou organizační kapacitu zahrnující governance, financování, personál, postupy, bezpečnost a plán řádného ukončení služby (CoreTrustSeal 2022). Data, která splňují FAIR ve službě bez takové kapacity, zůstávají exponovaná a tato expozice je na úrovni objektu neviditelná.

V analyzovaném materiálu jsou doloženy dva limity této vrstvy. Certifikace osvědčuje, že repozitář splňuje soubor požadavků; neosvědčuje, že je vhodný pro konkrétní datovou sadu, u níž jsou oborová vhodnost, podpora formátů, schopnost řídit přístup a dosah v komunitě samostatnými otázkami. A jak se služby federují a funkce sdílejí napříč poskytovateli, stává se doložení odpovědnosti za funkci vykonávanou jinde obtížnějším právě tam, kde na ní spočívá nárok na důvěryhodnost; samotné certifikační požadavky ukládají žadatelům deklarovat externě zajišťovaná a partnerská ujednání (CoreTrustSeal Standards and Certification Board 2025).

Tabulka 3-1 uvádí objekt, adresáta a otázku zodpovídanou každým ze tří souborů principů spolu s rozhodnutím, které každý z nich ponechává nevyřešené.

Tabulka 3-1. Objekt, otázka a zbytek tří souborů principů konstituujících normativní vrstvu.

Soubor principů	Objekt	Zodpovídaná otázka	Nevyřešené rozhodnutí
FAIR	datová sada a její metadata	lze objekt najít, zpřístupnit, kombinovat a opětovně použít stroji i lidmi	zda je opětovné použití legitimní; kdo nese náklady na dosažení daných vlastností
CARE	vztah mezi daty, lidmi a účelem	je použití legitimní a je prospěch spravedlivě rozdělen	jak se legitimita vyjádří v technickém režimu přístupu
TRUST	repozitář jako organizace	dokáže kustod dostát svým závazkům	zda se tento kustod hodí pro tuto datovou sadu

3.5 Kde se tyto tři soubory dostávají do konfliktu

Vrstva bývá předkládána jako soubor vzájemně se doplňujících pater a na úrovni deklarovaných hodnot jím také je. V provozu však tyto soubory generují konflikty, k jejichž vyřešení žádný z nich neobsahuje prostředky. Napříč doménami analyzovanými v této knize se opakují čtyři a jejich struktura je poučná: v každém z nich ukazují dva správně uplatněné požadavky k neslučitelným krokům.

Prvním je konflikt mezi otevřenou licencí a komunitním protokolem. Maximalizaci opětovného použití a respektování oprávnění ke kontrole nemůže naplnit táž licence a řešením zaznamenaným v korpusu je změna režimu přístupu, nikoli poznámka připojená k otevřeným datům. Druhým je konflikt mezi maximalizací opětovného použití a minimalizací údajů: opětovná použitelnost svědčí uchování bohatých záznamů, zatímco právo ochrany osobních údajů omezuje uchování na stanovený účel a přiznává práva na výmaz (European Parliament and Council 2016). Třetí vzniká tam, kde oborová přiměřenost a certifikovaná důvěryhodnost vybírají různé repozitáře, z nichž jeden drží komunitu a sémantiku a druhý doložitelnou schopnost přetrvat. Čtvrtý vzniká mezi interoperabilitou skrze obecná schémata a sémantickou přesností skrze doménové slovníky, kde je každý zisk vykoupen ztrátou v tom druhém.

Žádný z těchto konfliktů není nesprávným uplatněním principů. Každý vzniká ze správného uplatnění dvou z nich na tentýž případ, což je definiční znak nedourčené normativní vrstvy. Řešení vyžaduje vnější pravidlo, které konkurující si požadavky pro určitou třídu případů uspořádá podle priority, pojmenuje autoritu oprávněnou rozhodnout a zaznamená, jaké riziko po přijetí rozhodnutí zůstává. Taková pravidla dnes existují jako lokální praxe a jen zřídka jsou formulována tak, aby je bylo možné zkontrolovat či porovnat.

CHYBĚJÍCÍ DOKLAD

Zdokumentovaný soubor konfliktních případů z institucionální praxe, kódovaný podle typu konfliktu, přijatého řešení, rozhodující autority a zaznamenaného zbytkového rizika, umožňující stanovit rozložení typů konfliktu a otestovat, zda jsou řešení uvnitř instituce konzistentní

3.6 Nedourčenost jako mezera

Výsledek tohoto čtení lze uvést stručně. Normativní vrstva fixuje vlastnosti, které se od výsledků očekávají; nefixuje zobrazení podmínek případu na krok, jenž tyto vlastnosti vytváří. Neposkytuje uspořádání mezi vlastními komponentami, jsou-li v konfliktu, žádné přidělení rozhodovací autority, žádný výklad nákladů, za něž se každá vlastnost pořizuje, ani žádné kritérium odlišující realizované opatření od deklarovaného. Vyžadována jsou v důsledku toho tři samostatná posouzení — objektu podle FAIR, legitimacy a rozdělení prospěchu podle CARE a kustoda podle TRUST — a jejich složení do jediného rozhodnutí je ponecháno na instituci.

Tato nedourčenost není vadou souborů principů, které nebyly navrženy jako rozhodovací procedury, a argument této knihy nesměruje k jejich nahrazení. Směřuje k tomu, že vrstva nad nimi, v níž se podmínky zobrazují na rozhodnutí, nemá žádnou obdobnou kodifikaci a že její absence vysvětluje, proč se instituce s totožnými deklarovanými závazky v praxi rozcházejí. Srovnání napříč 23 doménami uváděné v této knize naznačuje, že tato divergence se soustřeďuje právě do čtyř výše uvedených typů konfliktu, což nasvědčuje tomu, že rozhodovací vrstva skromného rozsahu by pokryla většinu zbytku.

Tři kapitoly části I tak ustavily problém, způsobilost, která by nesla řešení, a bod, v němž se stávající normativní aparát zastavuje. Část II vykládá korpus, odstupňování důkazů a postup, jímž byly vytěženy rozhodovací

proměnné, aby bylo možné rámec předložený v části III posoudit proti materiálu, z něhož byl odvozen.

Č Á S T II

Výzkumný design a důkazní základna

Kapitola 4. Výzkumný design, korpus a analytická metoda

Předchozí část identifikuje problém, který nevyřeší žádný jednotlivý normativní zdroj. Rámce upravující výzkumná data jsou autoritativní ve vlastních doménách a vůči sobě navzájem převážně mlčí. Tento popis platí stejně pro principy FAIR, CARE a TRUST, pro právo ochrany osobních údajů, pro standardy dlouhodobého uchování i pro národní metodiku hodnocení výzkumu (Wilkinson et al. 2016; Carroll et al. 2020; Lin et al. 2020; European Parliament and Council 2016). Tato kapitola stanoví, jak se takový problém zkoumá, a nikoli pouze popisuje. Obhajované tvrzení zní, že integrovaný rámec lze odvodit z dokumentových zdrojů kódovacím postupem natolik explicitním, aby jej bylo možné zpochybnit, a že toto odvození je odlišné od jakéhokoli doložení, že výsledek ob stojí.

Toto tvrzení odpovídá na hlavní výzkumnou otázku knihy jen zčásti, a to záměrně. Hlavní otázka se ptá, jakými dimenzemi a proměnnými lze heterogenní normativní požadavky integrovat do jednoho rámce, který podpírá institucionální posouzení i rozhodování na úrovni projektu. Rámec předložený bez uvedeného odvození je názorem, jakkoli věrohodným; rámec validovaný bez uvedeného odvození nelze revidovat, selže-li validace, protože původ jednotlivých proměnných není znám. Design vyložený zde proto fixuje nejprve odvození a každé validující tvrzení odkládá do části IV.

4.1 Výzkumné otázky a jejich operacionalizace

Hlavní otázka se rozkládá do čtyř podřazených otázek, z nichž na každou lze odpovědět z identifikovatelné třídy zdrojového materiálu. První se ptá, které požadavky se opakují napříč normativními doménami pod odlišnými slovníky, a odpovídá se na ni mezidoménovým mapováním termínů. Druhá se ptá, kde se domény dostávají do konfliktu, a nikoli pouze liší, a odpovídá se na ni nalezením dvojic požadavků, jež nelze splnit současně bez rozhodnutí. Třetí se ptá, které vlastnosti datové činnosti určují, zda lze požadavek vůbec ověřit, a čtvrtá se ptá, jaké institucionální podmínky činí rámec provozuschopným.

Operacionalizace převádí každou otázku na pozorovatelný rys korpusu. Opakování je operacionalizováno jako výskyt funkčně ekvivalentního požadavku nejméně ve dvou vrstvách zdrojů, jedné normativní a jedné

aplikované. Konflikt je operacionalizován jako explicitní formulace kompromisu ve zdroji, nebo jako zdokumentovaný rozhodovací bod, v němž se jeden požadavek uvolňuje ve prospěch jiného. Ověřitelnost je operacionalizována jako přítomnost či absence pojmenovaného záznamu — logu, verze, schválení, metadatového pole —, jímž by bylo možné soulad dodatečně doložit.

Čtvrtá otázka, týkající se institucionálních podmínek, se čistě dokumentové operacionalizaci vzpírá. Dokumenty uvádějí, co instituce zamýšlí; nezakládají, co dělá. S touto asymetrií se v celém textu zachází jako s hranicí předloženého designu, nikoli jako s vadou, již by bylo možné argumentačně odklidit, a empirická vrstva potřebná k jejímu uzavření je označena tam, kde vyvstává.

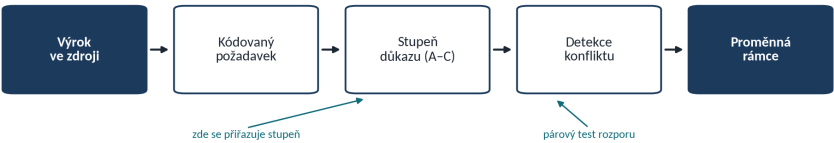
CHYBĚJÍCÍ DOKLAD

Data na úrovni institucí o skutečném provádění deklarovaných datových politik — doby odezvy služeb, četnost eskalací, podíl projektů s implementovaným ujednáním o dlouhodobém uchování —, umožňující stanovit, zda dokumentový závazek predikuje provozní praxi.

4.2 Přiřaditelná datová činnost jako jednotka analýzy

Jednotkou analýzy není datová sada ani instituce, nýbrž **datová činnost**: ohraničená operace s výzkumnými daty, k níž lze připojit odpovědnost, doložitelný záznam a zbytkové riziko. Sběr na základě nástroje souhlasu, transformace pojmenovaným skriptem, uložení do repozitáře, schválení řízeného přístupu i migrace dlouhodobě uchovávaného objektu jsou v tomto smyslu datovými činnostmi. Tato volba má důsledky, protože činí dimenze rámce vlastnostmi operací, a nikoli atributy objektů. Datovou sadu lze popsat jako FAIR; pouze o činnosti lze doložit, že ji někdo v určitém čase a podle určitého pravidla provedl.

Tato jednotka rovněž vysvětluje, proč analýza konverguje k záznamům, a nikoli k záměrům. Činnost, která nezanechá stopu, nelze po dokončení posoudit, ať byla její kvalita v daném okamžiku jakákoli. Kódování proto pojímá důkazní stopu jako kódovaný atribut každé činnosti, nikoli jako samostatné téma náležející auditu. Důsledek pro rámec je rozveden v kapitole 6, kde se ukazuje, že provenience je podmínkou, za níž se zbývající dimenze stávají vůbec posouditelnými.



Obr. 4-1. Analytický řetězec od formulace ve zdroji přes kódovaný požadavek k proměnné rámci, ukazující, ve kterém kroku se uplatňuje odstupňování důkazů a detekce konfliktů. Podklad: kódovací protokol popsany v oddílech 4.4 a 4.5. Vlastní zpracování autora.

4.3 Korpus, jeho vrstvy a jeho hranice

Korpus sestavený pro tuto studii je stratifikován podle funkce, kterou zdroj plní, nikoli podle jeho tématu. Rozlišují se čtyři vrstvy: závazné či normativní zdroje, oficiální metodiky a infrastrukturní specifikace, recenzovaná syntéza a aplikovaný institucionální materiál. Na stratifikaci záleží proto, že zdroje odlišné funkce opravňují k odlišným inferencím — nařízení zakládá povinnost, model zralosti zakládá srovnání, případová studie zakládá proveditelnost. Tabulka 4-1 uvádí jednotlivé vrstvy, pravidlo zahrnutí uplatněné na každou z nich a typ inference, k níž každá vrstva opravňuje.

Tabulka 4-1. Vrstvy analyzovaného korpusu s pravidlem zahrnutí a typem inference, k níž každá vrstva opravňuje.

Vrstva	Pravidlo zahrnutí	Inference, k níž opravňuje
Právo a standardy v platnosti	Použitelné na výzkumná data v evropském a českém prostředí; v platnosti k datu ověření	Povinnost; nevyjednatelné omezení
Oficiální metodiky, specifikace a infrastrukturní dokumentace	Vydané orgánem odpovědným za službu nebo specifikaci; identifikovaná verze	Referenční praxe; požadavek interoperability

Vrstva	Pravidlo zahrnutí	Inference, k níž opravňuje
Recenzované přehledy, syntézy a empirické studie	Systematický nebo scoping přehled, kontrolovaný či kvaziexperimentální design, nebo zavedená metodika	Zobecnitelná pravidelnost s uvedenými limity designu
Aplikovaný institucionální a národní materiál	Zdokumentovaná služba, politika nebo příručka pojmenované výzkumné organizace či národní infrastruktury	Pouze proveditelnost a implementační varianta

Vyloučení se uplatňuje ze tří důvodů. Zdroje, jejichž status je návrh nebo otevřená veřejná konzultace, se ponechávají, nikdy se však nepovažují za ustálené a při každé citaci se označují statusem. Zdroje uvádějící počty služeb, počty uživatelů nebo objem fondů bez data ověření jsou vyloučeny z jakéhokoli kvantitativního tvrzení. Propagační nebo dodavatelský materiál postrádající technickou specifikaci je vyloučen zcela. Překonané verze se ponechávají pouze tam, kde vysvětlují stabilní podkladový princip, a v takovém případě se uvádí verze i její omezení. Datem ověření časově proměnlivých zdrojů je pro celý korpus 27. července 2026.

Aplikovaná vrstva má dvě složky vyžadující samostatné pojednání. Českou národní složku tvoří vodítka, metadatová doporučení, dokumentace repozitářů, materiál k autorskoprávním rozhodnutím a jádrový metadatový model vydané národní technickou knihovnou a související infrastrukturou. Analyzuje se jako implementační vrstva pro české prostředí, nikoli jako mezinárodní norma. Mezinárodní složku tvoří zdokumentovaná portfolia služeb a kurikula výzkumných organizací a infrastruktur. Obojí se čte jako důkaz o tom, co instituce o své praxi publikovala, což je slabší důkaz než důkaz o praxi samotné.

CHYBĚJÍCÍ DOKLAD

Zdokumentovaný vyhledávací protokol pro aplikovanou vrstvu — dotazované databáze a registry, dotazovací řetězce, datum provedení, počet záznamů prověřených a vyloučených v každém kroku — postačující k tomu, aby nezávislý badatel korpus rekonstruoval. Předložený korpus byl sestaven účelovým výběrem proti kritériím v tabulce 4-1 a jeho úplnost proto není doložena.

4.4 Kódovací schéma a odvození dimenzí

Kódování postupuje ve třech průchodech týmž materiálem. První průchod vytěžuje formulace požadavků, definované jako jakákoli formulace, která specifikuje, co musí platit, aby byla datová činnost přijatelná; každá vytěžená formulace si podržuje svůj zdroj, svou vrstvu a svůj statusový štítek. Druhý průchod přiřazuje každý požadavek činnosti, kterou omezuje, a modu. Mody rozlišují, kdo rozhoduje, co se zaznamenává, jak je objekt popsán, za jakých podmínek je přístupný a jak dlouho přežívá. Třetí průchod seskupuje požadavky, jejichž modus a cíl se napříč vrstvami shodují, a právě tato seskupení, nikoli jakékoli předchozí pojmové schéma, tvoří dimenze rámce.

Ze třetího průchodu vzešlo šest dimenzí, které organizují část III. Tabulka 4-2 je pojmenovává, uvádí kódovací modus, z něhož byla každá odvozena, a určuje kapitolu, v níž je každá rozvinuta. Pořadí je analytické, nikoli sekvenční: governance se pojednává jako první, protože určuje, kdo smí řešit konflikty, jež pozdější dimenze odkrývají, a životní cyklus s proveniencí jako druhé, protože určují, zda lze jakékoli pozdější tvrzení ověřit.

Tabulka 4-2. Šest dimenzí integrovaného rámce s kódovacím modem, z něhož byla každá odvozena.

Dimenze	Kódovací modus	Rozvinuto v
Governance, role a odpovědnost za výsledek	Kdo rozhoduje, kdo je konzultován, kdo je informován	kapitola 5
Životní cyklus a provenience	Co se zaznamenává a o kterých agentech, činnostech a entitách	kapitola 6
Popis, metadata a interoperabilita	Jak je objekt popsán a jakým slovníkem	část III
Právní, etické a přístupové podmínky	Za jakých podmínek je objekt přístupný	část III
Kvalita, reprodukovatelnost a výpočetní spolehlivost	Jakým důkazem lze výsledek znovu získat	část III
Dlouhodobé uchování, důvěryhodnost repozitáře a udržitelnost	Jak dlouho objekt přežívá a pod čí zárukou	část III

Dva rysy tohoto postupu omezují, co jím lze podepřít. Kódování prováděl jediný analytik, takže žádná míra shody mezi kodéry neexistuje, a zejména

hranice mezi druhou a třetí dimenzí spočívá na interpretačním soudu, který by jiný analytik mohl vést jinudy. Dimenze se proto předkládají jako argumentované čtení korpusu, nikoli jako kategorizace s doloženou reliabilitou.

CHYBĚJÍCÍ DOKLAD

Nezávislé překódování náhodného vzorku nejméně 100 vytěžených formulací požadavků druhým analytikem s uvedeným koeficientem shody, umožňující stanovit, zda jsou hranice dimenzí v tabulce 4-2 reprodukovatelné, a nikoli idiosynkratické.

4.5 Syntéza, odstupňování důkazů a zacházení s konfliktem

Syntéza uvnitř každé dimenze sleduje pevnou čtyřprvkovou strukturu uplatněnou jednotně napříč všemi analyzovanými doménami: stabilní konsenzus, změna v období 2024–2026, otevřené otázky a limity důkazů a podmínky implementace. Právě jednotná struktura činí domény srovnatelnými, neboť každou z nich nutí deklarovat nejen to, co je ustálené, ale i to, kde jí docházejí důkazy. Srovnání napříč dvaceti třemi zde uváděnými doménami naznačuje, že autoritativní materiál je rozložen velmi nerovnoměrně. Dlouhodobé uchování, ochrana osobních údajů a metadatový popis jsou podepřeny standardy v platnosti, kdežto organizační uspořádání a účinnost institucionální podpory spočívají převážně na kazuistických zprávách.

Každé tvrzení přenesené do rámce si podržuje stupeň důkazu. Stupeň A označuje právní ustanovení, standard v platnosti, oficiální finální doporučení nebo široce přijímaný konsenzuální rámec. Stupeň B označuje recenzovanou studii, systematický přehled, vícepřípadové důkazy nebo zavedenou profesní metodiku. Stupeň C označuje jednopřípadovou zkušenost, nastupující komunitní praxi nebo doporučení, pro něž srovnávací důkazy chybějí. Stupeň se váže k důkaznímu opodstatnění, nikoli ke kvalitě konkrétní implementace, a uvádí se všude, kde je formulována proměnná rámce.

Konflikty mezi doménami jsou analyticky nejproduktivnějšími případy a neřeší se seřazením podle priority. Otevřená dostupnost stojí proti riziku zveřejnění, vyčerpávající provenience proti citlivosti logů, úplné dlouhodobé uchování proti nákladům. Tam, kde dva požadavky nelze splnit zároveň,

zaznamenává syntéza podmínky, za nichž je každá varianta proporcionální, a zbytkové riziko, jež každá z nich nechává neošetřené. Právě proto rámec produkuje rozhodovací struktury, a nikoli doporučení: doporučení konflikt zakrývá, kdežto podmíněná struktura jej odkrývá a pojmenovává autoritu příslušnou jej rozhodnout.

4.6 Vnitřní konzistence a oddělení konstrukce od validace

S vnitřní konzistencí se zachází jako s testovatelnou vlastností rámce, nikoli jako s estetickou ctností. Uplatňují se tři testy. Test přiřaditelnosti požaduje, aby každou proměnnou bylo možné připojit k pojmenované roli pro konkrétní datovou činnost. Test doložitelnosti požaduje, aby každá proměnná pojmenovala záznam, jímž by bylo možné její stav dodatečně doložit. Test uzavřenosti požaduje, aby každý konflikt identifikovaný v oddílu 4.5 byl směřován k dimenzi obsahující autoritu příslušnou jej rozhodnout. Proměnná, která v kterémkoli testu neobstojí, se buď přeformuluje, nebo se uvádí jako otevřený prvek rámce.

Validace, navržená v části IV, je samostatnou operací na samostatném souboru materiálu a žádná část z části III se nepředkládá jako validovaná zdroji, z nichž byla vybudována. Design má dvě složky. Expertní přezkum konstruktů odděluje technickou správnost od provozní použitelnosti a rozděluje přezkum mezi specialistu na správu výzkumných dat, repozitářového či technického specialistu a právního nebo etického posuzovatele, přičemž přijetí je podmíněno absencí neopravené věcné chyby. Aplikační test poté aplikuje rámec na autentický projektový materiál nejméně ve dvou oborech a posuzuje, zda jsou výsledná rozhodnutí zdokumentovaná, obhajitelná a dohledatelná k proměnným rámce.

Dvě vlastnosti tohoto designu zasluhují předběžné zdůraznění. Spokojenost účastníků se nepojímá jako důkaz validity, neboť rámec může být prožíván jako užitečný a přitom produkovat rozhodnutí, která v důkazním testu neobstojí. A posouzení aplikovaných výstupů je navrženo tak, aby probíhalo proti stanovené rubrice a provedl je posuzovatel zaslepený vůči původu výstupu, aby se oddělil příspěvek rámce od dovednosti osoby, která jej uplatňuje.

CHYBĚJÍCÍ DOKLAD

Provedení expertního přezkumu konstruktů — počet a oborové složení posuzovatelů, přezkoumávaná tvrzení a naložení s každou kritickou připomínkou — na podporu jakéhokoli tvrzení, že rámec je technicky správný.

CHYBĚJÍCÍ DOKLAD

Provedení aplikačního testu s 12–20 účastníky nejméně ze dvou oborů, včetně výchozího posouzení, zaslepeného bodování vytvořeného artefaktu podle rubriky a následné kontroly přenosu do vlastního projektu účastníka po šesti až osmi týdnech, na podporu jakéhokoli tvrzení, že rámec mění praxi, a nikoli znalosti.

4.7 Etické a právní podmínky a limity designu

Konstrukční fáze analyzuje publikované dokumenty a infrastrukturní specifikace a nezahrnuje žádné osobní údaje ani lidské účastníky; pro část III se proto žádné etické schválení neuplatní. Validační fáze účastníky a jejich projektový materiál zahrnuje a spočívá na třech podmínkách. Účast je dobrovolná a oddělitelná od jakéhokoli posuzování zaměstnání či studia. Projektový materiál se analyzuje ve formě, z níž nelze v reportování identifikovat projekt ani účastníka, a jakékoli zpracovávané osobní údaje zůstávají podřízeny použitelnému režimu ochrany osobních údajů (European Parliament and Council 2016). Tam, kde účastníci přinášejí materiál upravený dohodou s třetí stranou, design takový materiál vylučuje, místo aby usiloval o výjimku.

Vše následující kvalifikují tři limity. Korpus je vážen k evropskému a českému regulačnímu prostředí, takže právní a infrastrukturní dimenze se přenášejí hůře než dimenze governance a provenience. Časově proměnlivé zdroje byly ověřeny k jedinému datu a specifikace zde označené jako aktuální budou překonány; proměnné rámce jsou formulovány tak, aby takové změny přežily, ilustrace však nikoli. A dokumentové důkazy zakládají, co je požadováno a co je tvrzeno, nikdy však to, co je konáno, což je právě ta mezera, kterou značky v této kapitole identifikují a kterou má část IV začít uzavírat.

S fixovaným designem se část III obrací k rámci samotnému, počínaje dimenzí, jež určuje, kdo je příslušný rozhodnout, dostanou-li se zbývající dimenze do konfliktu.

Č Á S T I I I

Integrovaný rámec

Kapitola 5. Governance, role a odpovědnost za výsledek

Governance je první dimenzí integrovaného rámce proto, že rozhoduje otázku, kterou ostatní dimenze předpokládají. Každý konflikt, jež pozdější dimenze odkrývají — mezi otevřeností a rizikem zveřejnění, mezi náklady na dlouhodobé uchování a jeho důkazní hodnotou, mezi smluvním omezením a opětovným použitím —, ústí v rozhodnutí přijaté někým, kdo je k jeho přijetí oprávněn. Tvzení, jež tato kapitola obhájí, zní, že data stewardship je řízenou funkcí, a nikoli podpůrnou službou, a že vlastnosti, které jej činí řiditelným, lze přidělit každé datové činnosti jednotlivě, nikoli instituci jako celku. Tam, kde tyto vlastnosti přiděleny nejsou, jmenování specializovaných pracovníků přerozděluje práci, aniž by přerozdělovalo pravomoc.

S hlavní výzkumnou otázkou to souvisí specifickým způsobem. Má-li rámec podpírat posouzení stejně jako rozhodování, potřebuje proměnné, jejichž stav lze zjistit, aniž by se instituce dotazovala na své záměry. Organogramy, politiky a názvy pracovních pozic jsou slabým důkazem, protože zaznamenávají záměr, a nikoli praxi. Proměnné governance odvozené zde jsou proto formulovány jako vlastnosti činností a záznamů: kdo rozhodl, na jakém základě, v jakém limitu a s jakou stopou.

5.1 Rozhodovací vrstva a poradní vrstva

Nejdůležitější rozlišení v institucionální správě výzkumných dat odděluje vrstvu, která radí, od vrstvy, která rozhoduje. Poradní práce identifikuje problém, strukturuje varianty, shromažďuje dokumentaci, zaznamenává rozhodnutí a ověřuje, že navazující povinnosti byly naplněny. Rozhodovací pravomoc instituci zavazuje: schvaluje právní závěry, uděluje bezpečnostní výjimky, přijímá vysoké zbytkové riziko, podepisuje smlouvy, vydává etické schválení a alokuje zdroje. Analýza korpusu naznačuje, že se obě vrstvy v praxi běžně směšují a že právě toto směšování je mechanismem, jímž data steward začíná nést riziko, k jehož nesení role nemá mandát.

Nejde přitom o pochybení, nýbrž o strukturní mlčení. Výzkumník, který stojí před otázkou týkající se osobních údajů, dostane odpověď od kohokoli, kdo je právě k dispozici, a orientační posouzení pak funguje jako schválení, protože nic v procesu je neoznačuje za předběžné. Rámec proto zachází s

hranicí mezi poradou a rozhodnutím jako s explicitní proměnnou každé datové činnosti, vyjádřenou jako práh, za nímž věc přechází na pověřence pro ochranu osobních údajů, právní útvar, bezpečnostní funkci, etickou komisi nebo kurátora. Technická možnost se tím nepřevádí na právní, etické či vědecké oprávnění a hranice je právě tím, co obojí drží odděleně.

Pro návrh rolí z toho plynou tři důsledky. Hodnota data stewarda spočívá v identifikaci problému, ve strukturování variant, v koordinaci, v dokumentaci a v ověření navazujících kroků, přičemž nic z toho rozhodovací pravomoc nevyžaduje. Rozhodovací pravomoc uplatněná bez dokumentace je zpětně nerozlišitelná od toho, že žádné rozhodnutí přijato nebylo. A popis role, který uvádí pouze kompetence, aniž by uváděl limity, ponechává hranici k vyjednávání případ od případu pod časovým tlakem.

5.2 Odpovědnost za výsledek jako vlastnost datové činnosti

Rámec zachází s odpovědností za provedení, odpovědností za výsledek, konzultací a informováním jako se čtyřmi přiřaditelnými vlastnostmi každé datové činnosti, nikoli jako s popisem organizačního útvaru. Uplatněno na autoritativní kopii dlouhodobého senzorického měření to vede k oddělenému přiřazení pro kalibraci, zálohování, metadata, publikování a případné předání, vždy s pojmenovanou stranou odpovědnou za provedení a jedinou stranou odpovědnou za výsledek. Rozlišení mezi odpovědností za provedení a odpovědností za výsledek je právě tím, co přežívá delegování: práce se přesouvá, zodpovídání se nikoli. Přiřazení konzultace a informování explicitně určuje, kteří specialisté jsou zapojeni předem a které funkce se o výsledku dozvědí dodatečně.

Tento přístup řeší opakující se nejednoznačnost v literatuře o plánech správy dat. Je-li plán pojímán jako nástroj governance, činnosti, které plán nepokrývá, vypadávají z governance zcela a schválení plánu se stává zástupným ukazatelem institucionální kontroly, který takovou zátěž neunes. Je-li plán pojímán jako jeden z několika doložitelných záznamů, dokumentuje záměry, zatímco logy, schválení, přezkumy přístupu a záznamy o uložení dokumentují provedení. Rámec přejímá druhé čtení a plán správy dat nese dále jako důkazní proměnnou, nikoli jako proměnnou kontrolní.

Čitelnost přiřazení je rovněž vázána na okamžik činnosti, nikoli na dodatečnou rekonstrukci z korespondence. Přiřazení zaznamenané pouze v projektové žádosti je nedostupné tomu, kdo činnost provádí o tři roky

později. Z toho důvodu test přiřaditelnosti, formulovaný v oddílu 4.6, žádá, aby každou proměnnou bylo možné připojit k pojmenované roli pro konkrétní činnost v záznamu, k němuž se ten, kdo činnost provádí, dostane.



Obr. 5-1. Matice přiřazení pro jednu datovou činnost, ukazující role odpovědné za provedení, odpovědné za výsledek, konzultované a informované vedle vzniklého doložitelného záznamu a eskalačního práhu. Podklad: kódování governance popsané v oddílu 4.4. Vlastní zpracování autora.

5.3 Eskalace jako strukturní test modelu

Institucionální uspořádání správy výzkumných dat se konvenčně srovnávají jako organizační modely a korpus podporuje známou typologii centralizovaného, distribuovaného a federovaného uspořádání vedle projektově vnořené správy. Srovnání napříč korpusem naznačuje konvergenci k federovaným uspořádáním ve výzkumně intenzivních univerzitách, kde se ústřední metodika a infrastruktura kombinují s data stewardy ukotvenými v oborech (Science Europe 2021; EOSC Association 2024). Tato konvergence spočívá na institucionálních případových zprávách, nikoli na srovnávacích studiích výsledků, a rámec ji proto odstupňuje stupněm B. Tabulka 5-1 uvádí podmínky, za nichž je každé uspořádání proporcionální, a zbytkové riziko, jež každé z nich ponechává.

Tabulka 5-1. Organizační uspořádání institucionální správy výzkumných dat s podmínkami proporcionality a zbytkovým rizikem.

Uspořádání	Podmínka proporcionality	Zbytkové riziko	Stupeň důkazu
Centralizované	Menší nebo metodologicky homogenní instituce; služba v rané fázi zralosti	Odstup od oborové praxe; kapacitní úzké hrdlo v centru	B
Distribuované	Silné fakulty se zavedenou lokální podporou	Nekonzistentní pravidla a duplicitní úsilí	B
Federované	Multidisciplinární výzkumně intenzivní instituce	Koordinační náklady; nejasné financování sdílené vrstvy	B
Projektově vnořené	Jednotlivý vysoce rizikový nebo datově intenzivní projekt	Ztráta znalostí při skončení projektu	C

Typologie je nicméně slabší částí analýzy, protože jednotlivé modely určují méně, než literatura naznačuje. Fungující uspořádání odlišuje od nominálního přítomnost průchodné eskalační cesty: definovaný práh, pojmenovaný příjemce, očekávaný interval odezvy a záznam o výsledku. Instituce s federovaným organogramem a bez eskalační cesty si vede hůře než centralizovaná instituce, v níž se každá nevyřešená otázka dostane ke kompetentní autoritě ve známém intervalu. Rámec proto nese eskalaci jako samostatnou proměnnou a s organizační formou zachází jako s kontextovou podmínkou.

CHYBĚJÍCÍ DOKLAD

Srovnávací výsledková data napříč institucemi provozujícími odlišná organizační uspořádání — doba od žádosti k vyřešení, podíl eskalovaných žádostí, podíl eskalací uzavřených zaznamenaným rozhodnutím —, aby bylo možné otestovat, zda typ uspořádání predikuje výkon governance po zohlednění kvality eskalace.

5.4 Katalog služeb jako provozní podoba governance

Governance se stává pozorovatelnou v bodě, kdy nabývá podoby popsanych služeb. Katalog služeb pro každou službu specifikuje její vstupy, výstupy, vlastníka, očekávaný interval vyřešení a eskalační bod, v němž věc předává dál. Analýza uváděná zde naznačuje, že katalog je provozním vyjádřením modelu governance, nikoli administrativním doplňkem: konzultace k plánu,

bezpečnostní klasifikace, návrh metadat, podpora anonymizace a kurátorská kontrola uložení jsou odlišné služby s odlišnými vlastníky a jejich oddělený popis je právě tím, co činí vlastnictví přiřaditelným. Tam, kde katalog chybí, nabízí instituce místo služby kontaktní bod a hranice poskytovaného se vyjednává v každém jednotlivém případě.

Katalog rovněž disciplinuje vztah mezi nástroji a governance. Systematické mapování krajiny nástrojů ukazuje výraznou fragmentaci — přehled omezený na vědy o živé přírodě analyzoval více než 300 nástrojů (Aksenova et al. 2024) — a zavedení platformy samo o sobě pravomoc nepřiděluje. Nástroj nasazený nad nedefinovaným procesem reprodukuje nejednoznačnost tohoto procesu ve strojově čitelné podobě. Pořadí kroků, které důkazy podpírají, je proto opačné než pořadí obvyklé: definice služby, vlastnictví procesu, hranice odpovědnosti a eskalační cesta předchází volbě nástroje.

Katalog s sebou nese jeden charakteristický způsob selhání, který stojí za uvedení. Služby popsané bez očekávaného intervalu vyřešení umožňují neomezený odklad, a přitom zůstávají formálně v souladu, přičemž právě odklad je nejčastějším způsobem, jímž governance výzkumných dat selhává, aniž by bylo porušeno jakékoli pravidlo.

5.5 Kontinuita po skončení projektu a po odchodu osoby

Odpovědnost za výsledek připojená k projektu nebo k jednotlivci s ním zaniká. Korpus pojímá konec projektu, odchod pracovníka a ukončení služby jako tři instance jediného problému: převodu odpovědnosti za výsledek za data, rozhodnutí o přístupu, software, klíče a dokumentaci na pojmenovaného nástupce. Kontinuita je proměnnou governance, a nikoli archivní, protože odpovídá na otázku, kdo se zodpovídá dál, a nikoli na otázku, kde se bity nacházejí. Instituce, která nedokáže pojmenovat nástupce pro rozhodnutí o řízeném přístupu k ukončenému projektu, má nepřidělenou povinnost, ať je stav jejího repozitáře jakýkoli.

O tom, zda kontinuita drží, rozhodují dvě strukturní podmínky. První je rozpočtová: kurátorství a dlouhodobé uchování, které nejsou financovány z rozpočtu projektu, se stávají nefinancovanými povinnostmi útvaru, jenž je zdědí, a nefinancované povinnosti se plní výběrově. Druhá je dokumentační: nástupce, který nedokáže rekonstruovat základ dřívějších rozhodnutí o přístupu, buď všechny další žádosti odmítá, nebo je vyřizuje bez jakéhokoli základu, přičemž oba výsledky jsou selháními governance, nikoli úsudku.

Rámec nese obě podmínky jako proměnné, neboť každou z nich lze posoudit ze záznamů, které instituce již drží.

CHYBĚJÍCÍ DOKLAD

Longitudinální data o datových sadách, jejichž odpovědná osoba opustila instituci — podíl sad s dokumentovaným nástupcem, podíl sad, u nichž byly následné žádosti o přístup zodpovězeny, a doba do první nezodpověditelné žádosti —, aby bylo možné stanovit praktickou rychlost rozpadu nepřidělené odpovědnosti za výsledek.

5.6 Měření institucionální podpory bez počítání plánů

Měření institucionální podpory je bodem, v němž se rámce governance nejčastěji stávají sebestopratujícími. Počítání plánů správy dat, školení nebo konzultací měří činnost podpůrné funkce, a nikoli stav dat, a plán, který je napsán a nikdy proveden, počet zvyšuje a spolehlivost všeho navazujícího snižuje. Analýza korpusu naznačuje, že výsledkově orientované indikátory jsou dostupné a zároveň se užívají zřídka: včasnost, s níž je podpora zapojena vzhledem k datové činnosti, obnovitelnost dat při skutečném testu obnovy, úplnost metadat vůči stanovenému profilu, interval vyřízení žádostí o přístup a podíl projektů s provedeným ujednáním o dlouhodobém uchování. Tyto indikátory sdílejí vlastnost, kterou počítání nemá: každý z nich pojmenovává záznam, jenž může být v rozporu s tím, co instituce sama o sobě uvádí.

Výsledkové indikátory nicméně nesou známé zkreslení. Indikátory, které určují zdroje, přetvářejí chování, jež měří, a cíl vyjádřený jako podíl projektů se schváleným plánem produkuje schválené plány, a nikoli spravovaná data. Rámec to zmírňuje tím, že každý indikátor páruje s doložitelným záznamem, z něhož je odvozen, takže indikátor nelze naplnit bez toho, aby záznam existoval. Toto párování je zároveň tím, co spojuje governance s auditem: audit ověřující existenci dokumentů zakládá pouze to, že dokumenty existují, kdežto audit, který vzorkuje soubory, oprávnění, obnovy a metadata, zakládá, co bylo vykonáno (National Institute of Standards and Technology 2024).

CHYBĚJÍCÍ DOKLAD

Párování institucionální data o indikátorech založených na plánech a o výsledkových indikátorech pro tutéž množinu projektů, aby bylo možné otestovat tvrzení této studie, že se obojí rozchází a že počty plánů stav dat nadhodnocují.

5.7 Proměnné governance přenášené dále

Tabulka 5-2 uvádí proměnné, jimiž tato dimenze přispívá do integrovaného rámce, vždy se záznamem, z něhož se jejich stav čte. Formulace je záměrně strohá: každá proměnná je vlastností činnosti a každá proměnná pojmenovává stopu. Proměnné, které se nepodařilo připojit ke stopě, byly v oddílu 5.3 uvedeny jako kontextové podmínky, a nikoli připuštěny do rámce.

Tabulka 5-2. Proměnné governance v integrovaném rámci a jejich doložitelné záznamy.

Proměnná	Definice	Doložitelný záznam
Rozhodovací hranice	Práh, na němž činnost přechází od porady k zavazujícímu rozhodnutí	Popis role s uvedenými limity; záznam o eskalaci
Přiřazení odpovědnosti za výsledek	Jediná strana odpovědná za výsledek pro každou datovou činnost	Matice přiřazení vedená u projektového záznamu
Průchodnost eskalace	Existence pojmenovaného příjemce a očekávaného intervalu pro každý práh	Eskalační log s výsledkem
Definice služby	Vstupy, výstupy, vlastník a interval vyřešení každé služby	Položka katalogu služeb
Přiřazení kontinuity	Pojmenovaný nástupce pro data, rozhodnutí o přístupu, software a klíče	Předávací záznam; financované ujednání o dlouhodobém uchování
Výsledkové měření	Indikátory odvozené ze záznamů, nikoli z počtů dokumentů	Test obnovy, log žádostí o přístup, validační zpráva k metadatům

Přenos této dimenze omezují dvě hranice platnosti. Malé instituce neuživí diferencované role a proměnné jsou pro ně naplněny pojmenovanou externí kompetencí, nikoli vnitřními útvary; proměnným je lhotejné, zda je kompetentní strana zaměstnána lokálně. A ve federovaných infrastruktu-
rách může strana odpovědná za výsledek stát zcela mimo instituci, což činí smluvní specifikaci sdílené odpovědnosti jediným dostupným doložitelným záznamem (CoreTrustSeal 2022). Ani jedna z podmínek proměnné neoslabu-
je, obě však mění, kde se jejich důkaz nachází.

Dimenze rozvinutá zde určuje, kdo se zodpovídá za každou datovou činnost. Zda lze této zodpovědnosti dostát, závisí na tom, zda činnost vůbec zanechala záznam, a to je předmětem následující kapitoly.

Kapitola 6. Životní cyklus výzkumných dat a provenience

Dimenze governance rozhoduje, kdo se zodpovídá za datovou činnost. Nerozhoduje již, zda činnost zanechala cokoli, čím lze této zodpovědnosti dostát. Tato kapitola se ujímá druhé dimenze integrovaného rámce a obhazuje dvě spojená tvrzení: že životní cyklus výzkumných dat je strukturou vrstev, přechodů a režimů změny, a nikoli sledem fází, a že provenience je vlastností, na níž závisí posouditelnost všech ostatních dimenzí. Tam, kde odvození není zaznamenáno, zůstávají výroky o kvalitě, reprodukovatelnosti a právech tvrzeními, která žádné posouzení nemůže potvrdit ani vyvrátit.

Souvislost s hlavní výzkumnou otázkou je přímá. Rámec, který podpírá posouzení stejně jako rozhodování, vyžaduje proměnné, jejichž stav lze číst ze záznamů, jež činnost sama produkuje, a nikoli z toho, co projekt o vlastní praxi uvádí. Proměnné životního cyklu a provenience mají přesně tento charakter, protože každá z nich pojmenovává stopu: určenou autoritativní kopii, práh verzování, záznam o odvození, rozhodnutí o době uchování, předání kustodie. Dimenze rozvinutá zde tedy dodává důkazní substrát, který proměnné governance z kapitoly 5 předpokládají, ale samy generovat nedokážou.

6.1 Selhání fázového modelu

Model, který dělí práci s výzkumnými daty na plánování, sběr, zpracování, analýzu, publikování a archivaci, je didakticky úsporný a pro většinu projektů popisně nepravdivý. Výzkum je iterativní: předběžný výsledek přesměruje sběr, revidované pravidlo čištění zneplatní odvozenou tabulku, námitka recenzenta znovu otevře uzavřený model. Reprezentace přiměřená tomuto vzorci zachycuje opakování, větvení a návrat, což kurátorské modely životního cyklu v principu již umožňují (Higgins 2008). Vadou lineárního modelu není to, že zjednodušuje, nýbrž to, že zjednodušuje nesprávným směrem.

Směr této chyby lze určit přesně. Fázový model přiřazuje každému objektu pozici v čase, kdežto rozhodnutí o správě dat závisí na jeho statusu: co jej smí měnit, čeho je důkazem, jak dlouho je zapotřebí a která jeho kopie zavazuje. Dva soubory vytvořené v témž týdnu mohou obsazovat jednu fázi a

lišit se přitom ve všem, na čem pro rozhodnutí záleží. Rámec proto nahrazuje fázi dvojicí vrstvy a přechodu a s časem zachází jako s vlastností zaznamenanou uvnitř odvození, nikoli jako s organizující osou.

6.2 Vrstvy a autoritativní kopie

Výzkumná data se dělí do vrstev, které se liší důkazní funkcí i změnami, jež každá z nich připouští. Původní pořízení, pracovní kopie, harmonizované vrstvy, analytické datové sady a vydané balíčky se řídí odlišnými režimy: první se za žádných okolností nepřepisuje, střední vrstvy se upravují pod záznamem změn a poslední se v okamžiku citace zmrazuje. Uchovávat každý mezivýstup je nákladné a vzniká tak agregát, v němž žádná kopie není autoritativní; smazat je všechny znamená odstranit možnost auditu. Tabulka 6-1 uvádí jednotlivé vrstvy, funkci, kterou každá plní, a režim změny z této funkce plynoucí.

Tabulka 6-1. Datové vrstvy, jejich důkazní funkce a režim změny, který každá připouští.

Vrstva	Důkazní funkce	Přípustná změna	Hlavní důvod pro uchování
Původní pořízení	Zakládá, co bylo změřeno nebo pozorováno	Žádná; opravy se zaznamenávají v odvozené vrstvě	Nenahraditelnost
Pracovní kopie	Podpírá průzkum a opravy	Volná, pod záznamem změn	Regenerovatelnost ze vstupu a kódu
Harmonizovaná vrstva	Zakládá srovnatelnost napříč zdroji	Podle pravidel, se zaznamenaným mapováním	Náklady na opakovanou harmonizaci
Analytická datová sada	Podpírá publikované tvrzení	Zmrazena v okamžiku analýzy	Reprodukovatelnost výsledku
Vydaný balíček	Nese citaci a opětovné použití	Pouze nové vydání	Závazek persistence ze strany repozitáře

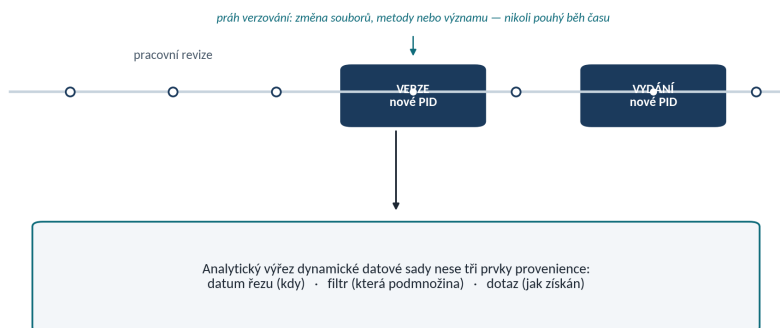
Rozhodnutím, které tabulka vynucuje, je určení autoritativní kopie pro každou vrstvu. Tam, kde žádná kopie určena není, dochází k rozbíhání tiše: kolují dvě harmonizované verze, obě jsou věrohodné a ani jedna není chybná způsobem, který by inspekce odhalila. Analýza korpusu naznačuje, že se toto určení explicitně činí zřídka a že jeho absence má větší důsledky než volba

úložiště nebo formátu, protože odstraňuje referenční bod, vůči němuž by bylo možné jakýkoli pozdější rozpor rozsoudit. Rámec proto nese autoritativní kopii jako proměnnou každé vrstvy, a nikoli jako atribut projektu.

6.3 Verze, vydání a pracovní revize

Tři pojmy se běžně směšují a lze je oddělit. Pracovní revize je změna vnitřní vůči vrstvě, verze je citovatelný stav objektu a vydání je verze, k níž se váže závazek persistence. Rozlišení se stává provozním v jediném bodě: v prahu, na němž změna generuje nový persistentní identifikátor, místo aby proběhla jako aktualizace pod identifikátorem stávajícím. Tento práh je rozhodnutím o tom, co identifikátor označuje, a nelze jej delegovat na nastavení repozitáře, aniž by o něm bylo implicitně rozhodnuto.

Dynamické datové sady činí práh nevyhnutelným. Průběžně aktualizovaná databáze neunesla stabilní citaci, není-li citovaný objekt specifikován jako časový řez, dotaz nebo zmrazené vydání, přičemž doporučení pro citování proměnlivých dat to řeší časovým razítkem a identifikací dotazu, nikoli zmrazením zdroje (Raubert et al. 2015). Reprodukovatelnost v takových případech nevyžaduje celou databázi, nýbrž dohledatelnost stavu, který analýza skutečně použila (National Academies of Sciences, Engineering, and Medicine 2019). Rámec proto zaznamenává datum řezu, filtr a dotaz jako prvky provenience analytické datové sady a s citací, jež je postrádá, zachází jako s neúplnou bez ohledu na to, zda se rozkládá.



Obr. 6-1. Vztah mezi pracovní revizí, verzí a vydáním u průběžně aktualizované datové sady, ukazující práh verzování, na němž se razí nový persistentní identifikátor, a tři prvky provenience — datum řezu, filtr, dotaz — připojené ke každému analytickému výtahu. Podklad: kódování životního cyklu a provenience popsané v oddílu 4.4. Vlastní zpracování autora.

6.4 Proč provenience podmiňuje ostatní dimenze

Provenience odpovídá na to, odkud objekt pochází, jakou činností, z jakých vstupů, prostřednictvím jakého agenta a v jakém čase. Je širší než historie souboru a užší než metodologický popis studie a model entita–činnost–agent ji reprezentuje neutrálně napříč ručním importem, skriptovanou transformací, kalibrací přístroje i formálním schválením (World Wide Web Consortium 2013). Specifikace balíčkování rozšiřují tutéž strukturu na přenositelné objekty nesoucí data, kód, workflow a kontext pohromadě (Soiland-Reyes et al. 2022). Aparát je tedy stabilní a dostupný; co literatura ponechává otevřené, je otázka, kolik z něj má daný projekt důvod instanciovat.

Argument předkládaný v této kapitole se týká postavení provenience mezi dimenzemi, nikoli její reprezentace. Tvrzení o kvalitě dat, o reprodukovatelnosti a o právech sdílejí společnou formu: každé z nich vypovídá něco o tom, jak objekt dospěl do stavu, v němž je. Tvrzení o kvalitě vypovídá, že na uvedené vstupy byly uplatněny uvedené postupy, tvrzení o reprodukovatelnosti vypovídá, že výsledek plyne z identifikovaných dat pod identifikovaným kódem a za identifikovaných podmínek, a tvrzení o právech vypovídá, že oprávnění vážící se ke vstupům dovolují to, co se nyní navrhuje. Žádné z těchto tří tvrzení není ověřitelné vůči objektu samotnému, neboť objekt nenese žádnou viditelnou známku vlastního odvození.

Asymetrie, která z toho plyne, je jádrem tohoto tvrzení. Provenienci lze zaznamenat i pro objekt, jehož kvalita je špatná, jehož výsledek se nereprodukuje a jehož licence zakazuje další šíření; záznam zůstává informativní a je vlastně tím, co uvedené nedostatky činí doložitelnými. Obrácené neplatí, protože sebevětší množství zajišťování kvality, práce na reprodukovatelnosti či licenční dokumentace nezakládá nic ověřitelného o objektu, jehož odvození není známo. Tato studie proto řadí provenienci před dimenze, které podmiňuje, a s přítomností záznamu o odvození zachází jako s předpokladem pro připuštění jakéhokoli posuzovacího tvrzení o ostatních dimenzích, nikoli jako s jedním kritériem mezi nimi.

6.5 Granularita úměrná riziku

Zaznamenat každou událost ve výpočetním pipeline je proveditelné a pro většinu výzkumu neúměrné. Praktickou otázkou je minimální důkazní řetězec, který korpus konzistentně rozkládá do šesti prvků: identifikátor a verze každého vstupu, uplatněný kód a konfigurace, čas, odpovědná role, identifikátor výstupu a kontrolní součet. Tabulka 6-2 uvádí varianty granularity, podmínku, za níž je každá proporcionální, a zbytkové riziko, jež ponechává. Odstupňování odráží, že varianty jsou zdokumentovány ve zprávách z praxe, a nikoli srovnány měřením výsledků.

Tabulka 6-2. Varianty granularity provenience, podmínky proportionality a zbytkové riziko.

Varianta	Podmínka proportionality	Zbytkové riziko	Stupeň důkazu
Úplná událostní provenience	Regulované zpracování nebo zpracování s vysokými důsledky	Objem a složitost zastírají rozhodující kroky	B
Provenience v kontrolních bodech	Běžný výzkumný projekt s definovanými body transformace	Drobné změny mezi kontrolními body zůstávají neviditelné	B
Automatizovaný log workflow	Výpočetní pipeline na stabilní platformě	Závislost na platformě; ztráta záznamu při migraci	B
Narativní soubor historie	Malý projekt s několika málo transformacemi	Nestrukturovaný a náchylný k opomenutí	C

Systematickou slabinou všech variant v tabulce jsou ruční kroky. Kopírování buněk mezi tabulkovými soubory, interaktivní filtrování a úpravy obrazu nezanechávají stopu, pokud není vytvořena záměrně, a automatizovaný log věrně zaznamenává tu část práce, která prošla pipeline, zatímco o části, jež jím neprošla, mlčí. Tam, kde se krok automatizaci vzpírá, nese jeho provenienci strukturovaný protokol uvádějící parametry, a nikoli pouze název softwaru. Srovnání napříč korpusem naznačuje, že rozdíl mezi zaznamenaným a skutečným odvozením je téměř výhradně soustředěn právě do těchto ručních přechodů.

CHYBĚJÍCÍ DOKLAD

Pro vzorek publikovaných analytických datových sad pokus o rekonstrukci pouze z uloženého záznamu provenience, měřící podíl obnovitelných transformačních kroků a podíl neobnovitelných kroků, jež byly provedeny ručně — na podporu tvrzení, že ruční přechody tvoří většinu rozdílu mezi zaznamenaným a skutečným odvozením.

Proveniencí má také právní složku, jež se od té technické často odděluje ke škodě obou. Datovou sadu kombinující zdroje různého původu lze dále šířit pouze tehdy, zůstávají-li licence, rozsah souhlasu a podmínky přístupu vážící se ke každému vstupu dohledatelné napříč uplatněnými transformacemi. Tam, kde tato stopa chybí, není odvozená datová sada tím pádem nezákonně držena, ale její další šíření se stává rozhodnutím přijatým bez informace, kterou rozhodnutí vyžaduje. Rámec proto zachází s licenčním manifestem jako s prvkem proveniencí, nikoli jako se samostatnou právní dokumentací.

6.6 Výběr k uchování a meze regenerace

Výběrové uchovávání spočívá na tvrzení, že mezivýstup, který lze regenerovat, není třeba uchovávat. Tvrzení je platné a jeho podmínka je náročná: regenerace předpokládá, že výpočet je deterministický, že kód a konfigurace jsou uchovány a že běhové prostředí zůstává dostupné nebo rekonstruovatelné. Uchovávat vstupy a kód namísto výstupů proto nahrazuje závislost na objemu úložiště závislostí na spustitelnosti prostředí a tato záměna je výhodná pouze tam, kde je první závislost skutečně řízena. Výběr k uchování je v rámci vyjádřen čtyřmi kritérii: důkazní hodnotou, náklady na náhradu, doloženou regenerovatelností a případnou závaznou povinností uchování.

Rozhodující váhu nese slovo doložená. Tvrzení, že výstup je regenerovatelný, je hypotézou o budoucím spuštění a periodické opakované spuštění je

jediným důkazem, který ji převádí na fakt o současném prostředí. Plánování dlouhodobého uchování v referenčním modelu to pojímá jako monitorovací povinnost archivu, a nikoli jako vlastnost uloženého objektu (Consultative Committee for Space Data Systems 2012). Výběr k uchování opřený o neověřená tvrzení o regenerovatelnosti se proto zaznamenává jako rozhodnutí přijaté za nekvantifikovaného předpokladu.

CHYBĚJÍCÍ DOKLAD

Opakované testy spuštění na definované množině uložených výpočetních datových sad v pevných intervalech po uložení, zaznamenávající podíl úspěšných regenerací a příčiny selhání — pro stanovení toho, jak rychle se rozpadá předpoklad regenerovatelnosti, o němž se výběrové uchování opírá.

6.7 Konec životnosti jako fáze životního cyklu

Modely životního cyklu končící uložením do archivu nedokážou vyjádřit události, jež datové fondy ničí nejspolehlivěji. Projekty končí, infrastruktury se vyřazují z provozu, pracovníci odcházejí a instituce se reorganizují, přičemž každá z těchto událostí převádí kustodii dat, rozhodnutí o přístupu, softwaru, šifrovacích klíčů a dokumentace, bez níž není nic z předchozího interpretovatelné. U dlouhodobě běžících observatoří a infrastrukturních fondů je ujednání o nástupnictví podmínkou tvrzení o udržitelnosti, a nikoli jeho administrativním dodatkem, a certifikace repozitářů právě proto pojímá plánování nástupnictví jako kritérium důvěryhodnosti (CoreTrustSeal 2022). Rámec proto reprezentuje konec životnosti jako fázi s vlastními rozhodnutími, vlastní stranou odpovědnou za výsledek a vlastním doložitelným záznamem.

Pojetí ukončení jako fáze má jeden důsledek, který stojí za samostatné uvedení. Převádí smazání z události, k níž dochází, když nikdo nevznese námitku, na rozhodnutí s uvedeným základem, autoritou a záznamem, tedy na týž status, jaký rámec přiděluje dlouhodobému uchování. Smazání a uchování se tím stávají symetrickými výsledky jediného rozhodnutí o výběru k uchování, a nikoli akcí a výchozím stavem. Asymetrie současné praxe — kde se o uchování rozhoduje, kdežto smazání nastává opotřebením — je mechanismem, jímž se fondy ztrácejí, aniž by o jejich ztrátě bylo kdy rozhodnuto.

6.8 Proměnné životního cyklu a provenience přenášené dále

Tabulka 6-3 uvádí proměnné, jimiž tato dimenze přispívá, vždy se záznamem, z něhož se jejich stav čte. Stejně jako v dimenzi governance každá proměnná pojmenovává stopu a vlastnosti, které se k žádné připojit nepodařilo, byly ponechány jako kontextové podmínky.

Tabulka 6-3. Proměnné životního cyklu a provenience v integrovaném rámci a jejich doložitelné záznamy.

Proměnná	Definice	Doložitelný záznam
Určení vrstvy	Vrstva, do níž objekt náleží, a režim změny z toho plynoucí	Položka datového registru s vrstvou a pravidlem změny
Autoritativní kopie	Kopie, která zavazuje, rozchází-li se kopie navzájem	Určení v registru s umístěním a kustodem
Práh verzování	Podmínka, za níž změna generuje nový persistentní identifikátor	Politika verzování; záznam o ražbě identifikátoru
Záznam o odvození	Vstupy, činnost, agent, čas, výstup a kontrolní součet pro každou transformaci	Graf provenience, log workflow nebo strukturovaný protokol
Protokol ručního kroku	Strukturovaný popis transformací provedených mimo pipeline	Laboratorní nebo transformační protokol s parametry
Stopa práv	Licence, rozsah souhlasu a podmínky přístupu nesené napříč odvozením	Licenční manifest propojený se záznamem provenience
Stav regenerovatelnosti	Zda lze neuchovaný výstup v současnosti reprodukovat	Datovaný výsledek testu opakovaného spuštění
Ujednání o nástupnictví	Pojmenovaný příjemce dat, rozhodnutí o přístupu, softwaru, klíčů a dokumentace	Dohoda o předání; financované ujednání o dlouhodobém uchování

Přenos této dimenze omezují dvě hranice platnosti. Tam, kde se výzkum provádí v důvěryhodných výzkumných prostředích, leží záznam o odvození uvnitř hranice, kterou vnější posouzení nepřekročí, a proměnná je pak naplněna atestovaným auditem provedeným uvnitř prostředí, nikoli inspektovatelným grafem. A v oborech, jejichž materiál není výpočetní, nese odvození analytického tvrzení vědecký aparát, a nikoli transformační logy, což mění formu záznamu, aniž by odpadl požadavek, aby nějaký existoval.

Dimenze ustavená zde určuje, co datová činnost zanechává. Zda je činnost předem utvářena rozhodnutími, vůči nimž lze pozdější záznamy kontrolovat, je otázkou, již se zabývá následující kapitola.

Kapitola 7. Plánování jako závazek: plány správy dat a strojově zpracovatelné plánování

Předchozí dimenze ustavuje, co datová činnost zanechává. Tato kapitola se ujímá bodu, v němž je činnost utvářena předem, a obhájí tvrzení o tom, čím plánování přispívá: plán správy dat má důkazní hodnotu pouze tam, kde zaznamenává rozhodnutí, jež lze později kontrolovat vůči stopám, které činnost produkuje. Takto pojatý plán není ani artefaktem souladu, ani nástrojem kontroly, nýbrž datovaným soupisem závazků, z nichž některé nelze revidovat, aniž by se ztratilo to, co měly zajistit. Kapitola identifikuje, které závazky mají tento nevratný charakter, a určuje, kdy strukturování plánu pro strojové zpracování přidává jistotu, místo aby formalizovalo nejasný postup.

Vztah k hlavní výzkumné otázce se odvíjí od posouditelnosti. Kapitola 5 ustavila, že dokumentový soulad a realizovaná praxe se rozcházejí a že počty dokumentů tento rozchod nedokážou zachytit. Plán nicméně zůstává jediným nástrojem, který fixuje záměry před činností, což z něj činí bod, v němž proměnné rámce nabývají svého počátečního stavu. Tato dimenze proto pojímá plán jako původ stavů proměnných a záznamy z kapitoly 6 jako důkazy, vůči nimž se tyto stavy později potvrzují nebo vyvracejí.

7.1 Dvojitě čtení plánu

Plán jako administrativní výstup se píše jednou, v okamžiku podání žádosti, a to v rejstříku kalibrovaném na posuzovatele, nikoli na projekt. Jeho výroky jsou obecné, protože obecnost je bezpečná: data budou bezpečně uložena, metadata budou následovat komunitní standardy, data budou uložena ve vhodném repozitáři. Takové výroky jsou nefalzifikovatelné a právě jejich nefalzifikovatelnost je důvodem, proč se píše, neboť o konkrétním závazku lze ukázat, že selhal, kdežto o obecném nikoli. Posuzování plánů podle strukturovaných kritérií konkrétnosti a proveditelnosti odpovídá právě na tento problém (Science Europe 2021).

Plán jako záznam rozhodnutí má jinou gramatiku. Každá položka pojmenovává systém, četnost, odpovědnou roli a záznam, jímž lze naplnění doložit,

čímž se příslib převádí na výrok s pravdivostní hodnotou. Analýza korpusu naznačuje, že právě tento převod, a nikoli jakýkoli nárůst délky či pokrytí, odděluje plány podírající pozdější ověření od plánů, které je nepodírají. Rámec přejímá druhé čtení a s prvním zachází jako se zdokumentovaným způsobem selhání, nikoli jako s méně kvalitní variantou téhož nástroje.

Toto rozlišení rovněž fixuje vztah plánu k sousedním dokumentům. Protokol upravuje vědecký postup, preregistrace fixuje hypotézy a analýzu, posouzení dopadu se týká rizika pro osoby a bezpečnostní plán hrozeb vůči systémům (European Parliament and Council 2016). Plán, který je vstřebá všechny, se stává neudržovatelným, a plán, který je ignoruje, jejich obsah nekonzistentně opakuje. Rámec proto pojímá plán jako propojující dokument identifikující, které rozhodnutí je zaznamenáno kde, a nikoli jako nádobu na každé rozhodnutí.

7.2 Závazky, které nelze přijmout zpětně

Většina plánovacích rozhodnutí je revidovatelná za mírné náklady: úložiště lze migrovat, metadata obohatit, repozitář před uložením změnit. Malý počet rozhodnutí revidovatelný není a právě jejich nevratnost dává plánování jeho důkazní sílu. Tabulka 7-1 tyto závazky identifikuje, uvádí, proč se každý z nich vzpírá zpětné nápravě, a pojmenovává záznam, v němž je závazek přijat.

Tabulka 7-1. Plánovací závazky, které nelze ustavit zpětně, s důvodem a se záznamem o závazku.

Závazek	Proč se vzpírá zpětné nápravě	Záznam o závazku
Rozsah souhlasu	Souhlas se uděluje pro uvedené účely před sběrem; pozdější opětovné použití nad tento rozsah nemůže výzkumník autorizovat	Informační list a formulář souhlasu, datované
Právní postavení vstupů	Podmínky přijaté při pořízení zavazují navazující užití; opětovné vyjednání po integraci je dostupné jen zřídka	Smlouva o užití dat; licenční manifest
Zachycení provenience	Transformaci, která nezanechala stopu, nelze po změně prostředí rekonstruovat	Konfigurace workflow; šablona protokolu
Třída úložiště a zabezpečení	Třída určuje, kde data smějí být od první kopie dále; reklasifikace neodčiní předchozí expozici	Rozhodnutí o klasifikaci se schválením

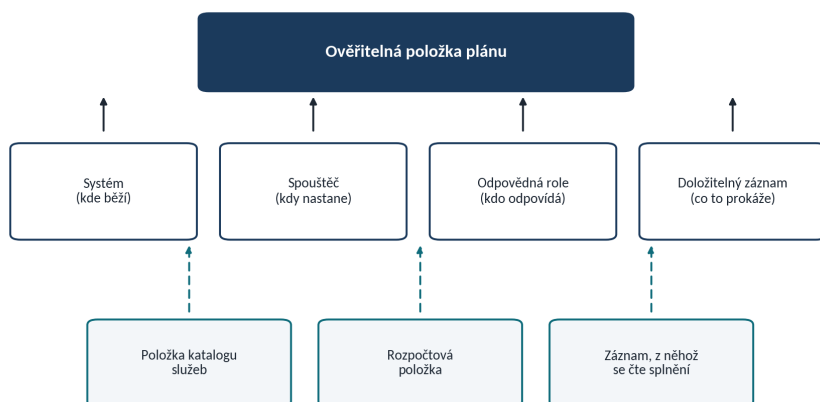
Závazek	Proč se vzpírá zpětné nápravě	Záznam o závazku
Politika identifikátorů a verzování	Již vydané citace nelze přesmérovat, aniž by se poškodil záznam	Politika verzování fixovaná před prvním vydáním

Závazky v tabulce 7-1 sdílejí jeden strukturní rys: každý z nich upravuje okamžik pořízení nebo prvního vydání a každý omezuje populaci pozdějších rozhodnutí, jež v době přijetí závazku nejsou známa. Nejzřetelnějším případem je rozsah souhlasu, neboť úzká formulace vylučuje opětovné použití, které nikdo nepředvídal, kdežto expanzivní nemusí být dost konkrétní, aby byla platná. Rámec toto napětí neřeší, protože náleží právní a etické dimenzi, zaznamenává však, kde má být napětí vyřešeno a která autorita je k jeho rozhodnutí příslušná. Plánování zde nepřispívá předvídatostí, nýbrž umístěním rozhodnutí do jediného okamžiku, kdy je lze ještě přijmout.

7.3 Anatomie ověřitelné položky plánu

Položka plánu, která podpírá pozdější posouzení, má nad rámec svého věcného obsahu čtyři komponenty. Pojmenovává systém nebo službu, v níž činnost probíhá, četnost či spouštěč jejího opakování, roli odpovědnou za výsledek a záznam, jímž by bylo možné její provedení doložit. Položka postrádající čtvrtou komponentu může být pravdivá, ale její pravdivost nelze zjistit bez dotazu na projekt, čímž se posouzení vrací k vlastnímu hlášení. Rámec proto pojímá důkazní komponentu jako definiční rys závazku, nikoli jako jeho zjemnění.

Tato struktura rovněž odhaluje závazky, jimž projekt nemůže dostát. Položka specifikující čtvrtletní kontroly integrity na úložné službě bez reportovacího rozhraní selže v bodě pojmenování svého záznamu, a selže během plánování, nikoli až při závěrečném přezkumu. Propojení položek plánu s rozpočtem a s katalogem služeb popsáním v oddílu 5.4 plní tutéž funkci pro zdroje, neboť závazek ke službě, kterou instituce neposkytuje, je detekovatelný již v okamžiku sepsání. Plánování je v tomto čtení testem proveditelnosti provedeným na papíře, tedy na nejlevnějším místě, kde jej lze provést.



Obr. 7-1. Struktura ověřitelné položky plánu, ukazující čtyři komponenty — systém, spouštěč, roli odpovědnou za výsledek, doložitelný záznam — a jejich provázání s položkou katalogu služeb, rozpočtovou linií a záznamem, z něhož se naplnění později čte. Podklad: kódování plánování popsané v oddílu 4.4. Vlastní zpracování autora.

7.4 Strojově čitelný a strojově zpracovatelný

Obě vlastnosti se běžně ztotožňují a přitom se liší způsobem, na němž záleží pro to, co lze zajistit. Strojově čitelný plán je takový, jehož obsah dokáže systém parsovat: strukturovaný dokument, export formuláře, soubor validní vůči schématu. Strojově zpracovatelný plán je takový, s jehož obsahem dokáže systém nakládat — validovat jej vůči pravidlu, předat jinou službou, použít jej ke spuštění žádosti nebo jej porovnat s pozorovaným stavem (Miksa et al. 2019). Čitelnost je vlastností formátu; zpracovatelnost je vlastností formátu spolu se sémantikou, která činí jeho hodnoty interpretovatelnými pro jiný systém než ten, jenž je vytvořil.

Toto rozlišení má provozní důsledek. Rozčlenění plánu do polí přináší čitelnost okamžitě a zpracovatelnost pouze tam, kde hodnoty pocházejí z autoritativních zdrojů, které jiné systémy rozpoznávají: osoba identifikovaná identifikátorem výzkumníka, poskytovatel podpory a grant záznamem v registru, repozitář svým registračním záznamem, publikovaná datová sada svým persistentním identifikátorem. Pro tento účel existují společné datové modely pro výměnu plánů, jež definují interoperabilní minimum, a nikoli úplný institucionální záznam; referenční specifikací je RDA DMP Common Standard, verze 1.1 (Miksa, Walk and Neish 2020). Rámec proto odděluje

interoperabilní minimum od lokálního rozšíření a jen prvním přiznává schopnost nést mezisystémové závazky.

Druhý důsledek se týká toho, co struktura nepřináší. Syntakticky validní, avšak zastaralá hodnota se beze změny propaguje do každého systému, který ji spotřebovává, a její struktura jí propůjčuje zdání ověření, jehož se jí nedostalo. Validace syntaxe tedy není validací obsahu a rámec pro každou interoperabilní hodnotu zaznamenává její zdroj, čas poslední aktualizace a stav ověření. Tato studie pojímá směšování struktury s pravdivostí jako hlavní riziko pro governance, jež automatizace přináší, a nikoli jako vadu konkrétních nástrojů.

7.5 Podmínky, za nichž automatizace přidává jistotu

Nástroje pro plánování sahají od textových šablon přes strukturované formuláře po znalostní systémy přizpůsobující otázky předchozím odpovědím a po plně integrovaná uspořádání, jež si vyměňují data s grantovými systémy, informačními systémy o výzkumu, úložnými službami a repozitáři. Tabulka 7-2 uvádí podmínky, za nichž je každá varianta proporcionální, a zbytkové riziko, jež každá ponechává. Odstupňování důkazů odráží, že srovnání spočívá na implementačních zprávách a specifikačních dokumentech, a nikoli na výsledkových studiích kvality či realizace plánů.

Tabulka 7-2. Implementační varianty plánování správy dat, podmínky proporcionality a zbytkové riziko.

Varianta	Podmínka proporcionality	Zbytkové riziko	Stupeň důkazu
Textová šablona	Malý samostatný projekt; žádný požadavek na institucionální výměnu	Slabá validace; položky nepropojené se záznamy	C
Strukturovaný formulář	Institucionální standardizace napříč mnoha projekty	Mechanické vyplňování; pole naplněná bez rozhodnutí	B
Znalostní systém	Složitě portfolio služeb; heterogenní třídy dat	Zátěž údržby znalostního modelu	B
Plně integrovaný strojově zpracovatelný plán	Zralá infrastruktura s vlastněnými integracemi	Propagace chybných hodnot; governance integrací	B

Podmínku oddělující jistotu od pouhé digitalizace lze formulovat nezávisle na zvolené variantě. Automatizace přidává jistotu tam, kde proces již má definovaného vlastníka, definované vstupy, definovaný výstup, definovaný způsob kontroly a stanovené podmínky aktualizace; chybí-li kterákoli z těchto náležitostí, reprodukuje automatizovaná verze nejednoznačnost vyšší rychlostí a se zdáním větší autority. Analýza napříč korpusem naznačuje, že se tento vzorec opakuje v každé doméně, v níž byla automatizace zkoumána, což nasvědčuje tomu, že jde o vlastnost automatizace ve slabě specifikovaných procesech, a nikoli o rys plánování správy dat.

Jedna další hranice se týká specificky plánování. Systém dokáže identifikovat riziko, vyžádat si informace, které rozhodnutí vyžaduje, a směřovat případ ke kompetentní autoritě, a to konzistentněji, než to činí lidský proces. Nedokáže určit právní základ, posoudit, zda je anonymizace dostatečná, ani zhodnotit vědeckou přiměřenost, protože každé z toho je úsudkem, jehož autoritu již přidělila dimenze governance. Automatizace tedy přenáší informaci a nikdy zodpovědnost, a uspořádání, v němž výstup systému funguje jako schválení, přemístilo rozhodnutí, aniž by přemístilo odpovědnost za výsledek.

CHYBĚJÍCÍ DOKLAD

Srovnání položek plánu vzniklých za použití textových šablon, strukturovaných formulářů a znalostních systémů u srovnatelných projektů, skórovaných podle rubriky konkrétnosti a proveditelnosti, aby bylo možné otestovat, zda nástroj ovlivňuje konkrétnost závazků, nebo pouze úplnost polí.

7.6 Revize jako důkaz, nikoli jako selhání

Plán, který je při ukončení projektu totožný s verzí podanou spolu se žádostí, dokládá jedno ze dvou: buď projekt, v němž nenastalo nic nepředvídaného, nebo plán, do nějž nikdo nenahlédl. Druhá možnost je podstatně častější a korpus uvádí týž vzorec napříč doménami. Revize spouštěná událostmi — dokončeným pilotem, změnou metody, novou smlouvou, změnou úložiště nebo personálu — zaznamenává skutečnou trajektorii projektu a historie verzí s uvedenými důvody je silnějším důkazním objektem než kterákoli jednotlivá verze. Rámec proto pojímá záznam o revizích jako proměnnou a absenci revize jako signál vyžadující vysvětlení.

Toto čtení řeší napětí v tom, jak s odchylkou nakládají poskytovatelé podpory a instituce. Je-li plán čten jako příslib, je odchylka nesouladem a racionální reakcí je psát plány natolik vágní, aby žádná odchylka nebyla detekovatelná. Je-li plán čten jako datovaný záznam rozhodnutí, je odchylka, která je zdůvodněná, schválená a verzovaná, důkazem řízení a způsobem selhání se stává nezdokumentovaná odchylka, nikoli odchylka jako taková. Struktury pobídek, jež obě čtení vytvářejí, se ostře liší a první z nich předpovídá právě tu obecnost, jaká se v podávaných plánech pozoruje.

CHYBĚJÍCÍ DOKLAD

Longitudinální záznam verzí plánu napříč kohortou projektů, kódovaný podle spouštěče revize, schválení a zdůvodnění, aby bylo možné otestovat tvrzení této studie, že revize spouštěná událostmi koreluje s realizovanými opatřeními silněji než úplnost plánu při podání.

7.7 Proč počítání plánů neměří správu dat

Indikátory uvádějící počet napsaných plánů nebo podíl projektů, jež plán mají, měří dosah povinnosti, a nikoli stav dat. Jejich necitlivost má tři oddělitelné zdroje. Plán může být úplný, aniž by kterákoli položka pojmenovávala záznam, takže úplnost je naplněna nefalzifikovatelnými výroky. Plán může být přesný při podání a během několika měsíců překonaný, takže jeho obsah popisuje stav, který již neplatí. A plán napsaný podpurným útvarem za projekt dokládá způsobilost tohoto útvaru, a nikoli jakékoli rozhodnutí přijaté projektem.

Co lze měřit místo toho, plyne z proměnných. Položka plánu pojmenovávající záznam umožňuje párované pozorování: závazek na jedné straně, záznam na druhé a určení, zda záznam existuje a zda ukazuje to, co položka tvrdila. Toto srovnání je oproti počítání nákladné a jeho cena je důvodem, proč se provádí zřídka, a nikoli důvodem, proč by nebylo dostupné. Tato studie pojímá párované srovnání jako minimální měření na úrovni výsledku pro dimenzi plánování a počty plánů jako statistiku pokrytí bez jakékoli výsledkové interpretace.

7.8 Proměnné plánování přenášené dále

Tabulka 7-3 uvádí proměnné, jimiž tato dimenze přispívá. Každá z nich je vlastností plánu jako objektu a každá pojmenovává záznam, vůči němuž se její stav čte.

Tabulka 7-3. Proměnné plánování v integrovaném rámci a jejich doložitelné záznamy.

Proměnná	Definice	Doložitelný záznam
Specifičnost položky	Zda každá položka pojmenovává systém, spouštěč, roli odpovědnou za výsledek a záznam	Položka plánu, jak je napsána
Pokrytí nevratných závazků	Zda jsou závazky z tabulky 7-1 fixovány před činností, kterou upravují	Datovaná rozhodnutí o souhlasu, smlouvě, klasifikaci a verzování
Provázání plánu	Identifikace toho, které rozhodnutí je zaznamenáno ve kterém sousedním dokumentu	Křížové odkazy na protokol, posouzení dopadu a bezpečnostní plán
Interoperabilní minimum	Hodnoty pocházející z autoritativních registrů, se zdrojem a časem aktualizace	Strukturovaný export plánu; záznam o rozklíčování identifikátoru
Stav ověření hodnoty	Zda byla vyměřovaná hodnota zkontrolována na aktuálnost, nikoli pouze na syntaxi	Validační log s časovým razítkem
Stopa revizí	Verze spouštěné událostmi, s důvodem a schválením	Historie verzí plánu
Důkaz realizace	Párované srovnání každého závazku s jeho pojmenovaným záznamem	Závěrečná verze plánu s uvedenými odchylkami

Dimenzi kvalifikují dvě hranice platnosti. Konsorciální a doktorská prostředí vyvolávají otázku, zda je namíste jeden plán, plán s přílohami za dílčí projekty, nebo samostatné propojené plány, a korpus ji řeší podle toho, nakolik se data, role, smlouvy a infrastruktura mezi jednotlivými částmi liší, a nikoli podle velikosti konsorcia. A publikování plánů se dotýká bezpečnostního a smluvního rizika, na něž odpovídá redigovaná veřejná verze, zatímco objektem posouzení zůstává plná interní verze; výše uvedené proměnné se čtou z interní verze a posouzení provedené pouze na veřejných verzích měří něco užšího, než jak se jeví.

Plánování fixuje závazky ohledně popisu, kvality a přístupu dříve, než data existují. Zda tyto závazky vedou k datům, jež dokáže jiný výzkumník interpretovat a opětovně použít, je předmětem následující kapitoly.

Kapitola 8. Popis, kvalita a podmínky opětovného použití

Dvě předchozí dimenze ustavují, kdo se zodpovídá za datovou činnost a jaký záznam po ní zůstává. Ani jedna nerozhoduje, zda datové sadě porozumí někdo, kdo ji nevytvořil. Tato kapitola se ujímá popisu a kvality jako dimenze určující interpretovatelnost a obhájí dvě tvrzení: že vhodnost k použití je vztahem mezi datovou sadou a deklarovaným účelem, a nikoli vlastností, kterou datová sada má, a že popis vznikající v průběhu činnosti je systematicky silnější než popis rekonstruovaný dodatečně. Druhé tvrzení je uvedeno spolu se svými hranicemi platnosti, neboť jde o empirické zobecnění, nikoli o nutnou pravdu.

Souvislost s hlavní výzkumnou otázkou spočívá v tom, že popis a kvalita jsou dimenzí, v níž jsou náklady na správu dat nejviditelnější a nejspornější. Požadavky na popis se snadno množí a výzkumníci je prožívají jako neohraňované, což činí z hranice mezi tím, co popis pokrývá, a tím, co ponechává stranou, rozhodnutí, jež rámec podpírá, místo aby se mu vyhýbal. Proměnné odvozené zde proto vyjadřují popis jako ohraničenou povinnost definovanou zamýšleným užitím a zamýšleným uživatelem a kvalitu jako deklarovanou zbytkovou chybu vůči stanovenému účelu.

8.1 Metadata jako nositel interpretovatelnosti

Metadata nejsou jedním popisem, nýbrž několika, adresovanými různým čtenářům a odpovídajícími na různé otázky. Citační a vyhledávací metadata objekt identifikují a spojují jej s osobami, poskytovateli podpory, projekty a souvisejícími výstupy; oborová metadata popisují proměnné, měření, vzorky, přístroje a metody; technická metadata podpírají dlouhodobé uchování; metadata provenience vysvětlují odvození. Obecné schéma slouží první vrstvě dobře a není zamýšleno tak, aby neslo druhou, kdežto oborový standard nese druhou a do vyhledávání napříč doménami se integruje slabě (DataCite Metadata Working Group 2024). Pojímat kterékoli z obou jako samo o sobě postačující je opakovanou chybou, již má rámec zabránit.

Vrstvy se liší nejen obsahem, protože se liší i v tom, kdo nese náklady opomenutí. Chybějící citační metadata se zjistí okamžitě, neboť objekt nelze najít ani ocitovat, a opomenutí se tak samo napravuje. Chybějící oborová

metadata zjistí až ten, kdo data opětovně používá, často po letech, a informace potřebná k jejich doplnění může být v té době již neobnovitelná. Povinnosti popisu vymáhané napříč vrstvami jednotně proto soustřeďují úsilí tam, kde má selhání nejmenší důsledky, což je jeden z mechanismů, jimiž popis poháněný souladem produkuje formálně úplné a obsahově prázdné záznamy.

8.2 Hranice mezi jádrovým a oborovým profilem

Aplikační profil řeší otázku vrstvení tím, že vybírá prvky z jednoho či více standardů a pro každý z nich fixuje, zda je povinný, v jakém formátu se vyjadřuje, z jakého slovníku pocházejí jeho hodnoty a podle jakého pravidla se validuje. Bez profilu nabývají shodně pojmenovaná pole napříč projekty rozbíhavých významů a záznamy se stávají formálně úplnými, a přitom vzájemně nesrovnatelnými. Profil je tedy artefaktem, v němž se hranice mezi společným jádrem a oborovou vrstvou skutečně vede, a její vedení je rozhodnutím s vlastníkem, nikoli technickou konfigurací. Tabulka 8-1 uvádí varianty kombinování obecného a oborového popisu spolu s podmínkou, za níž je každá proporcionální.

Tabulka 8-1. Varianty kombinování obecných a oborových metadat s podmínkami proportionalit a zbytkovým rizikem.

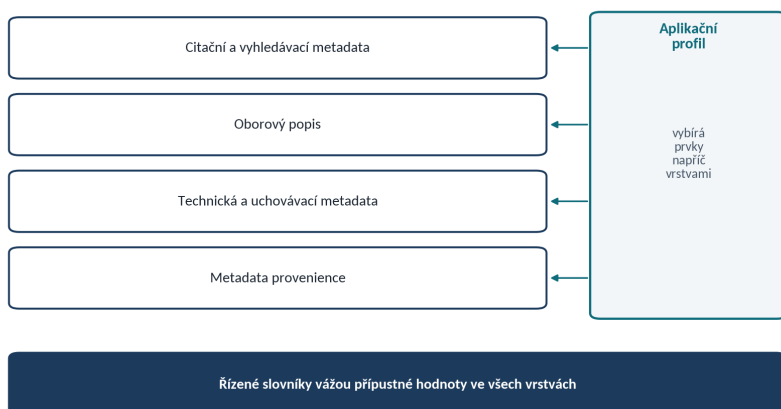
Varianta	Podmínka proportionality	Zbytkové riziko	Stupeň důkazu
Pouze obecné schéma	Uložení v generalistickém repozitáři s připojenou dokumentací	Vědecký detail nedostatečný pro opětovné použití	B
Pouze oborový standard	Uložení ve specializovaném repozitáři sloužícím jedné komunitě	Slabé napojení na vyhledávání napříč doménami	B
Dvojí vrstva: bohatý záznam a obecný export	Opětovné použití očekávané uvnitř oboru i mimo něj	Údržba převodníků a ztráta při exportu	B
Volný textový popis	Pouze doplňkové vysvětlení	Není interoperabilní; není strojově zpracovatelný	C

Dvoji vrstva nese náklad, který literatura identifikuje, ale kvantifikuje jen zřídka. Analýza čtrnácti převodníků z oborových schémat pro výzkumná data do obecného slovníku Schema.org ukazuje, že popisné prvky se mapují dobře, kdežto oborově specifické, právní a strukturní prvky zůstávají nenamapované nebo namapované nekonzistentně (Wu et al. 2023), přičemž ztracené prvky jsou zpravidla ty, jež nesou vědecký význam proměnných. Ponechat původní záznam vedle exportu je uspořádáním, které korpus konzistentně uvádí jako výhodnější, a to na základě úvahy, že ztrátový export je obnovitelný, kdežto ztrátový originál nikoli. Rámec proto pojímá autoritativní popis jako oborový záznam a obecnou vrstvu jako odvozený pohled, což je souběžné se zacházením s autoritativními kopiemi v oddílu 6.2.

8.3 Řízené slovníky jako podmínka interoperability

Popis se stává srovnatelným za hranicí projektu jen tam, kde jeho hodnoty pocházejí z termínů existujících nezávisle na projektu. Řízený slovník hodnoty sjednocuje, kdežto ontologie navíc reprezentuje vztahy mezi pojmy ve formě podpírající inferenci; druhá není automaticky výhodnější než první, neboť její nároky na údržbu jsou podstatně vyšší. Použitelným činí obojí soubor podmínek nezávislých na vyjadřovací síle: persistentní identifikátory pro každý pojem, dokumentace, explicitní licence, komunita, která zdroj udržuje, a stanovený postup pro řízení změn (World Wide Web Consortium 2009). Slovník, který tyto podmínky nesplňuje, produkuje terminologickou shodu rozpouštějící se v okamžiku, kdy zanikne hostitelská služba.

Pro rámec z toho plynou dva důsledky. Lokální termíny jsou ve specializované práci nevyhnutelné a nejsou vadou, ale lokální termín bez persistentního identifikátoru a bez definice je pro toho, kdo data opětovně používá, nerozlišitelný od překlepu. A vícejazyčná prostředí vyžadují paralelní názvy, popisy a jazykové značky, nikoli překlad hodnot, neboť přeložené hodnoty tiše porušují korespondenci, k jejímuž zaručení řízené slovníky existují. Rámec proto zaznamenává užití slovníku jako dvojici proměnných: který slovník upravuje které pole a zda se jeho termíny rozklíčují na stabilní identifikátory.



Obr. 8-1. Vrstevnatá architektura popisu, ukazující citační a vyhledávací metadata, oborový popis, technická metadata a metadata pro dlouhodobé uchování a metadata provenience, s aplikačním profilem vybírajícím prvky napříč vrstvami a řízenými slovníky vážícími přípustné hodnoty. Podklad: kódování popisu popsané v oddílu 4.4. Vlastní zpracování autora.

8.4 Zajišťování kvality jako návrh a kontrola kvality jako detekce

Zajišťování kvality a kontrola kvality jsou ve vodítkách oboru oddělovány a ve většině praxe splývají. Zajišťování působí dříve, než data existují: návrh přístroje, návrh formuláře, kalibrace, školení, standardní operační postupy a validační pravidla vestavěná do sběru. Kontrola působí až poté: kontroly rozsahu, detekce duplicit, inspekce rozdělení, dvojité pořízení a nezávislý přezkum. Na rozlišení záleží proto, že obojí má odlišnou nákladovou strukturu i odlišné meze a že kontrola nedokáže obnovit informaci, kterou by zajišťování bylo zachovalo.

Tuto asymetrii stojí za to formulovat přesně. Kontrola zjistí, že hodnota je nevěrohodná; nezakládá, jaká hodnota to byla, a tam, kde je původní pozorování neobnovitelné, se zjištěná chyba stává chybějícím údajem, nikoli údajem opraveným. Zajišťování ztrátě předchází, vyžaduje však rozhodnutí před sběrem, tedy v okamžiku, kdy je budoucí analýze rozuměno nejhůře. Tato studie pojímá rovnováhu mezi obojím jako proměnnou podmíněnou nahraditelností pozorování: tam, kde je opakované pozorování možné, je kontrola rozumnou náhradou zajišťování, a tam, kde je pozorování jedinečné, nikoli.

Postup opravy je bodem, v němž se kvalita a provenience potkávají. Původní hodnoty se nepřepisují; oprava se provádí v odvozené vrstvě a nese svůj důvod, datum a autora, což zachovává schopnost odlišit chybu zápisu od platné extrémní hodnoty. Chybějící hodnoty jsou stejně tak různorodé — neměřeno, selhání přístroje, odmítnutí, neaplikovatelné, pod mezí detekce — a jediný nediferencovaný kód zahazuje rozlišení, na němž analýza často závisí. Rámec proto pojímá kód důvodu u chybějících hodnot a záznam oprav jako prvky popisu, a nikoli jako vnitřní pracovní praxi.

8.5 Vhodnost k použití jako vlastnost relativní k účelu

Dimenze kvality, jako je přesnost, úplnost, konzistence, aktuálnost, reprezentativnost, validita a dohledatelnost, mají smysl pouze tehdy, jsou-li vypovídány o nějakém užití. Řada měření přiměřená k detekci sezonního trendu může být nepřiměřená ke kalibraci modelu citlivého na krátkodobou varianci a žádná vlastnost samotné řady oba případy neodliší. Vhodnost k použití je proto vztahem a výrok o kvalitě, který vynechává účel, vůči němuž byla posouzena, sděluje méně, než jak se jeví. Rámec žádá, aby byl účel deklarován spolu s výrokem o kvalitě, což je právě to, co převádí výrok z tvrzení o datech na tvrzení, že může ten, kdo data opětovně používá, testovat vůči jinému účelu.

Důsledek se týká zbytkové chyby. Každá datová sada nese chybu, o níž její tvůrci vědí a již se rozhodli neodstranit, buď proto, že odstranění je nemožné, nebo proto, že je vzhledem k deklarovanému účelu neúměrné. Uvést tuto zbytkovou chybu explicitně — její velikost tam, kde je kvantifikovatelná, její směr tam, kde je znám, a její mechanismus tam, kde je identifikován — je informativnější než tvrzení o přesnosti a je to zároveň falzifikovatelné, což tvrzení o přesnosti zpravidla není. Tabulka 8-2 ilustruje vztah mezi účelem, dimenzemi, jež se stávají rozhodujícími, a podobou, kterou výrok o zbytkové chybě nabývá.

Tabulka 8-2. Podoba výroků o kvalitě relativní k účelu, s dimenzemi, které každý účel činí rozhodujícími.

Deklarovaný účel	Rozhodující dimenze	Podoba výroku o zbytkové chybě
Ověření publikovaného výsledku	Dohledatelnost, konzistence	Odchytky mezi analytickou datovou sadou a uloženou verzí

Deklarovaný účel	Rozhodující dimenze	Podoba výroku o zbytkové chybě
Opětovné použití v sekundární analýze	Reprezentativnost, úplnost	Mechanismus výběru a struktura neodpovědí
Integrace do kombinované datové sady	Validita, srovnatelnost jednotek a kategorií	Pravidla mapování a nevyřešené konflikty kategorií
Dlouhodobá analýza trendu	Aktuálnost, konzistence v čase	Změny metody a přístrojů a jejich účinek na řadu

Samostatnou zmínku zasluhuje agregace, protože se běžně čte chybně. Kombinování zdrojů kvalitu nezvyšuje, neboť kombinovaná datová sada dědí nejslabší záruky svých vstupů, pokud harmonizace nezaznamenává původní jednotky, mapování kategorií a nejistotu. Objem nekompenzuje systematické zkreslení, které s růstem vzorku přetrvává nezmenšené a s tím, jak se standardní chyby zmenšují, se stává hůře zjištělným. Rámec proto zaznamenává harmonizační rozhodnutí jako prvky popisu a s nezaznamenanou harmonizací zachází jako s vadou kvality bez ohledu na kvalitu zdrojů.

8.6 Proč popis vznikající v průběhu předčí popis pořízený dodatečně

Ústřední argument této dimenze se týká okamžiku, v němž popis vzniká. Popis vytvořený během činnosti pořizuje agent, který informaci drží, v kontextu, kde jsou hodnoty přímo pozorovatelné a kde lze chyby opravit vůči stavu práce. Popis vytvořený dodatečně je rekonstrukcí a rekonstrukce podléhá třem selháním, jež popis pořizovaný souběžně nepostihují: informace držená pouze v pracovní paměti se rozpadá, o tom, co je vybaveno, rozhoduje výraznost, a nikoli potřeba toho, kdo data opětovně použije, a nezůstává žádná stopa, vůči níž by bylo možné rekonstrukci zkontrolovat. Tato studie předkládá tvrzení, že retrospektivní popis je z těchto důvodů systematicky slabší, a nikoli proto, že se píše pod časovým tlakem na konci projektu, byť tomu tak obvykle rovněž je.

Tvrzení je empirické a důkazy pro ně jsou v oboru nepřímé. Zdokumentován je výsledek — uložené sady s řídkým oborovým popisem a žádosti o vyjasnění ze strany těch, kdo je opětovně používají, k bodům, jež uložený záznam nechává nezodpovězené —, a nikoli mechanismus spojující okamžik vzniku s tímto výsledkem. Rámec proto nese toto tvrzení ve střední síle a s

časováním popisu zachází jako s rozhodovací proměnnou, a nikoli jako s ustáleným zjištěním.

CHYBĚJÍCÍ DOKLAD

Párované srovnání úplnosti popisu a žádostí o vyjasnění ze strany těch, kdo data opětovně používají, mezi datovými sadami popsány během sběru a datovými sadami popsány při uložení, při kontrole oboru, složitosti dat a repozitáře, aby bylo možné otestovat, zda časování vzniku popisu predikuje náklady opětovného použití.

Tři hranice platnosti toto tvrzení omezují a jsou jeho součástí. Tam, kde instrumentované pipeline automaticky vydává parametry, verze a konfigurace, popis vzniká v průběhu, i když je dokument, který jej předkládá, sestaven až při uložení, a tvrzení se neuplatní, protože relevantní záznam byl souběžný. Tam, kde je datová sada malá, stabilní a pořízená jedním operátorem v krátkém intervalu, je rozpad zanedbatelný a rozdíl mezi oběma okamžiky se blíží nule. A tam, kde použitelný standard popisu během sběru neexistoval, není retrospektivní popis podle současného standardu slabší než alternativa, neboť onou alternativou je nepopsat nic. Tvrzení se tedy týká okamžiku, v němž je informace zachycena, a nikoli okamžiku, v němž je dokument napsán, přičemž směšování obojího je nejpravděpodobnějším zdrojem zdánlivých protipříkladů.

Praktickou váhu dává tomuto tvrzení vztah mezi kvalitou popisu a náklady opětovného použití. Každý prvek, který záznam vynechává, přesouvá práci na toho, kdo data opětovně používá a kdo pak musí chybějící informaci buď odvodit, nebo si ji vyžádat od tvůrce, jenž může být nedostupný, anebo od opětovného použití upustit. Tento přesun je v účtech produkujícího projektu neviditelný a neobjevuje se v žádném indikátoru správy dat, což znamená, že náklady tenkého popisu nese zcela někdo jiný než útvar rozhodující o tom, kolik popisu vznikne. Rámec pojímá tuto externalitu jako strukturní vysvětlení trvalé nedopopsanosti, a nápravu proto umísťuje do institucionálních požadavků při uložení, nikoli do apelů adresovaných tvůrcům.

CHYBĚJÍCÍ DOKLAD

Měření úsilí těch, kdo data opětovně používají — doba do první analýzy, počet žádostí o vyjasnění, podíl opuštěných pokusů o opětovné použití — v závislosti na úplnosti aplikačního profilu, stanovující velikost nákladů přesunutých neúplným popisem.

8.7 Proměnné popisu a kvality přenášené dále

Tabulka 8-3 uvádí proměnné, jimiž tato dimenze přispívá. Každá pojmenovává záznam a každou lze číst, aniž by se produkující projekt dotazoval na své záměry.

Tabulka 8-3. Proměnné popisu a kvality v integrovaném rámci a jejich doložitelné záznamy.

Proměnná	Definice	Doložitelný záznam
Aplikační profil	Požadované prvky, jejich formát, slovník a validační pravidlo	Zdokumentovaný profil s verzí
Oddělení vrstev	Zda vedle obecného exportu existuje autoritativní oborový záznam	Záznam v repozitáři a jeho odvozený export
Vazba slovníku	Který slovník upravuje které pole a zda se jeho termíny rozklíčují	Definice pole s URI slovníku
Časování popisu	Zda jsou prvky popisu zachyceny v průběhu, nebo rekonstruovány	Časová razítka záznamů popisu vůči záznamům o činnosti
Opatření zajišťování kvality	Preventivní opatření fixovaná před sběrem	Validační pravidla, kalibrační záznamy, dokumenty postupů
Opatření kontroly kvality	Detekční opatření uplatněná po sběru	Validační zprávy; záznam oprav
Deklarovaný účel	Užití, vůči němuž byla vhodnost posouzena	Výrok o kvalitě doprovázející uložení
Výrok o zbytkové chybě	Ponechaná známá chyba, s velikostí, směrem nebo mechanismem	Výrok o kvalitě; dokumentace harmonizace

Dimenzi kvalifikují dvě hranice platnosti. V oborech, jejichž primárním materiálem je text nebo interpretace, nese oborovou vrstvu vědecký aparát a popis pramenů, a nikoli dokumentace na úrovni proměnných, a výše uvedené proměnné jsou naplněny v jiné formě, nikoli prominuty. A v oblastech, kde standardy popisu zůstávají neustálené, zaznamenává proměnná profilu lokální profil s uvedeným ujednáním o údržbě, což je slabší než přijetí komunitou, ale silnější než nezdokumentovaná konvence.

Popis a kvalita určují, zda datovou sadu dokáže interpretovat někdo, kdo ji nevytvořil. Zda smí být zpřístupněna, komu a za jakých podmínek, určuje dimenze ochrany, kterou rozvíjí následující kapitola.

Kapitola 9. Ochrana, právní režim a řízený přístup

Popis rozhoduje o tom, zda lze datovou sadu interpretovat; nerozhoduje o tom, zda smí opustit prostředí, v němž vznikla. Tato kapitola se ujímá ochrany jako dimenze upravující ukládání, přístup, přenos a zpřístupnění a obhájí dvě spojené teze. První zní, že klasifikace dat je podmiňující proměnnou této dimenze, a to v tom smyslu, že úložné prostředí, model přístupu, kontrola výstupu z prostředí i režim publikování jsou jejími funkcemi, nikoli nezávislými volbami. Druhá zní, že zásada „tak otevřeně, jak je to možné, tak uzavřeně, jak je to nezbytné“ funguje jako rozhodovací pravidlo pouze tehdy, je-li podmínka nezbytnosti rozložena na posouditelné proměnné, neboť v nerozložené podobě ospravedlní jakýkoli výsledek.

S hlavní výzkumnou otázkou to souvisí tím, že ochrana je dimenzí, v níž jsou konfliktní požadavky nejhustší a v níž jsou náklady nepřiděleného rozhodnutí nejvyšší. Retenční povinnost a povinnost výmazu mohou dopadat na tytéž záznamy, otevřená licence může být právně dostupná a přitom eticky neslučitelná s komunitním protokolem a bezpečnostní klasifikace může zakazovat to, co vyžaduje podmínka poskytovatele podpory. Takové konflikty nelze vyřešit uvnitř žádného z nástrojů, které je generují, a rámec proto zachází s ochranou jako se souborem proměnných s pojmenovanou rozhodovací autoritou a zaznamenaným zbytkovým rizikem, a nikoli jako s kontrolním seznamem souladu.

9.1 Klasifikace jako podmiňující proměnná

Klasifikace se často redukuje na binární rozlišení mezi osobními a neosobními údaji a právě tato redukce je hlavním zdrojem chybného zařazení. Důvěrnost vzniká stejnou měrou ze smluvního omezení, z obchodního tajemství, z ochrany kritické infrastruktury a z kulturní citlivosti, přičemž nic z toho není v testu na osobní údaje viditelné. Klasifikace omezená na důvěrnost je neúplná i ve druhém ohledu, neboť požadavky na integritu a dostupnost se u objektů shodné citlivosti liší (National Institute of Standards and Technology 2024). Rámec proto zachází s klasifikací jako s vektorem nad dopadem na osoby, na instituci, na veřejnou bezpečnost a na důkazní hodnotu výzkumu.

O tom, jak klasifikace vstupuje do rozhodovacího pravidla, rozhodují dvě její vlastnosti. Klasifikace se připojuje k datovým objektům a v některých případech k proměnným, nikoli k projektům, protože jediný projekt běžně kombinuje veřejné senzorické záznamy, neveřejné výkresy infrastruktury a identifikovatelné kontaktní údaje. A klasifikace je závislá na čase: materiál důvěrný před podáním patentové přihlášky může být po něm publikovatelný, agregace může riziko identifikace snížit a propojení je může zvýšit. Klasifikace zaznamenaná jednou na počátku projektu tedy popisuje stav, který již nenastává, a právě tento mechanismus vede k tomu, že nadměrná ochrana běžných dat a nedostatečná ochrana citlivých fragmentů vznikají současně. Tabulka 9-1 uvádí, co přiřazená třída určuje.

Tabulka 9-1. Důsledky určené přiřazenou třídou dat, se zbytkovým rizikem, které každá kombinace ponechává.

Třída	Schválené prostředí	Model přístupu	Kontrola výstupu z prostředí	Režim zpřístupnění
Veřejná	Běžné institucionální úložiště	Otevřený uvnitř týmu	Žádná	Otevřené uložení
Interní	Institucionální síťové úložiště	Podle rolí, časově omezený	Přenos zaznamenáván v logu	Uložení s podmínkami uvedení autorství
Důvěrná	Institucionální úložiště s auditem, nebo schválený cloud	Podle rolí, s vyžadovaným schválením	Export schvalovaný pojmenovanou rolí	Řízený přístup
Vysoce citlivá	Důvěryhodné výzkumné prostředí	Jmenovitý, posouzený, vázaný na projekt	Kontrola výstupů před zpřístupněním	Pouze metadata, nebo řízený přístup na základě smlouvy

Zbytková rizika se liší druhem, a nikoli pouze mírou. Běžné úložiště ponechává slabý audit a slabou kontrolu exportu, schválená cloudová platforma ponechává smluvní a jurisdikční expozici a důvěryhodné výzkumné prostředí ponechává náklady a omezení workflow, jež nese výzkum, a nikoli instituce. Protože každá varianta v tabulce 9-1 ponechává zbytkové riziko, závisí obhajitelnost volby na tom, zda je uvedeno, které riziko bylo přijato a kým.

9.2 Právní režim a meze souhlasu

Osobní údaje nejsou definovány přítomností přímého identifikátoru a kombinace místa, věku, povolání a události může osobu identifikovat i bez jakéhokoli jména (European Parliament and Council 2016). Pseudonymizované údaje zůstávají osobními údaji, protože dodatečná informace umožňující přiřazení nadále existuje. Výzkumná ustanovení nařízení jsou podmíněná, nikoli povolující: předpokládají právní titul, omezení účelu, minimalizaci, transparentnost, zabezpečení a záruky vyžadované podle článku 89. Pojímát výzkumný kontext jako výjimku samu o sobě je nejzávažnějším právním omylem, který korpus zaznamenává.

Souhlas zaujímá zvláštní postavení, které rámec rozděluje do tří vrstev. Etický souhlas s účastí ve studii, právní titul pro zpracování podle článku 6 a podmínka umožňující zpracování zvláštních kategorií podle článku 9 jsou odlišné nástroje odpovídající na odlišné otázky a podepsaný formulář o účasti vyřizuje pouze první z nich. Souhlas jako právní titul je odvolatelný, což staví dlouhodobé uchování a sekundární opětovné použití do strukturně nestabilní pozice, a tituly veřejného či oprávněného zájmu ve spojení se zárukami podle článku 89 bývají pro výzkum prováděný veřejnými institucemi robustnější. Rámec proto zaznamenává právní titul jako proměnnou s pojmenovaným vlastníkem, zpravidla pověřencem pro ochranu osobních údajů, a nikoli jako pole vyplňované výzkumným týmem.

Důkazní status současných vodítek v této oblasti vyžaduje výslovné uvedení. Evropská vodítka ke zpracování pro účely vědeckého výzkumu a k anonymizaci byla v průběhu roku 2026 v konzultaci a k datu ověření korpusu nedosáhla konečné podoby (European Data Protection Board 2026). Vodítka ve stavu návrhu naznačují výkladový směr, aniž by zakládala pravidlo, a rámec nese toto rozlišení jako vlastnost zdroje, nikoli jako výhradu v textu.

9.3 Anonymizace jako kontextové posouzení rizika

Anonymizace je stavem, v němž osoba není identifikovatelná prostředky, jejichž použití lze rozumně předpokládat, a je proto relativní vůči prostředí, nikoli vlastní danému souboru. Odstranění přímých identifikátorů tento stav nezakládá a posouzení postupuje třemi testy: zda lze záznam vyčlenit, zda jej lze propojit s jiným záznamem nebo datovou sadou a zda lze s dostatečnou jistotou odvodit určitý atribut. Bohatá kvalitativní, genetická, prostorová a

longitudinální data po odstranění přímých identifikátorů běžně neobstojí nejméně v jednom testu a absolutně nulové riziko u nich zpravidla doložit nelze. Posouzení tak závisí na tom, jaké externí zdroje existují, což se v čase mění a co není vlastností, kterou by projekt produkující data kontroloval.

Z toho plyne asymetrie, s níž rámec zachází jako s rozhodnutím, a nikoli jako s technickým detailem. Anonymizace natolik silná, aby vyhověla všem třem testům, může zničit analytickou hodnotu, kvůli níž byla data pořízena, a obhajitelnou alternativou je pseudonymizace ve spojení s přísnějším prostředím přístupu, nikoli nezdokumentované oslabení ochrany. Syntetické deriváty podpírají vývoj, výuku a testování kódu a nejsou ekvivalentem anonymizace původních záznamů. Zaznamenávanou proměnnou je samo rozhodnutí o anonymizaci, spolu s uplatněnými testy, zvažovanými externími zdroji a přijatou ztrátou informace.

CHYBĚJÍCÍ DOKLAD

Testování reidentifikace na vzorku uložených výzkumných datových sad prohlášených za anonymní, s uvedením podílu sad neobstojících v testech vyčlenění, propojitelnosti a odvození, aby bylo možné stanovit, jak často se odstranění přímých identifikátorů pokládá za dostačující

9.4 Řízený přístup a Five Safes

Dichotomie otevřeného a uzavřeného ponechává nezařazenou širokou třídu dat, která nesou vysokou analytickou hodnotu a netriviální riziko identifikace. Řízený přístup je režimem pro tuto třídu a je službou, nikoli nastavením: veřejná metadata, uvedené podmínky, posouzení žadatele a navrhovaného účelu, smlouva o užití dat, zabezpečené prostředí, audit a kontrola výstupů. Rámec Five Safes poskytuje strukturu posouzení tím, že odděluje bezpečné osoby, projekty, data, prostředí a výstupy, a jeho analytickým přínosem je teze, že ochrana vzniká z kombinace, a nikoli z kterékoli jednotlivé vrstvy (UK Data Service 2025). Datová sada příliš riziková na to, aby byla zpřístupněna při jedné konfiguraci těchto pěti složek, může být zpřístupnitelná při jiné, čímž se zdánlivě binární soud mění v úlohu návrhu.

Tabulka 9-2 uvádí režimy přístupu, jež rámec rozlišuje, spolu s podmínkou, za níž je každý z nich proporcionální. Nejsou uspořádány pouze podle přísnosti, neboť pouze metadatový záznam zachovává vyhledatelnost

tam, kde přístup není možný, a důvěryhodné výzkumné prostředí zachovává analýzu tam, kde není možný přenos.

Tabulka 9-2. Režimy přístupu k datům nesoucím riziko identifikace nebo důvěrnosti, s podmínkami proporcionality a zbytkovým rizikem.

Režim	Podmínka proporcionality	Zbytkové riziko	Stupeň důkazu
Otevřené anonymní zpřístupnění	Realistické riziko identifikace je nízké a otestované	Budoucí externí zdroje mohou umožnit propojení	A/B
Stažení po schválení	Střední riziko; identifikovatelný žadatel a smlouva	Omezená kontrola následného užití	B
Důvěryhodné výzkumné prostředí	Vysoké riziko nebo citlivé propojování	Náklady; omezení analytického workflow	A/B
Pouze metadata	Přístup není právně ani eticky možný	Vyhledatelné, avšak přímo nepoužitelné	B

Výrok o dostupnosti dat je bodem, v němž se zvolený režim stává viditelným, a jeho informační obsah odlišuje režim od odmítnutí. Výrok, který pojmenovává repozitář, podmínky přístupu, rozhodující orgán a důvod omezení, popisuje režim, jež lze posoudit; nerozlišující formule nabízející data na odůvodněnou žádost nepopisuje žádný.

9.5 Kontinuita jako vlastnost ochrany

Diskusi o ochraně dominuje důvěrnost, zatímco selhání, jež ukončují výzkumné projekty, jsou častěji selháními dostupnosti a integrity. Záloha je oddělenou kopií určenou k obnově, synchronizace šíří výmaz a zašifrování stejně rychle, jako šíří obsah, a archiv sleduje dlouhodobou interpretovatelnost, a nikoli provozní obnovu. Kontinuita sahá za soubory ke znalostem a k oprávněním, neboť ztráta šifrovacího klíče, odchod jediného správce nebo ukončení cloudové smlouvy produkují týž výsledek jako fyzické zničení. Rámec proto zachází s kontinuitou jako se součástí ochrany, a nikoli jako se sousední infrastrukturní službou.

Rozhodujícím bodem je měřená veličina. Existence zálohovací úlohy nezakládá nic o tom, zda je kopie úplná, čitelná a použitelná se softwarem a klíči potřebnými k jejímu výkladu, a odpovídajícími proměnnými jsou

přípustná ztráta dat, doba obnovení a datum i výsledek posledního testu obnovy. Oba cíle jsou vlastnostmi třídy dat, a nikoli instituce, protože nenahraditelný pozorovací vstup a regenerovatelný mezivýstup opodstatňují odlišné hodnoty, a jednotný denní harmonogram nevyjadřuje žádné posouzení. Nejméně jedna administrativně oddělená nebo neměnná kopie pokrývá způsob selhání, jež vytváří kompromitovaný účet nebo chyba správce a jež kopie uvnitř jediné administrativní domény nepokrývají.

CHYBĚJÍCÍ DOKLAD

Audit výzkumných projektů zaznamenávající, zda jsou cíle obnovy stanoveny pro jednotlivé třídy dat a zda byl ve stanoveném období proveden a zdokumentován úplný test obnovy, aby bylo možné stanovit rozdíl mezi poskytovaným zálohováním a doloženou obnovitelností

9.6 Práva jako vrstvy, nikoli jako vlastnictví

Výraz vlastnictví dat nepopisuje žádný jednotný právní vztah a zastírá vrstvení, jež zpřístupnění ve skutečnosti upravuje. Jediný výzkumný balíček může spojovat nechráněné skutečnosti, autorský text, obrazový materiál třetích stran, strukturu databáze zakládající zvláštní právo pořizovatele databáze (*sui generis*) tam, kde lze doložit podstatný vklad, osobní údaje a software, přičemž ke každé z těchto složek náleží jiný nositel práv (European Parliament and Council 1996). Pracovněprávní vztahy určují, kdo smí udělovat oprávnění k materiálu vzniklému při plnění pracovních povinností jako zaměstnanecké dílo, a smluvní výzkum přidává ustanovení o vstupních znalostech, výsledcích, přístupových právech, publikování a o stavu věcí po ukončení smlouvy. Pro účely správy dat jsou operativní otázky užší než vlastnictví: kdo smí autorizovat archivaci, kdo smí autorizovat publikování a za jakých podmínek je přípustné další užití.

Volba licence plyne z tohoto soupisu, a nikoli z výchozího nastavení. Licence převádí pouze ta oprávnění, jimiž poskytovatel skutečně disponuje, a nemůže založit práva k nechráněným skutečnostem, odstranit povinnosti podle práva ochrany osobních údajů ani překonat práva třetích stran a autoritu komunity (Carroll et al. 2020). Objekty odlišného druhu vyžadují licence odlišných rodin, takže jediný nerozlišující licenční soubor pokrývající data, dokumentaci i kód vytváří nejednoznačnost právě tam, kde strojové zpracování vyžaduje jistotu; licenční manifest zaznamenávající pro každou

logickou část nositele práv, zdroj oprávnění a vyloučený materiál ji odstraňuje. Rámec zaznamenává licenci jako rozhodnutí s uvedeným důsledkem, neboť permissivní dedikace maximalizující právní opětovné použití je nedostupná tam, kde ukladatel nedisponuje plnými právy, a požadavek na uvádění autorství se napříč odvozenými kompilacemi kumuluje.

9.7 Jak učinit podmínku nezbytnosti posouditelnou

Zásada „tak otevřeně, jak je to možné, tak uzavřeně, jak je to nezbytné“ je široce přijímána a sama o sobě rozhodovacím pravidlem není, protože neobsahuje žádný test nezbytnosti, na niž se odvolává. Jakoukoli míru omezení lze ospravedlnit tvrzením o nezbytnosti a jakoukoli míru otevřenosti jejím popřením, takže zásada ratifikuje kteroukoli praxi, již instituce již sleduje. Příspěvek, který tato studie v tomto bodě přináší, spočívá v rozložení nezbytnosti na šest proměnných, z nichž každá je samostatně posouditelná a každá má vlastníka: právní titul, zbytkové riziko identifikace, práva třetích stran, smluvní omezení, komunitní protokol a bezpečnostní klasifikace. Otevřenost je pak výsledkem, jenž nastává tehdy, stojí-li každá proměnná ve svém nulovém stavu, a každá odchylka od otevřenosti je přiřítelná pojmenované proměnné, a nikoli obecnému odvolání na citlivost.

Rozklad má dva důsledky, jichž nerozložená zásada dosáhnout nemůže. Omezení přiřtené proměnné je přezkoumatelné, takže zveřejnění patentu nebo dokončení posouzení anonymizace spouští reklasifikaci namísto toho, aby omezení setrvalo setrvačností. A omezení nepřičítelné žádné ze šesti proměnných se odkrývá jako nevyslovená preference rizika, a nikoli jako požadavek. Tabulka 9-3 uvádí proměnné ochrany přenášené dále do integrovaného rámce.

Tabulka 9-3. Proměnné ochrany v integrovaném rámci a jejich doložitelné záznamy.

Proměnná	Definice	Doložitelný záznam
Třída dat	Třída podle dopadu přiřazená jednotlivému datovému objektu	Registr klasifikace s datem přiřazení
Právní titul	Titul pro zpracování, odlišený od etického souhlasu	Záznam o činnostech zpracování; posouzení pověření pro ochranu osobních údajů
Riziko identifikace	Výsledek testů vyčlenění, propojitelnosti a odvození	Posouzení anonymizace s uvedenou ztrátou informace

Proměnná	Definice	Doložitelný záznam
Režim přístupu	Otevřený, stažení po schválení, řízené prostředí, nebo pouze metadata	Nastavení přístupu v repozitáři; výrok o dostupnosti dat
Kontrola výstupu z prostředí	Zda jsou výstupy kontrolovány před opuštěním prostředí	Log kontroly výstupů
Cíle obnovy	Přijatelná ztráta dat a doba obnovení podle třídy	Plán obnovy; datovaná zpráva o testu obnovy
Inventář práv	Nositel práv a zdroj oprávnění pro každou vrstvu objektu	Manifest licencí a uvádění autorství
Přičtení omezení	Proměnná nezbytnosti, z níž plyne každá odchylka od otevřenosti	Záznam o rozhodnutí s uvedením proměnné, vlastníka a data přezkumu

Dimenzi omezují tři hranice platnosti. Právní závěry jsou závislé na jurisdikci a proměnné identifikují třídu rozhodnutí a jeho autoritu, nikoli odpověď; rámec proto lokalizuje eskalační bod a právní posouzení nenahrazuje. V oborech pracujících se znalostmi drženými komunitou se právní dostupnost a etická přípustnost rozcházejí a proměnná komunitního protokolu není redukovatelná na inventář práv (Carroll et al. 2020). A v malých výzkumných prostředích může být důvěryhodné výzkumné prostředí nedostupné za jakýchkoli přijatelných nákladů, a v takovém případě rámec zaznamenává pouze metadatový režim s uvedeným důvodem namísto nepodloženého tvrzení o anonymitě.

Ochrana určuje, co smí být zpřístupněno a komu. Zda zpřístupněné podpírá tvrzení, jež jsou z něj vyvozována, určuje dimenze reprodukovatelnosti, kterou rozvíjí následující kapitola.

Kapitola 10. Reprodukovatelnost a výpočetní integrita

Dimenze ochrany určuje, co smí být zpřístupněno; neurčuje, zda zpřístupněný materiál podpírá tvrzení, jež jsou z něj vyvozována. Tato kapitola se ujímá reprodukovatelnosti jako dimenze spojující uvedený výsledek s daty, kódem a podmínkami, jež jej vytvořily, a obhazuje jednu hlavní tezi: reprodukovatelnost je vlastností důkazu, a nikoli jeho etickým doplňkem. Výsledek, jež nelze regenerovat z materiálu, který jej doprovází, není výsledkem, u něhož zůstala správnost neověřena, nýbrž výsledkem, jehož obsah je zčásti nedourčen, neboť nezdokumentované volby učiněné při jeho vytváření jsou konstitutivní pro vykázané číslo. Kapitola proto uvádí, co rámeček měří, namísto toho, co se praktikům doporučuje.

S hlavní výzkumnou otázkou to souvisí tím, že reprodukovatelnost je dimenzí, v níž předmět správy sahá nejdál za datovou sadu. Data sama analýzu neobnovují a kód sám o sobě neběží; analytický objekt zahrnuje verze vstupů, transformační kód, konfiguraci, verze závislostí, náhodná semena a prováděcí prostředí, rozprostřené napříč systémy spravovanými odlišnými službami a odlišnými vlastníky. Rámeček proto vyjadřuje reprodukovatelnost proměnnými, jež se připojují k výpočetnímu objektu jako celku, což je jediná úroveň, na níž může být tato vlastnost naplněna i posouzena.

10.1 Dvě často zaměňované vlastnosti

Reprodukovatelnost a replikovatelnost jsou samostatné vlastnosti se samostatnou důkazní silou a jejich zaměňování oslabuje obě. Reprodukovatelnost nastává tehdy, když táž data, týž kód a tytéž podmínky poskytnou týž výsledek, a testuje výpočetní řetězec mezi pozorováním a tvrzením. Replikovatelnost nastává tehdy, když nová studie sbírající nová data za srovnatelných podmínek dospěje ke konzistentnímu závěru, a testuje empirickou generalizaci (National Academies of Sciences, Engineering, and Medicine 2019). Selhání prvního lokalizuje vadu v pipeline nebo v její dokumentaci; selhání druhého může naznačovat skutečnou mez zkoumaného jevu, a pojmímat kterékoli z nich jako verdikt o tom druhém znamená chybně přiřadit diagnózu.

Rozlišení má důsledek pro to, co může správa dat dodat. Reprodukovatelnost je v moci týmu, který data produkuje, protože závisí na artefaktech, jimiž tým disponuje, a je proto posouditelná při uložení; replikovatelnost závisí na nezávislém úsilí a pozdějších důkazech a žádné uložení ji nezakládá. Rámec si proto nárokuje jurisdikci nad prvním a s druhým zachází jako s výsledkem, jež první umožňuje, což zároveň vysvětluje, proč reprodukovatelnost není zástupným ukazatelem správnosti. Kód, který věrně reprodukuje chybnou transformaci, reprodukuje s naprostou věrností i chybu, takže tato vlastnost dokládá dohledatelnost, a nikoli validitu.

10.2 Výpočetní prostředí jako součást výzkumného objektu

Analýza se provádí nad zásobníkem verzí knihoven, nad operačním systémem, nad hardwarovou architekturou a ve stále větší míře nad externími službami, přičemž každá vrstva může výstup pozměnit. Dokumentace tohoto zásobníku sahá od seznamu balíčků přes lock file až po obraz kontejneru nebo virtuální stroj a jednotlivé varianty se liší tím, co zachycují, nikoli pouze tím, kolik úsilí stojí. Kontejner podstatně zvyšuje přenositelnost a nezaručuje dlouhodobou spustitelnost, neboť závisí na registru, na základním obrazu, na architektuře a na kterékoli externí službě, již analýza volá. Tabulka 10-1 uvádí varianty spolu s podmínkou, za níž je každá proporcionální.

Tabulka 10-1. Varianty zaznamenání výpočetního prostředí, s podmínkami proporcionality a zbytkovým rizikem.

Varianta	Podmínka proporcionality	Zbytkové riziko	Stupeň důkazu
Zdokumentovaný seznam balíčků	Malá analýza v běžném jazyce	Verze se rozcházejí; výsledky nelze později regenerovat	C
Lock file s rozřešenými verzemi	Standardní ekosystém balíčků	Operační systém a externí služby nezachyceny	B
Obraz kontejneru s definicí sestavení	Složitě závislosti nebo systémové knihovny	Závislost na běhovém prostředí, registru a architektuře	B

Varianta	Podmínka proporcionality	Zbytkové riziko	Stupeň důkazu
Virtuální stroj nebo emulace	Historicky kritické nebo neudržovatelné workflow	Vysoké náklady; zátěž uchování přenesená do budoucna	B/C

Pro rámec z toho plynou dva důsledky. Bitově identický výstup není u řady numerických úloh dosažitelný napříč hardwarem, pořadími paralelního provádění a verzemi knihoven, což z přijatelné numerické tolerance činí proměnnou analýzy, a nikoli přiznání selhání. A obraz uchovaný bez své definice sestavení, bez zdroje a bez metadat zachovává provedení, přičemž ničí možnost prostředí rekonstruovat nebo upravit, takže rámec zaznamenává specifikaci prostředí a obraz jako odlišné artefakty. Uchování obojího je uspořádáním, jež korpus konzistentně uvádí jako preferované, na základě téže úvahy, jaká upravuje autoritativní kopie v oddílu 6.2.

10.3 Stupně volnosti na cestě od dat k výsledku

Mezi surovým pozorováním a vykázaným odhadem leží sled obhajitelných voleb: které případy vyloučit, které proměnné odvodit, kterou transformaci uplatnit, které kovariáty zahrnout, který model odhadnout a kdy sběr ukončit. Každá volba je jednotlivě rozumná a jejich soubor definuje mnoho cest z týchž dat k odlišným číslům. Je-li cesta vybrána poté, co jsou známy výsledky několika z nich, a vykázána tak, jako by byla fixována předem, je důkazní síla vykázaného čísla nadhodnocena. Účinek nastává i bez jakéhokoli úmyslu klamat (Gelman and Loken 2014; Simmons, Nelson and Simonsohn 2011). Opakované testování v průběhu narůstajícího sběru, ukončené v bodě dosažení významnosti, mění chybovost testu týmž způsobem a je přípustné pouze v rámci sekvenčního designu, který s tím počítá.

Rámec zachází s analytickou flexibilitou jako se zaznamenanatelnou veličinou, a nikoli jako s otázkou ctnosti výzkumníka. Zaznamenanatelným prvkem je spárování každého rozhodovacího bodu s okamžikem, v němž byl učiněn, a s informací dostupnou v tomto okamžiku, což odlišuje volbu fixovanou před spatřením dat od volby na nich podmíněné. Multiverzní analýza a analýza křivky specifikací záznam rozšiřují tím, že uvádějí rozdělení výsledků napříč obhajitelnými specifikacemi, a jejich příspěvek je diagnostický: ukazují, zda závěr závisí na jediné analytické volbě (Steege et

al. 2016). Nenahrazují věcně odůvodněný primární model, neboť křivka výsledků bez odůvodněné primární specifikace vykazuje variabilitu, aniž by ji interpretovala.

10.4 Předem stanovený analytický plán jako oddělovač druhů tvrzení

Předem stanovený analytický plán odděluje plán analýzy od znalosti výsledků a jeho funkcí není potlačení průzkumu, nýbrž jeho označení. Plán zaznamenaný před sběrem, s časovým razítkem a v nezměnitelném záznamu, umožňuje pozdějšímu čtenáři odlišit konfirmatorní tvrzení od exploratorního; druhé z nich nezakazuje a exploratorní práce zůstává mechanismem, jímž vzniká většina hypotéz (Nosek et al. 2018). Odchylka od plánu bývá odůvodněna podmínkami zjištěnými v průběhu sběru a transparentně vykázaná odchylka je silnějším důkazem než dodržení plánu, který se stal nepoužitelným. Plán umožňuje přiřadit každé vykázané tvrzení k jedné ze dvou kategorií s odlišnou důkazní vahou.

Dosah tohoto nástroje omezují tři hranice platnosti. Předchozí stanovení plánu kvalitu nezakládá a slabě odůvodněný předem registrovaný model může být méně přiléhavý než transparentně opravená analýza. Ne každý obor testuje hypotézy a kvalitativní, exploratorní, modelovací a simulační výzkum vyžaduje jinou podobu analytického plánu, popisující zamýšlený postup, kritéria interpretace a podmínky revize. A předchozí stanovení plánu vstupuje do interakce s dimenzí plánování z kapitoly 7, neboť statistický plán nepropojený s plánem správy dat produkuje dva dokumenty popisující jeden projekt bez definovaného vztahu mezi nimi.

CHYBĚJÍCÍ DOKLAD

Srovnání předem registrovaných analytických plánů s odpovídajícími publikovanými analýzami na vzorku studií, se zaznamenáním podílu odchylek vykázaných jako odchylky, aby bylo možné stanovit, zda předem stanovený plán v praxi funguje jako oddělovač druhů tvrzení

10.5 Nejistota jako vykazovaná veličina

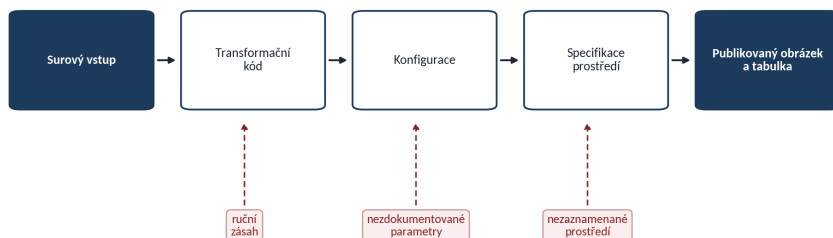
Nejistota výzkumného výsledku pochází z měření, z výběru, ze specifikace modelu, z imputace, z kalibrace a z budoucí variability, přičemž uvedení standardní chyby pokrývá jeden z těchto zdrojů. Jediný bodový odhad

doprovázený prahovým rozhodnutím zbývající zdroje spíše potlačuje, než řeší, a rámec zachází s uznanými zdroji i se zdroji ponechanými stranou jako se složkami výsledku. Analýza citlivosti vůči obhajitelným alternativním předpokladům je provozní podobou tohoto uznání a její výstup náleží k primárnímu odhadu, a nikoli do doplňkového materiálu. Plánování orientované na přesnost, v němž se velikost vzorku odvozuje ze šířky intervalu, jejíž závěr vyžaduje, tutéž logiku rozšiřuje zpětně do fáze návrhu.

Grafické sdělování je místem, kde k potlačení nejistoty dochází nejčastěji a kde je pro čtenáře nejméně patrné. Souhrnné statistiky mohou zakrýt multimodalitu, shlukování a odlehlou strukturu, které graf původních hodnot odkrývá, a právě proto vizualizace funguje jako součást analýzy, a nikoli jako její výzdoba (Anscombe 1973). Chybové úsečky nenesou žádný pevný význam, neboť mohou označovat směrodatnou odchylku, standardní chybu, konfidenční interval, predikční interval nebo věrohodnostní interval, a nedefinovaná úsečka je neinterpretovatelná, a nikoli přibližně interpretovatelná. Důkazy naznačují, že napříč úlohami není žádná jediná grafická podoba pro sdělování nejistoty optimální, a rámec proto zaznamenává definici intervalu, velikost vzorku a parametr vyhlazení nebo binování namísto toho, aby předepisoval typ grafu.

Vizualizaci s již ustavenými dimenzemi spojují dvě další vlastnosti. Obrázek generovaný skriptem z verzované datové sady nese provenienci, kdežto obrázek ručně upravený v grafickém editoru přerušuje řetězec mezi zobrazenou hodnotou a rozhodnými daty. Rámec proto zachází s ruční úpravou jako s událostí, jež má být zdokumentována, a nikoli jako s neutrálním úkonem. A přístupnost je vlastností důkazu, nikoli zdvořilostí, neboť význam zakódovaný výhradně barvou je pro část čtenářstva nedostupný a přímé popisky, kontrast, tvar a alternativní text jej nesou tam, kde barva nemůže (World Wide Web Consortium 2024).

Provenience je přerušena všude, kde krok nelze odvodit ze samotného záznamu



Obr. 10-1. Řetězec od surového vstupu přes transformační kód, konfiguraci a specifikaci prostředí až po vykázaný obrázek a tabulku, s vyznačením bodů, v nichž se provenience přerušuje ručním zásahem, nezdokumentovanými parametry nebo nezaznamenaným prostředím. Podklad: kódování reprodukovatelnosti popsané v oddílu 4.4. Vlastní zpracování autora.

10.6 Výzkumný software jako výsledek s vlastní identitou

Výzkumný software je současně součástí metody a výsledkem sám o sobě a pojímat jej pouze jako přílohu datové sady znamená chybně popsat obojí. Od dat se liší tím, že je spustitelný, složený a průběžně se měnící, takže pravidla upravující statické objekty se na něj přenášejí jen nedokonale. Vyhledatelnost vyžaduje identifikaci konkrétních vydání, a nikoli projektu, a opětovná použitelnost vyžaduje licenci, dokumentaci, informace o závislostech a výrok o omezeních (Barker et al. 2022). Commit v systému správy verzí není pro čtenáře použitelným cílem citace, kdežto označené vydání, archivované a vázané na persistentní identifikátor, jím je. Mechanismus poskytuje praxe citování softwaru a strojově čitelný citační soubor činí zamýšlenou citaci explicitní namísto odvozované (Smith, Katz and Niemeyer 2016).

Identifikace kódu získala vlastní infrastrukturu a její meze jsou stejně důležité jako její schopnosti. Obsahově adresovaný identifikátor archivovaného zdrojového kódu identifikuje přesně ten stav kódu, který byl spuštěn. Metadatová schémata takové identifikátory nyní připouštějí jako typ souvisejícího identifikátoru, což umožňuje propojit datovou sadu, článek a stav kódu do jediného propojeného objektu (DataCite Metadata Working Group 2026; Soiland-Reyes et al. 2022). Archivovaný zdrojový kód přesto není automaticky sestavitelný ani spustitelný, neboť k rekonstrukci chování jsou

zapotřebí závislosti, konfigurace, testovací data a dokumentace. Údržba po skončení projektu vyžaduje pojmenovaného vlastníka a financování a otevřená licence sama o sobě neprodukuje ani komunitu, ani opravy, a rámec proto zaznamenává horizont údržby a postup pro označení softwaru za neudržovaný.

Samotné schéma identifikátoru bylo mezitím standardizováno: SoftWare Hash IDentifier, vyvinutý pro archiv Software Heritage, byl v roce 2025 publikován jako mezinárodní norma (ISO/IEC 2025), čímž se komunitní konvence mění v citovatelný normativní odkaz.

10.7 Co rámec měří

Proměnné této dimenze jsou zvoleny tak, že každá pojmenovává záznam, jež lze prozkoumat, aniž by se produkující tým dotazoval na své záměry. Tabulka 10-2 je uvádí. Dvě z nich nesou většinu rozlišovací síly. První je test čistého spuštění, při němž se workflow provede z prázdného prostředí od začátku do konce a jeho výsledek se zaznamená. Druhou je oddělení neměnného vstupu, transformačního kódu, konfigurace a generovaného výstupu, jež činí směr odvození viditelným už ve struktuře samotného uložení.

Tabulka 10-2. Proměnné reprodukovatelnosti v integrovaném rámci a jejich doložitelné záznamy.

Proměnná	Definice	Doložitelný záznam
Oddělení objektů	Zda vstup, kód, konfigurace a výstup zaujímají odlišné vrstvy	Struktura uložení; manifest
Specifikace prostředí	Podoba, v níž je zaznamenán prováděcí zásobník	Lock file, definice sestavení nebo obraz s metadaty
Vazba verzí	Která verze dat, které vydání kódu a která množina závislostí výsledek vytvořily	Identifikátory vydání citované ve výstupu
Numerická tolerance	Odchylka připouštěná jako úspěšná reprodukce	Výrok doprovázející reprodukční balíček
Výsledek čistého spuštění	Výsledek provedení od začátku do konce z prázdného prostředí	Datovaný log spuštění
Časování rozhodnutí	Kdy byla každá analytická volba fixována vzhledem k datům	Analytický plán s časovým razítkem; záznam o odchylce

Proměnná	Definice	Doložitelný záznam
Výrok o nejistotě	Uznané zdroje a zdroje ponechané stranou	Analýza citlivosti; definice intervalů
Provenience obrázku	Zda jsou výstupy generovány skriptem z rozhodných dat	Generující skript; zdokumentovaný ruční zásah
Identita softwaru	Zda jsou vydání identifikována, licencována, zdokumentována a citovatelná	Identifikátor vydání; citační soubor; horizont údržby

Chráněná data dimenzi omezují, aniž by ji suspendovala. Tam, kde vstupní data nelze zpřístupnit, zachovávají úplný kód, zdokumentované workflow a minimální nebo syntetická ukázková datová sada prozkoumatelnost postupu a ověření proti skutečným vstupům lze provést uvnitř řízeného prostředí popsaneho v oddílu 9.4. Transparentnost tohoto druhu nevyžaduje porušení právního omezení, a její absenci proto nevysvětluje citlivost sama o sobě. Rámec v takových případech zaznamenává, která složka byla zadržena, v jakém režimu přístupu ji lze prozkoumat a kdo takové prozkoumání provedl.

Celou dimenzi omezují dvě hranice platnosti. V oborech, jejichž analýza je interpretativní, a nikoli výpočetní, vede řetězec od materiálu k tvrzení přes kódovací schémata, analytická memoranda a pramenný aparát, a proměnné jsou naplněny v této podobě, nikoli prominuty. A úplné uchování prostředí nese finanční a environmentální náklady škálující s počtem uchovávaných objektů, takže rámec zachází s hloubkou zachycení prostředí jako s veličinou úměrnou nenahraditelnosti výsledku, a nikoli jako s jednotným požadavkem.

Reprodukovatelnost zakládá, že výsledek lze regenerovat, dokud artefakty zůstávají neporušené a interpretovatelné. Zda takové zůstávají po dobu, po níž se od výzkumného důkazu očekává platnost, a zda je úsilí vynaložené na jejich udržení uznáváno, určuje dimenze dlouhodobého uchování a uznání, kterou rozvíjí následující kapitola.

Kapitola 11. Dlouhodobé uchování, publikování a uznání

Dosud ustavené dimenze určují, zda lze výsledek interpretovat, zpřístupnit a regenerovat v okamžiku uložení. Žádná z nich neurčuje, zda materiál zůstane interpretovatelný poté, co projekt, který jej vytvořil, skončí, ani zda je úsilí vynaložené na jeho udržení viditelné pro kohokoli, kdo instituci posuzuje. Tato kapitola se ujímá dlouhodobého uchování, publikování a uznání jako jediné dimenze a obhájí dvě teze. První zní, že dlouhodobé uchování je aktivním závazkem odlišným jak od zálohování, tak od publikování. Druhá zní, že uznání je strukturní proměnnou rámce, a nikoli otázkou profesní morálky, protože úsilí, které nelze posoudit, se neudrží.

S hlavní výzkumnou otázkou to souvisí tím, že tato dimenze nese nejdelší časový horizont rámce a jeho nejslabší vlastnictví. Rozhodnutí o retenční lhůtě, formátu a repozitáři nabývají účinku roky poté, co se projektový tým rozešel, a náklady dopadají na instituci, která se rozhodování neúčastnila. Tam, kde vynaložené úsilí neregistruje žádný systém hodnocení, je rovnovážným výsledkem uložení za nejnižších nákladů, jež vyhoví kontrole poskytovatele podpory, což je racionální odpověď na nastavené pobídky, a nikoli nedostatkem pečlivosti. Rámec proto zachází s uznáním jako s proměnnou, jež podmiňuje ostatní proměnné této dimenze.

11.1 Tři často slučované funkce

Zálohování, dlouhodobé uchování a publikování odpovídají na odlišné otázky a selhávají odlišnými způsoby a jejich slučování je nejčastější strukturní chybou institucionální praxe. Zálohování obnovuje provozní stav po ztrátě a měří se cíli obnovy, jak stanovil oddíl 9.5. Dlouhodobé uchování udržuje schopnost budoucí komunity objekt najít, ověřit a interpretovat, a je proto organizačním závazkem drženým lidmi, systémy a postupy, nikoli uspořádáním úložiště (Consultative Committee for Space Data Systems 2012; International Organization for Standardization 2025). Publikování činí objekt citovatelným a vyhledatelným, přičemž publikovaný objekt s persistentním identifikátorem nemusí být v žádném aktivním smyslu ani zálohován, ani dlouhodobě uchováván.

Referenční model přispívá dvěma konstrukty, jež rámec přejímá, a jedním termínem, jež hlídá. Informační balíčky oddělují to, co je předáno, to, co je uloženo, a to, co je zpřístupňováno, a plánování dlouhodobého uchování je průběžnou funkcí, nikoli událostí v okamžiku příjmu. Určená komunita fixuje znalostní základnu předpokládanou u budoucích uživatelů a převádí jinak neohrazenou otázku, kolik dokumentace postačuje, na otázku posouditelnou, neboť dostatečnost je relativní vůči uvedené komunitě. Slovo otevřený v názvu modelu označuje otevřený vývoj standardu, a nikoli neomezený přístup, a rámec právě z tohoto důvodu drží dlouhodobé uchování a přístup jako samostatné proměnné.

11.2 Výběr k uchování a retenční lhůta jako rozhodnutí s vlastníky

Uchovávat vše není v institucionálním měřítku udržitelné finančně ani environmentálně a dlouhodobé uchování to navíc oslabuje tím, že rozptyluje úsilí pryč od objektů, které je opodstatňují (National Digital Stewardship Alliance 2026). Výběru k uchování se tedy nelze vyhnout a otázkou zůstává, zda k němu dojde jako ke zdokumentovanému rozhodnutí, nebo jako k úbytku. Kritérii, jež korpus konzistentně uvádí, jsou vědecká a důkazní hodnota, nenahraditelnost, právní postavení, náklady na uchování a možnost objektu regenerovat, přičemž vážit je lze pouze vůči stanovené retenční lhůtě. Vyřazení rozhodnuté podle těchto kritérií a zaznamenané je aktem správy dat; vyřazení uplynutím smlouvy o úložišti je ztrátou, za níž nestojí žádné rozhodnutí.

O tom, zda výběr k uchování ob stojí ve styku s praxí, rozhodují dvě strukturní vlastnosti. Retenční lhůty vznikají z několika zdrojů zároveň — z podmínek poskytovatele podpory, z právních povinností, z oborových norem a z vlastní politiky instituce —, a rozcházejí-li se zdroje navzájem, je namíste pravidlo uspořádání a pojmenovaná autorita spíše než delší výchozí hodnota. A povinnost výmazu vzniklá podle práva ochrany osobních údajů může dopadat na materiál, na nějž dopadá i retenční povinnost, což je konflikt neřešitelný uvnitř kteréhokoli z obou nástrojů a směřovaný rámcem k eskalační autoritě identifikované v kapitole 5. Zaznamenávanými proměnnými jsou tedy retenční lhůta, její zdroj, rozhodnutí o výběru k uchování a role, která je přijala.

Institucionální audit zaznamenávající pro vymezenou populaci ukončených projektů, zda byla stanovena retenční lhůta, zda bylo zdokumentováno rozhodnutí o výběru k uchování a jaký podíl projektových dat zůstal po ukončení projektu na nesprávaném úložišti

11.3 Zastarávání, migrace a integrita

Formáty selhávají dvěma způsoby, jež vyžadují odlišnou odpověď: software schopný formát zobrazit přestane existovat, anebo se specifikace formátu ztratí či nikdy nebyla veřejná. Migrace do otevřeného nebo široce podporovaného formátu pokrývá první z nich a může pozměnit přesnost, funkčnost nebo význam, zvláště u výstupů proprietárních přístrojů; emulace zachovává původní chování za cenu nákladů, jež se opakují bez konce. Volbu mezi nimi určují významné vlastnosti, které si objekt zachovává, což je vědecký soud o tom, čeho je objekt důkazem, a nikoli technická preference. U proprietárních vědeckých formátů uchovává uspořádání nejkonzistentněji uváděné jako obhajitelné originál, zdokumentovaný export a dokumentaci konverze pohromadě.

OVĚŘENÍ integrity se běžně pokládá za rozhodující ve větším rozsahu, než jaký mu náleží. Kontrolní součet přepočtený a porovnaný dokládá, že se bitová posloupnost nezměnila, a nedokládá nic o tom, zda objekt zůstává interpretovatelný ani zda migrace zachovala jeho význam. Obě vlastnosti proto vyžadují oddělený důkaz: kontroly bitové neměnnosti s datovanými výsledky a validaci, že významné vlastnosti transformaci formátu přežily. Slovník metadat pro dlouhodobé uchování poskytuje strukturu, v níž se takové události, agenti, kteří je provedli, a práva, jež je upravují, zaznamenávají, takže historie uchování objektu je sama prozkoumatelným záznamem (Library of Congress 2015).

Dlouhodobá správa sahá za objekt až ke službě, jež jej drží. Záznamy o verzích, náhrobní záznamy stažených objektů, změny licence a plán nástupnictví popisují, co nastane, změní-li objekt nebo repozitář svůj stav. Repozitář, který nedokáže uvést, co se s jeho fondy stane v případě ukončení činnosti, se nevypořádal se způsobem selhání, jenž jeho ukladatele ohrožuje nejvíce. Certifikační požadavky se posunuly tímto směrem a posílily důkazy očekávané u udržitelnosti, sdílené odpovědnosti a plánování ukončení činnosti (CoreTrustSeal 2022). Tato studie zachází s ujednáním o ukončení

jako s vlastností rozhodnutí o uložení, a nikoli jako s provozní záležitostí vnitřní repozitáři.

11.4 Výběr repozitáře jako posouzení

Výběr repozitáře se běžně vyřizuje zvykem, volbou, kterou učinil kolega, nebo tým, která služba vydá identifikátor s nejmenším třením. Pojat jako posouzení se rozpadá na kritéria, jež lze stanovit předem a uplatnit na kandidáty. Patří k nim oborová vhodnost, povinné požadavky poskytovatele podpory, instituce a časopisu, podporované režimy přístupu, persistentní identifikace, kurátorství a kvalita metadat, limity velikosti a formátu, verzování, náklady, závazky služby a ujednání platná pro okamžik, kdy služba skončí. Registrace a certifikace jsou odlišné signály a bývají zaměňovány: záznam v registru naznačuje, že služba existuje a jak se popisuje, kdežto certifikace naznačuje, že zdokumentované požadavky posoudil externí orgán (re3data.org 2026; CoreTrustSeal 2022). Tabulka 11-1 uvádí kritéria a důkaz, jímž je každé z nich naplněno.

Tabulka 11-1. Kritéria výběru repozitáře, s důkazem naplňujícím každé z nich a se zbytkovým rizikem při jeho absenci.

Kritérium	Naplňující důkaz	Zbytkové riziko při absenci
Oborová vhodnost	Metadatový standard přijatý komunitou; specializované kurátorství	Záznam nalezen, avšak vědecky neinterpretovatelný
Povinné požadavky	Ověřené podmínky poskytovatele podpory, instituce a časopisu	Uložení povinnost nevyřizuje
Certifikace	Platná externí certifikace ověřená u vydávajícího orgánu	Důvěryhodnost tvrzená, nikoli posouzená
Persistentní identifikace	Identifikátory ražené a rozklíčované, s verzováním	Citace se láme; k verzi neexistuje stabilní odkaz
Režimy přístupu	Podpora otevřeného, embargovaného, řízeného i pouze metadatového režimu	Citlivá data vtlačena do nevhodného režimu
Ujednání o ukončení	Zdokumentovaný plán nástupnictví nebo převodu; export metadat	Fondy rozptýleny bez předání kustodie

Kritéria se dostávají do konfliktu a příspěvek rámce v tomto bodě spočívá v tom, že konflikt pojmenovává, namísto aby jej ukryl do pořadí. Oborový

repozitář často nabízí nejbohatší metadata a komunitu, u níž je opětovné použití dat nejpravděpodobnější, a přitom mu může scházet certifikace, kapacita pro velké objemy nebo podpora řízeného přístupu; certifikovaný generalistický repozitář nabízí posouzenou důvěryhodnost a slabší vědecký popis. Tam, kde konflikt vyřešit nelze, je dostupné rozdělené uložení propojené persistentními identifikátory, za cenu udržování tohoto propojení v čase. Důkazy podírající tyto varianty jsou odstupňovány stupněm B, neboť spočívají na případových zprávách a na dokumentaci služeb, a nikoli na srovnávacím měření výsledků opětovného použití.

CHYBĚJÍCÍ DOKLAD

Srovnávací měření opětovného použití, citovanosti a dlouhodobé dostupnosti datových sad uložených v oborových oproti certifikovaným generalistickým repozitářům, při kontrole oboru, aby bylo možné stanovit, zda má kompromis mezi oborovou vhodností a certifikovanou důvěryhodností měřitelné důsledky

11.5 Publikování dat jako samostatný vědecký akt

Uložení a publikování jsou v rámci odděleny proto, že nesou odlišné povinnosti. Datová sada publikovaná jako výsledek získává identifikátor, citaci, licenci, popsanou verzi a seznam autorů a stává se objektem, na nějž mohou jiné práce přesně odkazovat (Wilkinson et al. 2016). Tento akt je vědecký, nikoli administrativní. Výběr toho, co tvoří publikovatelný objekt, popis dostatečný pro neznámého uživatele a rozhodnutí o tom, která verze je autoritativní, jsou analytickými soudy, a jde právě o soudy, jež kapitoly 6 a 8 přiřazují životnímu cyklu a popisu. Datová sada připojená k článku jako doplňkový materiál žádnou z těchto vlastností nemá a nelze ji citovat, verzovat ani posuzovat samostatně.

Dva rysy publikování dat mají důsledky pro vrstvu posuzování. Citace dat je možná pouze tam, kde se identifikátor rozkládá a kde ji metadatový záznam podpírá, takže uznání za datovou sadu závisí na infrastruktuře, již autor nekontroluje (Cousijn et al. 2019). A příspěvek k datové sadě je zpravidla rozprostřen mezi sběr, kurátorství, popis a software, což autorská konvence článku vyjadřuje jen slabě; taxonomie rolí přispěvatelů existují proto, aby tyto příspěvky učinily odděleně viditelnými. Rámec zaznamenává publikovaný objekt, jeho identifikátor, jeho verzi a role přispěvatelů jako

proměnné, a to na základě úvahy, že uznání, jež nelze přiřadit, nelze ani posoudit.

11.6 Uznání jako strukturní proměnná

Reforma hodnocení výzkumu vychází z pozice, že úzký soubor publikačních metrik vyjadřuje výzkumný příspěvek nedostatečně a že data, software, metody, kurátorství a týmové role opodstatňují uznání samy o sobě (CoARA 2022; DORA 2013). Vlastní riziko této reformy je v korpusu patrné: nahrazení počtu článků počtem datových sad reprodukuje mechanismus, jež mělo napravit, neboť počet uložení je vůči kvalitě, dokumentaci a opětovnému použití stejně necitlivý jako počet článků. Odpovědné posuzování proto vyžaduje kvalitativní doklad příspěvku interpretovaný v oborovém kontextu a indikátory podpořený, a nikoli jimi nahrazený. Statistiky užití a citovanosti dat jsou užitečné a zároveň nejsou srovnatelné napříč obory a jejich manipulovatelnost je prozkoumána natolik, aby bránila jejich užití jako jediného základu.

Česká implementační vrstva poskytuje konkrétní instanci obecného pohybu a test jeho důsledků. Od roku 2026 je digitální datová kolekce uznána v národním systému hodnocení jako samostatný druh výsledku, s registrací v národním metadatovém adresáři jako související podmínkou (Úřad vlády České republiky 2025; EOSC CZ 2025). Plynou z toho dva účinky analytický oddělitelné od národního prostředí. Prvním je, že datový výstup se stává posouditelným objektem, takže kurátorské a popisné úsilí vstupuje do záznamu, jímž jsou instituce financovány. Druhým je, že požadavek na registraci metadat převádí kvalitu popisu z profesní normy na podmínku uznání, což mění, kdo uvnitř instituce má na ní zájem.

Tato studie zachází s českým uspořádáním jako s instancí, a nikoli jako s vodítkem, a její analytický zájem se upíná k mechanismu, jež toto uspořádání odkrývá. Uznání funguje jako proměnná, která podmiňuje ostatní proměnné této dimenze, protože výběr k uchování, rozhodnutí o formátech, posouzení repozitáře i popis spotřebovávají úsilí, jež je pro systém hodnocení postavený pouze na počtech publikací neviditelné. Tam, kde je uznání přítomno, popisují kritéria posouzení v rámci činnosti, kterou má instituce důvod udržet; tam, kde chybí, popisují tatáž kritéria nefinancované očekávání, a hranice platnosti celého rámce je odpovídajícím způsobem

ostrá. Předpověď, kterou tato studie z uvedeného mechanismu odvozuje, je testovatelná a důkazy potřebné k jejímu otestování dosud neexistují.

CHYBĚJÍCÍ DOKLAD

Měření objemu uložení, úplnosti metadat a poskytování řízeného přístupu na českých výzkumných organizacích před zavedením druhu výsledku datová kolekce a po něm, oproti srovnávací skupině bez takového uznání, aby bylo možné otestovat, zda uznání mění chování ve správě dat, a nikoli chování ve vykazování

11.7 Proměnné přenášené dále

Tabulka 11-2 uvádí proměnné, jimiž tato dimenze přispívá do integrovaného rámce. Každá pojmenovává záznam a každá je prozkoumatelná i poté, co projekt, který data vytvořil, skončil, což je podmínka, k jejímuž naplnění tato dimenze existuje.

Tabulka 11-2. Proměnné dlouhodobého uchování, publikování a uznání v integrovaném rámci a jejich doložitelné záznamy.

Proměnná	Definice	Doložitelný záznam
Určená komunita	Znalostní základna předpokládaná u budoucích uživatelů	Výrok politiky dlouhodobého uchování
Retenční lhůta	Lhůta, její zdroj a autorita, která rozhoduje spor	Registr retenčních lhůt; záznam o eskalaci
Rozhodnutí o výběru k uchování	Co se uchovává, co se vyřazuje a podle jakých kritérií	Datovaný záznam o výběru s uvedením rozhodující role
Významné vlastnosti	Vlastnosti, které si objekt zachovává napříč transformací	Posouzení formátu; dokumentace migrace
Důkaz integrity	Datované výsledky kontroly bitové neměnnosti a validace po migraci	Záznamy o událostech dlouhodobého uchování
Posouzení repozitáře	Kandidáti posouzení podle stanovených kritérií, s odůvodněnou volbou	Záznam o výběru odkazovaný v plánu správy dat
Ujednání o ukončení	Co se stane s fondy, skončí-li služba	Zdokumentovaný plán nástupnictví; ustanovení o exportu metadat
Publikovaný objekt	Identifikátor, verze, licence a role přispěvatelů	Záznam v repozitáři; citační metadata

Proměnná	Definice	Doložitelný záznam
Cesta uznání	Jak úsilí vynaložené na správu dat vstupuje do institucionálního hodnocení	Pravidla hodnocení; záznamy o vykázání výsledku

Dimenzi omezují tři hranice platnosti. Povinností dlouhodobého uchování u velmi velkých objemů dat nelze dostát uchováním všeho a odpovědí rámce je zdokumentovaný výběr k uchování namísto nenaplněného univerzálního závazku. Certifikační status je časově omezený a přezkoumatelný, takže záznam o výběru citující certifikaci bez data ověření uvádí méně, než se jeví. A uspořádání uznání jsou specifická pro jurisdikci a nová, takže mechanismus popsáný v oddílu 11.6 je formulován jako podmíněná předpověď, a nikoli jako ustavená pravidelnost, dokud nebude provedeno měření, jež výše uvedená značka specifikuje.

Pět dimenzí rozvinutých v části III doplňuje soubor proměnných rámce. Zda si tyto sdílené proměnné uchovávají rozlišovací sílu, jsou-li aplikovány na data výrazně odlišného druhu, a kde se jejich aplikace hroutí, je otázkou, jíž se následující část zabývá prostřednictvím oborových profilů a integrovaného případu.

Kapitola 12. Operacionalizace rámce: rozhodovací pravidla a úrovně zralosti

Sedm předchozích kapitol ustavuje soubor proměnných. Rámec tím ustaven ještě není, neboť souborem proměnných se rámec stává teprve tehdy, když je uvedena jeho struktura, deklarovány jeho vztahy závislosti, specifikováno jeho chování při konfliktu a jeho stavy jsou pozorovatelné i pro někoho jiného než pro popisovanou stranu. Tato kapitola tuto skladbu provádí a nese hlavní původní příspěvek knihy. Teze, kterou obhájí, zní, že 55 proměnných z kapitol 5–11 se skládá do jediné struktury o čtyřech úrovních, uspořádané obhajitelným vztahem podmíněnosti, uzavřené pod pěti rozhodovacími pravidly a čitelné jako profil pozorovatelných stavů, a nikoli jako skóre.

Integrace byla v kapitole 1 položena jako otázka, zda lze částečné požadavky převést na obhajitelná rozhodnutí. Struktura, jež ponechává prioritu mezi konkurujícími si požadavky neuvedenu, na ni neodpovídá, neboť konflikty identifikované v oddílu 4.5 jsou přesně tím místem, kde částečné nástroje selhávají. Následující výklad je proto falzifikovatelný v bodě konfliktu, a nikoli v bodě popisu, a zůstává konstrukcí, jejíž validace je předmětem části IV.

12.1 Čtyři úrovně: dimenze, komponenta, proměnná, indikátor

Rámec rozlišuje čtyři úrovně a toto rozlišení není prezentační. Dimenze je seskupením požadavků sdílejících kódovací modus a podmíněnost platí právě na této úrovni. Komponenta je souborem proměnných připojených k jednomu rozhodovacímu objektu a k jedné straně odpovědné za výsledek a na této úrovni instituce alokuje práci. Proměnná je vlastností datové činnosti s definovaným stavovým prostorem a s pojmenovaným doložitelným záznamem a na této úrovni se přijímá rozhodnutí.

Čtvrtá úroveň je v této kapitole nová a doplňuje to, co kapitoly 5–11 ponechaly otevřené. Tyto kapitoly pro každou proměnnou pojmenovávají záznam, z něhož se její stav čte, a záznam sám o sobě čtení nefixuje. Indikátor je záznamem spolu s operací, jež se na něm provádí, s výběrovou jednotkou a s datem pozorování. Jediné uspořádání podpírá čtení nestejné

síly — plán obnovy, datovaná zpráva o testu obnovy a její výsledek —, a vlastnosti rámce vztahující se k posouzení závisí na tom, které z nich je specifikováno.

Rámec ve složené podobě nese sedm dimenzí namísto šesti seskupení z tabulky 4-2. Plánování bylo odděleno proto, že jeho proměnné jsou závazky fixované před činností, a nikoli jejími popisy, a protože strana odpovědná za jejich výsledek se liší od strany odpovědné za proměnné governance, k nimž je kódování seskupilo. Kvalita byla přesunuta k seskupení popisu, neboť vhodnost k použití i interpretovatelnost se čtou z týchž záznamů. Obojí se dotýká spolehlivosti postupu odvození a posiluje důvody pro překódování provedené v oddílu 4.4. Tabulka 12-1 uvádí výsledek.

Tabulka 12-1. Sedm dimenzí, jejich komponenty, proměnné, jimiž každá přispívá, a jejich postavení v pořadí podmíněnosti.

Dimenze	Komponenty	Proměnné	Postavení
D1 Governance a odpovědnost za výsledek	Autorita; přiřazení a kontinuita; služba a měření	6	podmiňující
D2 Životní cyklus a provenience	Identita objektu; odvození; přetrvání způsobilosti	8	podmiňující
D3 Plánování jako závazek	Specifičnost závazku; výměna; revize	7	mezilehlé
D4 Popis a kvalita	Struktura popisu; určení kvality	8	podmíněná
D5 Ochrana a přístup	Klasifikace a právní titul; přístup a jeho ohraničení; práva	8	podmiňující
D6 Reprodukovatelnost	Objekt a prostředí; důkaz o spuštění; analytická disciplína	9	podmíněná
D7 Dlouhodobé uchování, publikování a uznání	Výběr k uchování; interpretovatelnost; kustodie a zpřístupnění; uznání	9	podmíněná, se zpětnou hranou

12.2 Pořadí podmíněnosti a jeho obhajoba

Tři dimenze podmiňují ostatní, každá v jiné měně. Governance přenáší autoritu, neboť rámec bez pojmenované autority produkuje analýzy, a nikoli určení. Životní cyklus a provenience přenášejí posouditelnost, neboť každá proměnná pojmenovává záznam a záznamy existují pouze tam, kde je spra-

vována identita objektu, jeho odvození a jeho kustodie. Ochrana přenáší množinu variant, protože třída dat určuje, která úložná prostředí, které režimy přístupu a které cesty publikování zůstávají dostupné.

Plánování zaujímá mezilehlé postavení, jež tabulka 12-2 zobrazuje, namísto aby je uhladila. Jeho proměnné jsou podmíněny těmito třemi dimenzemi, neboť závazek nelze přijmout bez autority, ověřit bez záznamu ani rozšířit nad to, co třída dat připouští, a jeho nevratné závazky pak podmiňují dimenze následující.

Tabulka 12-2. Vztahy podmíněnosti mezi dimenzemi a to, co každý vztah přenáší.

Dimenze	Podmíněna	Podmiňuje	Přenášeno
D1	—	D2–D7	oprávnění rozhodovat a přijímat zbytkové riziko
D2	D1	D3–D7	existence a provázanost záznamů, jež ostatní proměnné pojmenovávají
D5	D1, D2	D3, D4, D6, D7	přípustná množina variant, prostřednictvím třídy dat
D3	D1, D2, D5	D4, D6, D7	závazky, jež nelze revidovat bez ztráty
D4	D1–D3, D5	D6, D7	interpretovatelnost, bez níž je regenerace nedefinovaná
D6	D1–D5	D7	důkazní obsah toho, co se uchovává
D7	D1–D6	všechny, prostřednictvím uznání	zdroje a pozornost

Uspořádání se neopírá o tvrzení, nýbrž o tři testy. Odstranění podmiňující dimenze ponechává podmíněnou dimenzi bez určitého stavu, kdežto obráceně to neplatí. Žádná podmíněná proměnná nemění stav proměnné podmiňující, takže vztah je asymetrický všude, kde je tvrzen. A test uzavřenosti z oddílu 4.6 je naplněn, protože D1 nemá předchůdce.

Jeden vztah běží proti tomuto pořadí záměrně. Komponenta uznání v D7 podmiňuje každou dimenzi, protože úsilí neviditelné pro institucionální hodnocení se neudrží a výběr k uchování, popis, kurátorství i zachycení prostředí spotřebovávají právě takové úsilí (CoARA 2022). Zpětná hrana nevytváří v závislosti rozhodnutí žádný cyklus, neboť uznání určuje, zda je činnost zdrojově zajištěna, a nikoli to, co daná činnost určuje. Její přítomnost

přesto znamená, že kritéria posouzení popisují nefinancované očekávání všude tam, kde cesta uznání chybí.

CHYBĚJÍCÍ DOKLAD

Průřezové měření stavů proměnných uvnitř souboru projektů, testující, zda podmíněné proměnné kovariují se svými podmiňujícími proměnnými tak, jak tvrdí tabulka 12-2 — tedy pozorování, jež by uspořádání falzifikovalo.

12.3 Rozhodovací pravidla pro konkurující si požadavky

Pravidla tvoří uspořádaný postup uplatňovaný na konflikty mezi správně uplatněnými požadavky. Pravidlo 1 odděluje přípustnost od priority: požadavek, od něž se nelze odchýlit — právní povinnost, závazná retenční lhůta, rozsah platného souhlasu —, odstraňuje varianty, namísto aby uvnitř množiny soupeřil. Pravidlo 2 uspořádává zbytek podle nevratnosti a staví zveřejnění, výmaz a ztrátu spustitelného prostředí nad chudá metadata a opožděnou registraci. Pravidlo 3 se uplatní tam, kde obě varianty zůstávají vratné, a dává přednost požadavku nesoucím vyšší stupeň důkazu podle oddílu 4.5.

Zbytek vyřizují dvě další pravidla. Pravidlo 4 se uplatní tam, kde jsou stupně shodné, a dává přednost požadavku připojenému k vyšší třídě dat, přičemž náklady se posuzují vůči dopadu této třídy, a nikoli vůči rozpočtové položce. Pravidlo 5 upravuje to, co zbývá: nerozlišuje-li žádné předchozí pravidlo, přijímá rámec variantu zachovávající větší budoucí množinu variant — uchování před vyřazením, řízený přístup před otevřeným zpřístupněním — a zaznamenává rozhodnutí jako přijaté za nevyřešeného konfliktu se stanovenou podmínkou přezkumu. Pravidlo 5 je výchozím nastavením, a nikoli určením, a zacházet s ním jako s určením je nejpravděpodobnějším zneužitím rámce. Tabulka 12-3 uvádí pravidla a jejich authority.

Tabulka 12-3. Rozhodovací pravidla, autorita oprávněná každé z nich uplatnit a vzniklý záznam.

Pravidlo	Uplatní se, když	Určení	Rozhodující autorita	Záznam
1 Přípustnost	Od požadavku se nelze odchýlit	Odstraňuje varianty z množiny	Právní útvar, pověřenec pro ochranu osobních údajů nebo etická komise	Posouzení s uvedením daného ustanovení
2 Nevratnost	Varianty se liší napravitelností	Dává přednost napravitelnému porušení	Strana odpovědná za výsledek dané činnosti	Záznam o rozhodnutí s uvedeným důvodem
3 Stupeň důkazu	Obě varianty jsou napravitelné	Dává přednost požadavku s vyšším stupněm	Strana odpovědná za výsledek, na základě porady	Záznam o rozhodnutí s citovaným stupněm a zdrojem
4 Proporcionalita třídy	Stupně jsou shodné	Dává přednost vyšší třídě dat	Vlastník registru klasifikace	Odkaz na klasifikaci v záznamu
5 Zachování variant	Žádné pravidlo nerozlišuje	Přijímá vratnou variantu předběžně	Příjemce eskalace pojmenovaný v D1	Záznam o nevyřešeném konfliktu s podmínkou přezkumu

Lexikální uspořádání nese věcný závazek v tom, že vratnost převažuje nad důkazní silou. Slabě doložený požadavek, jehož porušení je trvalé, tak převažuje nad silně doloženým požadavkem, jehož porušení lze napravit, což obrací pořadí produkované čistě důkazním uspořádáním. Toto obrácení je obhájeno podmínkami rozhodnutí: budoucí užití je nejisté a spravovaným aktivem je právě množina variant.

Každé uplatnění pravidel 1–5 produkuje zbytkové riziko, s nímž rámec zachází jako s výstupem majícím vlastní záznam. Tento záznam nese šest polí: proměnné v konfliktu, přijatou variantu, variantu opuštěnou, riziko, jež ponechává, roli, která toto riziko přijímá, a podmínku, za níž se rozhodnutí vrací k přezkumu. Rozhodnutí zaznamenané bez pátého pole je zpětně nerozlišitelné od toho, že žádné rozhodnutí přijato nebylo, a šesté pole brání

tomu, aby omezení přijaté podle pravidla 5 přetrvávalo poté, co jeho důvod odpadl.

CHYBĚJÍCÍ DOKLAD

Uplatnění pravidel 1–5 na pevný soubor zdokumentovaných konfliktních případů několika nezávisle pracujícími posuzovateli, s uvedením podílu případů, jimž se dostane téhož určení a téhož přiřazení zbytkového rizika.

12.4 Zralost jako pozorovatelný stav

Úrovně zralosti se v tomto oboru běžně uvádějí jako aspirační nálepky, jejichž kritéria popisují záměr, a takovým škálám vyhoví instituce, která se umí dobře popsat (Paulk et al. 1993; Crowston and Qin 2011). Rámec namísto toho definuje pět úrovní jako pozorovatelné stavy jediné proměnné, přičemž každá je od úrovně nižší odlišena jedinou pozorovatelnou skutečností, kterou posuzovatel může hledat a nenalézt. Úrovně se připojují k proměnným, a nikoli k institucím, a institucionální zralost je rozdělením 55 proměnných napříč nimi. Profil se třemi proměnnými na úrovni 0 opodstatňuje jinou odpověď než rovnoměrně průměrný profil s týmhž průměrem a jediné číslo tento rozdíl zahazuje. Tabulka 12-4 uvádí pět úrovní.

Tabulka 12-4. Úrovně zralosti jako pozorovatelné stavy, s pozorovatelnou skutečností odlišující každou úroveň od úrovně nižší.

Úroveň	Stav	Rozlišující pozorovatelná skutečnost	Typický záznam
0 Nepřiřazená	Proměnná nemá čitelný stav	—	žádný
1 Deklarovaná	Dokument stanoví požadavek pro danou třídu činností	Dokument, jenž požadavek pojmenovává a na činnost dopadá	Politika; šablona; položka plánu
2 Přiřazená	Pro danou činnost existuje pojmenovaná role a práh	Záznam pojmenovávající roli a datum, nikoli útvar nebo záměr	Matice přiřazení; položka katalogu
3 Doložená	Pojmenovaný záznam u vzorkovaných činností existuje a odpovídá deklaraci	Párované srovnání závazku se stopou, úspěšné na vzorku	Log; validační záznam nebo záznam o uložení

Úroveň	Stav	Rozlišující pozorovatelná skutečnost	Typický záznam
4 Otestovaná	Určení obstálo v události, kterou instituce nenaplánovala	Datovaný výsledek obnovy, opakovaného spuštění, žádosti o přístup nebo incidentu	Zpráva o testu; záznam o incidentu s vyřešením

Většina institucionálních sebepopisů končí mezi úrovněmi 1 a 2, kdežto tvrzení rámce o posouzení se soustřeďuje mezi úrovněmi 2 a 3. Úroveň 3 je dosažitelná pouze prostřednictvím záznamu vzniklého jako vedlejší produkt činnosti, a nikoli jako artefakt posuzování, což je právě to, co činí její předstírání nákladným. Úroveň 4 přidává podmínku, že načasování pozorování nekontrolovala pozorovaná strana, a je jedinou úrovní testující určení proti skutečnosti, a nikoli proti dokumentaci. Rozlišovací síla těchto prahů je tvrzením o měření a změřena nebyla.

CHYBĚJÍCÍ DOKLAD

Posouzení 55 proměnných na institucích vybraných tak, aby se lišily v nezávislých indikátorech výkonu ve správě výzkumných dat, testující, zda prahy mezi úrovněmi 2 a 3 a mezi úrovněmi 3 a 4 oddělují instituce, jež nezávislé důkazy rozlišují.

12.5 Automatizace a strojové učení uvnitř rámce

Automatizace uplatněná na výzkumná data, včetně strojového učení jako metody i jako předmětu správy, další dimenzi netvoří a důvod je strukturní. Dimenze je opodstatněná tam, kde třída požadavků adresuje rozhodovací objekt, jež žádná stávající dimenze neadresuje. Požadavky vznikající zde se připojují vždy k objektu, jež sedm dimenzí již nese, jak uvádí tabulka 12-5 (Gebu et al. 2021; Mitchell et al. 2019; National Institute of Standards and Technology 2023). Mění se stav stávajících proměnných a umístění jejich důkazu.

Jedna vlastnost se strukturně mění a odpověď rámce na ni je nejsilnějším argumentem proti samostatné dimenzi. Automatizace přesouvá agenta datové činnosti z osoby na systém a záznam o odvození v D2 již pole agenta nese. Rozhodovací hranice mezi poradou a rozhodnutím z D1 se pak přenáší beze změny: automatická klasifikace, navržená hodnota metadat nebo doporučené určení přístupu jsou poradou a rozhodnutím je jejich přijetí

stranou, která má autoritu. Pojímat strojový výstup jako rozhodnutí opakuje kategoriální omyl, jímž orientační posouzení začíná fungovat jako schválení.

Tabulka 12-5. Kam se požadavky na automatizaci a strojové učení připojují.

Požadavek	Dimenze a proměnná
Provenience a oprávnění k trénovacímu materiálu	D2 stopa práv; D5 právní titul, inventář práv
Složení datové sady, vyloučení a omezení	D4 aplikační profil, deklarovaný účel, výrok o zbytkové chybě
Oddělení trénovací, validační a testovací části	D6 oddělení objektů, časování rozhodnutí
Verzování modelu, parametrů, kódu a prostředí	D6 specifikace prostředí, vazba verzí
Externí služba použitá na neveřejné vstupy	D5 třída dat, kontrola výstupu z prostředí; D1 rozhodovací hranice
Posouzení modelu, drift a adversariální testování	D4 opatření kontroly kvality; D6 výrok o nejistotě

Pro posuzování z toho plynou dva důsledky. Nástroje adresované umělé inteligenci zvyšují stupeň důkazu u některých požadavků a posouvají eskalační bod u některých tříd užití, přičemž ani jeden z těchto účinků nevyžaduje novou proměnnou (European Parliament and Council 2024). Automatizované monitorování kontrolních opatření mění náklady pozorování, a nikoli pozorovanou úroveň, takže činí úroveň 3 levnou, kdežto úroveň 4 produkuje pouze tam, kde bylo obstáno v nenaplánované události.

CHYBĚJÍCÍ DOKLAD

Kódování formulací požadavků z korpusu governance a dokumentace umělé inteligence vůči 55 proměnným, s uvedením podílu formulací přiřaditelných stávající proměnné a s charakteristikou případného zbytku — tedy test toho, zda je tabulka 12-5 vyčerpávající, a nikoli ilustrativní.

12.6 Status konstrukce

To, co část III předkládá, je konstrukce s vlastnostmi dvojího druhu. Vnitřní vlastnosti se zakládají prozkoumáním a tři testy z oddílu 4.6 vracejí následující: každou proměnnou lze připojit k pojmenované roli, každá

proměnná pojmenovává záznam a každý konflikt je směřován k D1. Vnější vlastnosti prozkoumáním založit nelze vůbec, neboť rozlišovací síla indikátoru, spolehlivost pravidla a oddělovací síla prahu zralosti jsou měřeními, a nikoli odvozeními. Žádné tvrzení o validitě se zde nevznáší a výše uvedené značky uvádějí, co by muselo být změřeno, aby se každé takové tvrzení stalo dostupným.

Část IV tuto práci zahajuje. Následující kapitola aplikuje rámec na institucionální prostředí a v celém svém průběhu odlišuje rekonstrukci proti zdokumentovaným zdrojům od empirické aplikace proti vlastním záznamům instituce.

Č Á S T I V

Validace

Kapitola 13. Aplikace na institucionální prostředí

Část III se uzavírá konstrukcí, jejíž vlastnosti byly založeny pouhým prozkoumáním. Tato kapitola otevírá část, v níž se konstrukce setkává s materiálem, jež sama nevytvořila, a obhazuje tvrzení o tom, co takové setkání může založit. Aplikace rámce na určité prostředí ukazuje, zda jsou jeho proměnné v tomto prostředí vyjádřitelné, zda se jeho postup uzavírá a kde jeho pravidla ponechávají konflikt otevřený; neukazuje, že je rámec validní, a žádná zde uváděná aplikace se za takový důkaz nevydává. Kapitola proto odděluje dvě operace, které literatura o vývoji rámců běžně slučuje, a uvádí, která z nich byla provedena.

S hlavní výzkumnou otázkou to souvisí posloupností. Kapitola 1 se ptá, za jakých podmínek lze částečné požadavky integrovat do rámce přinášejícího obhajitelná rozhodnutí, a konstrukce odpovídá pouze na první polovinu této otázky. Na druhou polovinu odpovídá styk s prostředím, v nichž si požadavky skutečně konkurují, a hodnota takového styku závisí zcela na tom, zda čtenář dokáže odlišit rekonstrukci od měření. Toto rozlišení je uspořádávajícím principem toho, co následuje.

13.1 Logika výběru případů

Prostředí byla vybrána tak, aby maximalizovala variabilitu ve čtyřech kontextových podmínkách, jež kapitola 14 identifikuje jako zdroje citlivosti: oborová konvence, právní režim, institucionální velikost a dostupná infrastruktura. Výběr tohoto druhu je teoretický, nikoli statistický, a podpírá tvrzení o rozpětí, napříč nímž zůstává rámec vyjádřitelný, a nikoli tvrzení o tom, jak jsou instituce rozloženy (Seawright and Gerring 2008). Design vybírající pouze příznivá prostředí nemůže vydat negativní výsledek, a soubor proto zahrnuje jedno prostředí, v němž se u několika proměnných rámce očekává, že čitelné nebudou. Tabulka 13-1 uvádí čtyři prostředí, to, co každé z nich obměňuje, a to, co je každé schopno testovat.

Tabulka 13-1. Prostředí vybraná pro aplikaci, s podmínkou, kterou každé obměňuje, a s určením, jež každé testuje.

Prostředí	Obměňuje	Testuje	Status
S1 Výzkumně intenzivní multioborová univerzita, federované uspořádání, národní systém uznání (Úřad vlády České republiky 2025)	velikost a infrastruktura v nejpříznivější podobě	zda je všech sedm dimenzí čitelných tam, kde existuje každý druh záznamu	rekonstrukce
S2 Malá instituce bez vlastního repozitáře, bez funkce ochrany osobních údajů a bez služby dlouhodobého uchování	velikost a infrastruktura v nejméně příznivé podobě	zda proměnné přežijí tam, kde je naplňuje vnější způsobilost	rekonstrukce
S3 Mezinárodní konsorciální projekt kombinující senzorová data, neveřejné stavební výkresy, rozhovorový materiál, licencované zdroje třetích stran a vlastní klasifikační model	právní režim a heterogenita objektů	zda se rozhodovací pravidla uzavrou při současně probíhajícím konfliktu	rekonstrukce
S4 Párová oborová prostředí, jedno interpretativní a jedno přístrojově vybavené	oborová konvence	zda si proměnné podržují rozlišovací sílu tam, kde se liší podoby záznamů	rekonstrukce

Volba prostředí S3 vyžaduje slovo zdůvodnění, neboť jednotku analýzy zde zaujímá projekt, a nikoli instituce. Rámec byl v kapitole 1 vázán vyjádřit institucionální i projektovou úroveň týmiž proměnnými a prostředí, v němž vzniká mnoho konfliktů naráz, je jediným prostředím, v němž může tato vázanost viditelně selhat. Projekt kombinující veřejné senzorové měření, neveřejné výkresy infrastruktury, osobní rozhovorový materiál, externě licencované vstupy a model použitý v bezpečnostně relevantním určení uvádí v činnost všech pět rozhodovacích pravidel. Prostředí, v nichž vzniká vždy jen jeden konflikt, nerozlišují mezi rámcem s uspořádáním a rámcem bez něj.

13.2 Aplikační postup

Postup má sedm kroků a řídí se pořadím podmíněnosti z tabulky 12-2, a nikoli pořadím kapitol části III. Datové objekty se nejprve vyčtou do registru se stabilními identifikátory, neboť každá proměnná se připojuje buď k

objektu, nebo k činnosti na něm prováděné. Podmiňující proměnné se poté přiřazují v pořadí: třída dat, autoritativní kopie a vrstva a strana odpovědná za výsledek u každé činnosti. Proměnné podmíněných dimenzí se čtou teprve potom, protože určení popisu nebo dlouhodobého uchování přijaté dříve, než je fixována třída, není určením, nýbrž preferencí.

Zbývající čtyři kroky produkují výstup. Stav každé proměnné se čte prostřednictvím jejího indikátoru, jenž specifikuje operaci, výběrovou jednotku a datum, namísto aby záznam pouze pojmenoval. Proměnné, jejichž stav číst nelze, se zaznamenávají s uvedením důvodu, přičemž se odlišuje nepřítomnost záznamu od nepřístupnosti záznamu, neboť obojí nese odlišné důsledky a pouze první z nich je zjištěním o daném prostředí. Konflikty se poté řeší pravidly 1–5 se zaznamenanou rozhodující rolí a zbytkovým rizikem, každé proměnné se přiřadí úroveň zralosti a uplatní se test uzavřenosti z oddílu 4.6, jenž potvrzuje, že každý konflikt skončil u pojmenované autority.

Ve třetím kroku jsou dostupné dva režimy čtení a rozdíl mezi nimi je předmětem následujícího oddílu. Při rekonstrukci se stavy čtou ze zdokumentovaného materiálu popisujícího prostředí daného typu. Při empirické aplikaci se stavy čtou z vlastních záznamů instituce v rámci sjednaného přístupu, s uvedeným výběrovým rámcem a s nezávislým kódováním. Postup je totožný; důkazní status jeho výstupu nikoli.

13.3 Rekonstrukce a empirická aplikace

Žádná aplikace tohoto rámce nebyla provedena proti záznamům pojmenované instituce. Čtyři prostředí z tabulky 13-1 jsou rekonstrukcemi vybudovanými ze zdokumentovaného materiálu a jako rekonstrukce se zde také uvádějí. Vyprávět je, jako by šlo o instituce, jejichž záznamy byly prozkoumány, by vyprodukovalo právě ten ilustrativní kazuistický důkaz, jež kapitola 4 identifikuje jako nejslabší vrstvu tohoto oboru, navíc s vadou vymyšlenosti. Kniha proto uvádí design, jenž by každou rekonstrukci převedl ve výsledek, namísto aby dodávala jeho zdání.

To, co rekonstrukce založit může, je omezené, nikoli však prázdné. Může ukázat, že každá proměnná rámce je v daném prostředí vyjádřitelná, že se postup uzavírá, že konflikty směřují k autoritě a že některé konflikty pravidla nerozhodnou. Jde o vlastnosti konstrukce pozorované za náročného popisu prostředí a níže se jako takové uvádějí. To, co rekonstrukce založit

nemůže, je rozložení stavů proměnných, rozlišovací síla kteréhokoli indikátoru, míra shody mezi posuzovateli a četnost, s níž pravidla ponechávají konflikt otevřený.

CHYBĚJÍCÍ DOKLAD

Empirická aplikace na třech až pěti výzkumných organizacích vybraných podle kritérií tabulky 13-1, zahrnující sjednaný přístup k projektovým záznamům, výběrový rámec datových činností stanovený posuzovatelem, a nikoli institucí, nezávislé kódování všech 55 proměnných dvěma posuzovateli, uvedený koeficient shody, rozložení úrovně zralosti a registr konfliktů s uplatněným pravidlem a se zaznamenaným zbytkovým rizikem — studie, jež by níže uvedené rekonstrukce převedla ve zjištění o institucích.

CHYBĚJÍCÍ DOKLAD

Aplikace rámce na probíhající projekt s více partnery po celou dobu jeho trvání, se stavy proměnných zaznamenanými ve třech bodech a s párovým srovnáním závazků plánu s realizovanými záznamy provedeným při ukončení projektu, umožňující stanovit, zda určení rámce zůstávají stabilní, mění-li se podmínky projektu.

13.4 Co rekonstrukce naznačují o úplnosti

Z rekonstrukcí plynou tři zjištění o úplnosti a každé z nich je vlastností souboru proměnných, a nikoli kteréhokoli prostředí. První se týká chybějícího objektu. Žádná z 55 proměnných nepojmenovává incident ani nápravné opatření, ačkoli ztráta důvěrnosti, integrity či dostupnosti, chybné zveřejnění, neodůvodněný výmaz a ztráta provenience jsou zdokumentovanými událostmi s identifikovatelnými vlastníky a záznamy (National Institute of Standards and Technology 2024). Škála zralosti z oddílu 12.4 se na úrovni 4 o takové události opírá, zatímco soubor proměnných nenese žádné místo, kam je zaznamenat, což je vnitřní nekonzistence, a nikoli opomenutí detailu.

Tuto mezeru lze zacelit uvnitř vlastního kritéria rámce pro připuštění dimenze. Řešení incidentů neadresuje žádný rozhodovací objekt, jež governance již neadresuje, neboť se týká toho, kdo rozhoduje, v jakém intervalu a s jakým záznamem, a náleží proto do D1 jako další proměnná, a nikoli jako osmá dimenze. Jejím doložitelným záznamem je registr incidentů nesoucí klasifikaci podle typu dopadu, posouzení případné oznamovací povinnosti, nápravná a preventivní opatření s vlastníky a daty a výsledek.

Rámec ve znění kapitoly 12 je právě v tomto ohledu neúplný a odkryla to rekonstrukce.

Druhé zjištění se týká zdánlivé duplicity. Tři proměnné adresují přenos odpovědnosti za výsledek pod různými názvy: přiřazení kontinuity v D1, ujednání o nástupnictví v D2 a ujednání o ukončení v D7. Rekonstrukce naznačují, že je lze odlišit podle strany, jejíž zánik je spouští — postupně osoby, projektu a služby —, a že se tomu odpovídajícím způsobem liší i záznamy. Podrzet je odděleně je tedy obhajitelné a rámec uvádí odlišující spouštěč, namísto aby všechny tři sléval v jedinou slabší proměnnou.

Třetí zjištění se týká nákladů. Pravidla 4 a 5 obě odkazují na proporcionalitu a proporcionalita předpokládá soud o tom, co je instituce schopna unést, přesto žádná proměnná rámce nenese kapacitu, financování ani jednotkové náklady. Korpus zachází s náklady jako s omezením rozhodnutí, a nikoli jako s vlastností, jež má záznam, a rámec toto zacházení dědí. Důsledek je přesný a přenáší se do kapitoly 14: rámec určuje, co je přiměřené riziku, a neurčuje, co je finančně únosné, takže určení, jež produkuje, může být zároveň správné a neproveditelné.

13.5 Interpretační nároky proměnných

Rekonstrukce rozdělují soubor proměnných do dvou skupin, jejichž rozdíl váží více než jakýkoli souhrnný výrok o použitelnosti. Jedna skupina se čte bez interpretace v tom smyslu, že dva posuzovatelé zkoumající týž záznam by vrátili tentýž stav: určená autoritativní kopie, režim přístupu nastavený u uložení, výsledek čistého spuštění, identifikátor a licence publikovaného objektu, datum poslední kontroly bitové neměnnosti. Proměnné této skupiny nesou posuzovací váhu rámce, protože jejich stavy lze zpochybnit jediné předložením jiného záznamu.

Druhá skupina vyžaduje soud, jež rámec strukturuje, aniž by jej odstraňoval. Granularita provenience přiměřená riziku, významné vlastnosti objektu, deklarovaný účel, vůči němuž se posuzuje vhodnost k použití, specifičnost položky plánu, přičtení omezení proměnné nezbytnosti i znalostní základna předpokládaná u určené komunity — u všech z nich předchází přiřazení stavu posouzení. Nejde o vadu, kterou by bylo možné odstranit návrhem, neboť každá z těchto proměnných kóduje rozhodnutí, jež může přijmout pouze způsobilá strana. Znamená to však, že úrovně zralosti jsou u druhé skupiny jisté jen do té míry, do jaké se posuzovatelé shodnou,

což je podmínka, s níž kapitola 14 zachází jako s předřazenou použitelností škály.

CHYBĚJÍCÍ DOKLAD

Nezávislé přiřazení stavů a úrovní zralosti u všech 55 proměnných alespoň třemi posuzovateli pracujícími z téhož projektového materiálu, se shodou vykázanou zvlášť pro obě výše identifikované skupiny, umožňující stanovit, které proměnné jsou použitelné v posuzování a které vyžadují rozhodčí určení.

13.6 Konflikty, jež rozhodovací pravidla nevyřešila

Byly zkonstruovány tři třídy konfliktu, v nichž pravidla 1–5 nevracejí žádné určení, a jejich existence je zjištěním o pravidlech, a nikoli o prostředích. V první z nich se setkávají dva požadavky, od nichž se nelze odchýlit: závazná retenční povinnost a žádost o výmaz adresovaná týměž záznamům. Pravidlo 1 odstraňuje varianty na obou stranách a množinu vyprazdňuje, takže rámec namísto rozhodnutí produkuje eskalaci se zdokumentovaným konfliktem. Právě takové chování test uzavřenosti požaduje a stojí za to říci bez obalu, že rámec, jenž by zde určení vydal, by si osoboval více, než mu náleží.

V druhé třídě jsou nevratná obě porušení. Zveřejnění datové sady o infrastruktuře nelze vzít zpět a lhůta zadržení přesahující užitečnou životnost měření ničí hodnotu, kvůli níž bylo měření pořízeno. Pravidlo 2 nerozlišuje, protože nevratnost je přítomna na obou stranách, pravidlo 3 rozlišuje pouze tam, kde se liší stupně důkazu, a pravidlo 4 odkazuje na třídu dat definovanou jen na jedné straně konfliktu. Výstupem rámce je pak pravidlo 5 s podmínkou přezkumu, jež zachovává množinu variant za cenu toho, že rozhodnutí zůstává předběžné.

Třetí třída má nejzávažnější důsledky a nejhůře se zvládá. Tam, kde se dostanou do konfliktu právní dovolení a komunitní protokol, se pravidlo 1 neuplatní, protože protokol není požadavkem, od něž by se v právním smyslu nebylo možné odchýlit, a pravidlo 3 staví právní nástroj nad něj podle stupně důkazu (Carroll et al. 2020). Určení, jež z toho plyne, je určením, které by podstatná část oboru odmítla, a odpovědí rámce je zaznamenat zbytkové riziko a roli, jež je přijímá, namísto skrývání hodnotového soudu uvnitř stupně. Tabulka 13-2 uvádí tři třídy a to, co rámec v každé z nich produkuje.

Tabulka 13-2. Třídy konfliktu, v nichž rozhodovací pravidla nevracejí žádné určení.

Třída	Proč pravidla nerozlišují	Výstup rámce
Dva požadavky, od nichž se nelze odchýlit	Pravidlo 1 vyprazdňuje množinu variant	Eskalace se zdokumentovaným konfliktem; žádné určení
Obě porušení nevratná	Pravidlo 2 neuplatnitelné; pravidlo 4 definované jen na jedné straně	Předběžná varianta podle pravidla 5, s podmínkou přezkumu
Právní dovození proti komunitnímu protokolu	Pravidlo 3 hodnotí protokol níže než právní nástroj	Určení se zaznamenaným zbytkovým rizikem a s pojmenovanou přijímající rolí

CHYBĚJÍCÍ DOKLAD

Registr autentických konfliktních případů shromážděných z výzkumných projektů, kódovaných podle třídy a podle pravidla, jež je vyřešilo, s uvedením podílu případů spadajících do tří tříd tabulky 13-2 — měření, na němž závisí jakékoli tvrzení o pokrytí pravidel.

Rekonstrukce tedy vrací smíšený výsledek, jenž se uvádí bez zjemňování. Rámec je vyjádřitelný napříč prostředími výrazně odlišné velikosti i oboru, jeho postup se uzavírá a jeho konflikty směřují k autoritě; je neúplný ve věci incidentů, mlčí o nákladech a je neurčitý pro charakterizovatelnou třídu konfliktů. Otázky, zda jeho určení obstojí, mění-li se samo prostředí, a které z jeho proměnných se přitom pohybují, se ujímá následující kapitola.

Kapitola 14. Robustnost, citlivost a hranice platnosti

Rámec, jenž produkuje určení, je odpovědný za stabilitu těchto určení za podmínek, které si ne zvolil. Tato kapitola uvedenou stabilitu posuzuje a obhajuje tvrzení o dvou částech: citlivost rámce na prostředí má vzorec, a nikoli difuzní podobu, a tento vzorec sleduje pořadí podmíněnosti ustavené v kapitole 12. Proměnné podmiňujících dimenzí si napříč prostředím podržují svou definici i své čtení, kdežto proměnné podmíněných dimenzí mění spolu s prostředím buď podobu svého záznamu, nebo svůj stav. Přenositelnost je tedy asymetrická způsobem, jež lze uvést přesně, a toto uvedení je výsledkem této studie, a nikoli výhradou k ní připojenou.

S hlavní výzkumnou otázkou to souvisí prostřednictvím šesté podřazené otázky z kapitoly 1, která se ptá, jak daleko se přenáší rámec odvozený z jednoho národního a institucionálního prostředí a co značí hranici, za níž se již nepřenáší. Odpověď se zde podává na úrovni proměnných, a nikoli na úrovni rámce jako celku, neboť rámec se nepřenáší ani nepřestává přenášet vcelku. Následující výklad identifikuje, které komponenty se pohybují, které drží a za jakých podmínek určení slábnou až do bodu, kdy přestávají být informativní.

14.1 Čtyři zdroje citlivosti

Oborová konvence mění podobu záznamu, aniž by měnila požadavek, jež tento záznam naplňuje. V přístrojově vybavených oborech nesou odvození logy workflow a konfigurační soubory, v oborech interpretativních vědecký aparát, kódovací schémata a analytické poznámky; obojí naplňuje proměnnou odvození a ani jedno druhé nenahrazuje. Za této podmínky se posouvá profil popisu, granularita, na níž je provenience přiměřená, a podoba důkazu o reprodukovatelnosti. Drží přiřazení odpovědnosti za výsledek, eskalační práh, klasifikace materiálu a rozhodnutí o době uchování.

Právní režim působí jinak, protože zasahuje přípustnost, a tedy množinu variant, a nikoli podobu důkazu. Určení dimenze ochrany jsou konstrukčně relativní vůči jurisdikci a rámec lokalizuje třídu rozhodnutí a autoritu příslušnou je přijmout, namísto aby dodával odpověď (European Parliament and Council 2016). Institucionální velikost působí na umístění způsoblosti:

malé instituce naplňují proměnné governance a ochrany prostřednictvím pojmenované vnější způsobilosti a jejich stavy se pak čtou ze smluv, a nikoli z organogramů. Dostupná infrastruktura působí na dosažitelné úrovni, neboť instituce bez repozitáře, bez řízeného prostředí nebo bez služby dlouhodobého uchování nemůže několik proměnných zvednout nad úroveň 1, ať jsou její záměry jakékoli. Tabulka 14-1 uvádí, co se za každé z těchto podmínek posouvá a co drží.

Tabulka 14-1. Zdroje citlivosti, s tím, co se u každého z nich posouvá a co drží.

Podmínka	Co se posouvá	Co drží	Účinek na určení
Oborová konvence	Podoba záznamu; profil; granularita provenience	Přiřazení, eskalace, klasifikace, doba uchování	mění se podoba důkazu; požadavek beze změny
Právní režim	Přípustnost; právní titul; riziko identifikace; retenční lhůta	Struktura inventáře práv; přičtení omezení	mění se množina variant; třída rozhodnutí beze změny
Institucionální velikost	Umístění způsobilosti; vlastní, nebo smluvní	Existence pojmenované strany odpovědné za výsledek	důkaz se přesouvá do smluv
Dostupná infrastruktura	Dosažitelná úroveň zralosti	Definice každé proměnné	strop na úrovních, nikoli na definicích

Poslední řádek nese rozlišení, jež nástroje zralosti v tomto oboru běžně stírají. Instituce, jejíž infrastruktura činí úroveň 3 u proměnné dlouhodobého uchování nedosažitelnou, je v jiném postavení než instituce, která infrastrukturu má a nevyužívá ji, a škála vykazující pro obě jediné číslo je popisuje shodně. Rámec obojí odděluje tím, že zaznamenává důvod, proč stav nebylo možné přecházet, jak specifikuje aplikační postup v oddílu 13.2. Toto oddělení je důvodem, proč se infrastrukturní strop uvádí jako zjištění o prostředí, a nikoli jako nedostatek instituce.

14.2 Které proměnné jsou stabilní a které závisejí na prostředí

Rekonstrukce z kapitoly 13 podpírají rozdělení souboru proměnných do tří tříd a toto rozdělení je analytickým soudem této studie, a nikoli změřenou

vlastností. Invariantní proměnné si podržují svou definici, druh svého záznamu i své čtení napříč všemi čtyřmi podmínkami a náleží k nim přiřazení odpovědnosti za výsledek, průchodnost eskalace, autoritativní kopie, existence záznamu o odvození, stopa revizí a přičtení omezení proměnné nezbytnosti. Formově variantní proměnné podržují požadavek, zatímco záznam se s prostředím mění, a náleží k nim aplikační profil, vazba slovníků, protokol manuálního kroku, specifikace prostředí, významné vlastnosti a výrok o nejistotě. Proměnné určené prostředím mají stavy fixované převážně prostředím, a nikoli případem, a náleží k nim režim přístupu, cíle obnovy, posouzení repozitáře, retenční lhůta, interoperabilní minimum a cesta uznání (Úřad vlády České republiky 2025).

Věcným zjištěním je rozložení těchto tříd napříč dimenzemi. Invariantní proměnné se soustřeďují v governance a v životním cyklu a provenienci, tedy ve dvou dimenzích, jež podmiňují všechny ostatní; proměnné určené prostředím se soustřeďují v ochraně a v dlouhodobém uchování, publikování a uznání, tedy v dimenzích nejbližších právnímu režimu a infrastruktuře. Rámec je tak nejstabilnější právě tam, kde nese autoritu a posouditelnost, a nejméně stabilní tam, kde nese odpovědi. Právě to činí rozhodovací strukturu exportovatelnou, zatímco její výstupy zůstávají místní.

Alternativní vysvětlení tohoto vzorce si zaslouží uvedení, neboť by výsledek podstatně oslabilo. Podmiňující proměnné mohou být invariantní proto, že jsou abstraktní, a nikoli proto, že jsou robustní, neboť požadavku pojmenovat stranu odpovědnou za výsledek lze vyhovět v jakémkoli prostředí právě tím, že se o činnosti této strany řekne jen málo. Při tomto čtení by stabilní jádro rámce bylo stabilní tak, jak je stabilní truismus, a jeho zdánlivá přenositelnost by byla vykoupena prázdnotou. Obě čtení lze empiricky odlišit, protože abstraktní proměnná nabývá všude téhož stavu, kdežto robustní proměnná rozlišuje mezi prostředími, jež se od sebe liší.

CHYBĚJÍCÍ DOKLAD

Nezávislé zařazení všech 55 proměnných do tří výše uvedených tříd dvěma analytiky pracujícími pouze z definic, s uvedeným koeficientem shody, spolu s rozložením stavů invariantních proměnných napříč prostředími — pozorování, jež odděluje robustnost od abstraktnosti.

14.3 Podmínky, za nichž určení slábnou nebo selhávají

Určení rámce oslabují čtyři podmínky a první z nich je absolutní. Tam, kde datová činnost nezanechává vůbec žádný záznam, stojí každá proměnná k ní připojená na úrovni 0 a rámec vrací konstatování nečitelnosti namísto posouzení. Nemůže diagnostikovat to, co nedokáže přečíst, a nespravované prostředí je pro něj proto neviditelné stejně, jako je neviditelné pro audit. Rámec posuzuje spravovaná prostředí a jeho mlčení o zbytku je hranicí nástroje, a nikoli zjištěním o praxi, kterou nevidí.

Druhou podmínkou je kustodie mimo posuzovanou instituci. Ve federovaných infrastrukturách, v komerčních službách a v důvěryhodných výzkumných prostředích leží doložitelný záznam za hranicí, kterou vnější posouzení nepřekročí, a proměnná je pak naplněna atestací namísto prozkoumání (CoreTrustSeal 2022). Stupeň důkazu tomu odpovídajícím způsobem klesá z přímého pozorování na ujištění třetí strany a rámec tuto záměnu zaznamenává, namísto aby s obojím zacházel jako s rovnocenným. Tam, kde není dostupné ani prozkoumání, ani atestace, je poctivým výstupem nepřechtená proměnná, a nikoli příznivá výchozí hodnota.

Třetí podmínkou je změna normativní a infrastrukturní vrstvy. Proměnné byly formulovány tak, aby takovou změnu přežily, indikátory však odkazují na konkrétní artefakty — certifikační schémata, metadatové registry, identifikátorové služby —, jejichž aktuálnost se rozpadá, takže záznam o výběru citující certifikaci bez data ověření uvádí méně, než se jeví. Čtvrtou podmínkou je právní neurčitost: tam, kde je použitelné právo sporné nebo nově účinné, nelze přijmout určení přípustnosti podle pravidla 1 a uspořádání se zastavuje hned v prvním kroku. Výstupem rámce je pak eskalace, což je správné chování, a nikoli určení.

14.4 Shoda hodnotitelů jako podmínka použitelnosti

Škála zralosti z oddílu 12.4 je použitelná jedinečně za podmínky, jež je předržena veškerému jejímu obsahu. Přiřadí-li dva posuzovatelé zkoumající též materiál téže proměnné odlišné úrovně, není výsledný profil měřením a srovnáním mezi institucemi ani v čase pak nemají žádnou interpretaci. Shoda tedy není žádoucí vlastností škály, nýbrž podmínkou, za níž škála vůbec existuje jako nástroj. Tato studie netvrdí, že je uvedená podmínka naplněna, protože měření, jež by to založilo, provedeno nebylo.

Rekonstrukce nicméně naznačují, kde se neshoda soustřeďuje, a tento vzorec je informativní. Na hranicích mezi úrovněmi 0, 1 a 2 se očekává shoda vysoká, neboť pozorovatelnými skutečnostmi na těchto prazích jsou existence dokumentu a přítomnost pojmenované role a data. Na hranicích mezi úrovněmi 2, 3 a 4, kde vstupuje soud o dostatečnosti záznamu, se očekává nižší a u interpretativních proměnných identifikovaných v oddílu 13.5 nejnižší. Škála vykazující jediný údaj o shodě napříč všemi proměnnými by právě tuto strukturu zakryla.

CHYBĚJÍCÍ DOKLAD

Studie shody s alespoň třemi posuzovateli nezávisle přiřazujícími úroveň zralosti všem 55 proměnným u téhož projektového materiálu, s vykázanou shodou pro každou proměnnou a pro každou hranici mezi úrovněmi proti prahu stanovenému předem, přičemž proměnné pod tímto prahem se uvádějí jako rozhodčí určení, a nikoli jako posouzení.

14.5 Riziko vykazování místo konání

Každý nástroj, jenž určuje zdroje, přetváří chování, které měří, a nástroj zralosti se obchází tím, že vzniká záznam namísto praxe — tedy vykazováním místo konání. Definice úrovní z kapitoly 12 byly zkonstruovány proti tomuto mechanismu: úrovně 3 a 4 jsou dosažitelné pouze prostřednictvím záznamů vzniklých jako vedlejší produkt činnosti a prostřednictvím událostí, jejichž načasování si posuzovaná strana nezvolila. Návrh zvyšuje náklady předstírání tak, aby se blížily nákladům samotné práce, což je vlastnost návrhu, a nikoli důkaz, a tři zbytkové zranitelnosti přetrvávají.

První je výběrový rámec. Tam, kde zkoumané činnosti vybírá instituce, lze záznamy vzniklé jako vedlejší produkt předem kurátorsky upravit, a odpovědi dostupnou rámci je výběrový rámec stanovený posuzovatelem a odvozený z registru objektů, jež instituce vede pro jiné účely. Druhou je inscenovaná událost: test obnovy provedený na malé, nedávno zapsané datové sadě vyhovuje liteře úrovně 4, aniž by cokoli testoval, a právě proto indikátor specifikuje výběrovou jednotku, a nikoli pouze operaci. Třetí je registr vedený kvůli posouzení: registr incidentů bez jediné položky je nerozlišitelný od prostředí bez incidentů, není-li spárován s nezávislými stopami, jakými jsou logy přístupů a požadavky na podporu.

Vůči těmto strategiím není imunní žádný nástroj zralosti a tvrzení o imunitě by bylo nejzřetelnějším dokladem toho, že nástroj nikdy nebyl

použit v prostředí, kde na výsledku něco záviselo. Tvrzení rámce je užší a testovatelné. Zní tak, že rozdíl mezi profilem ze sebesposouzení a profilem ověřeným posuzovatelem je měřitelnou veličinou, že se soustřeďuje v interpretativních proměnných a že roste s důsledky, jež jsou k výsledku připojeny.

CHYBĚJÍCÍ DOKLAD

Párové sebesposouzení a ověření posuzovatelem u týchž proměnných v týchž organizacích, s uvedením směru a velikosti rozdílu podle tříd proměnných, umožňující stanovit, nakolik lze na vykázanou zralost spoléhat a kde je nezávislé ověření nepostradatelné.

14.6 Co rámec neurčuje

Meze rámce plynou z vlastnosti, jež jej činí posouditelným, a jejich uvedení je součástí výsledku, a nikoli jeho zúžením. Rámec, jehož proměnné se čtou ze záznamů, může určit pouze to, co záznamy unesou. Právní závěry stojí mimo něj, neboť lokalizuje třídu rozhodnutí a příslušnou autoritu, aniž by dodával odpověď. Vědecká významnost souboru dat stojí mimo něj rovněž, protože výběr k uchování vyžaduje oborový soud, jež nedodá žádný záznam a jež rámec směřuje k pojmenované roli, namísto aby jej rozhodoval.

Za výslovné uvedení stojí čtyři další vyloučení. Rámec neurčuje, co je finančně únosné, jak ustavuje oddíl 13.4, takže určení, jež produkuje, může být zároveň správné a neproveditelné. Neurčuje, zda bylo rozhodnutí správné, nýbrž pouze zda je přijala způsobilá strana na uvedeném základě a se zaznamenaným zbytkovým rizikem, což je vlastnost slabší a obhajitelnější. Neurčuje ani volbu nástrojů, ani organizační formu, s nimiž kapitola 5 zachází jako s kontextovými podmínkami, a nikoli jako s proměnnými. A neurčuje etickou přijatelnost výzkumného účelu, o níž rozhodují orgány, jejichž autoritu rámec předpokládá.

14.7 Přenositelnost uvedená jako zjištění

Přenositelnost tohoto rámce je asymetrická a tato asymetrie je jeho vlastností s nejzávažnějšími důsledky. Rozhodovací struktura — proměnné podmiňujících dimenzí, uspořádání, pravidla a požadavek, aby bylo zbytkové riziko zaznamenáno — se přenáší napříč oborovou konvencí, právním režimem, velikostí i infrastrukturou, protože proměnné, jež ji nesou, jsou vůči všem čtyřem invariantní. Určení se nepřenáší, protože je produkují

podmíněné proměnné, jejichž stavy fixuje převážně prostředí. Rámec tohoto druhu lze proto exportovat jako strukturu a nelze jej exportovat jako soubor odpovědí.

Tento závěr má přímý důsledek pro to, jak se práce tohoto druhu publikuje. Publikovat určení jako doporučení by reprodukovalo selhání diagnostikované v kapitole 1, při němž se nástroje navržené pro jedno prostředí uplatňují v jiném, jako by jejich závěry putovaly s nimi. Přiměřeným výstupem je struktura spolu s podmínkami, za nichž každé určení platí, což je to, co uvádí část III a co ohraničuje tato kapitola. Zbývá otázka, zda konvergence institucionální praxe, o níž korpus referuje, připouští jiná vysvětlení než to, jež podala tato kniha, a té se ujímá část V.

Č Á S T V

Kritická diskuse

Kapitola 15. Kritická diskuse

Část IV se uzavírá rámcem, jenž je vyjádřitelný napříč prostředími výrazně odlišné velikosti i oboru, jehož postup se uzavírá a jehož přenositelnost je asymetrická. Žádná z těchto vlastností nerozhoduje o tom, zda rámec stojí za to mít. Tvrzení obhajované v této kapitole má dvě části: rámec zaujímá pozici, kterou nezaujímá žádná ze sousedních rodin nástrojů, a alternativní vysvětlení schopná tuto pozici rozpustit přezkoumání přžívají v oslabené, avšak neprázdné podobě. To, co nakonec zbývá, je užší než to, co ohlásila kapitola 1, a zúžení se uvádí, namísto aby bylo pohlceno souhrnem.

S hlavní výzkumnou otázkou to souvisí spíše statusem než obsahem. Kapitola 1 se ptá, za jakých podmínek lze částečné požadavky integrovat do rámce přinášejícího obhajitelná rozhodnutí, a části III a IV dodávají konstrukci spolu s jejími hranicemi. Zda tato konstrukce na otázku odpovídá, závisí na tom, zda integrace byla dosažena, a nikoli pouze simulována. Věc se níže rozhoduje s přesně uvedenými výhradami, neboť výhrada vyslovená jako obecné varování je nerozlišitelná od výhrady žádné.

15.1 Pozice mezi nástroji těžé rodiny

S tradicí modelů zralosti způsobilostí sdílí rámec stupňovitou strukturu a institucionální profil jako svůj výstup (Paulk et al. 1993; Crowston and Qin 2011; Cox et al. 2019). Odchyluje se tam, kde tato tradice umísťuje své úrovně. Úrovně se zde připojují k proměnným, a nikoli k institucím, a institucionální zralost je rozložením napříč 55 proměnnými, a nikoli stupněm (oddíl 12.4). Každou úroveň odděluje od úrovně nižší pozorovatelná skutečnost, kterou posuzovatel může hledat a nenalézt, což odstraňuje aspirační jazyk, jímž se popisům stupňů běžně vyhovuje.

S nástroji posuzování FAIR sdílí rámec ambici čtení, jež nezávisí na tom, co o sobě uvádí producent (Wilkinson et al. 2016; RDA FAIR Data Maturity Model Working Group 2020). Odchylka spočívá v posuzovaném objektu. Takové nástroje vypovídají svá skóre o digitálním objektu a mlčí o tom, kdo rozhodl o podobě, jíž tento objekt nabývá, kdežto každá proměnná zde je vlastností datové činnosti a záznamu, jež tato činnost zanechala, takže datová sada může skórovat vysoko, ačkoli rozhodnutí, které ji vyprodukovalo, nepřijal nikdo identifikovatelný.

Nejbližším metodickým příbuzným je certifikace repozitářů, neboť posuzuje shodu kustoda s dokumentovanými požadavky prostřednictvím záznamů, a nikoli tvrzení (CoreTrustSeal 2022; Lin et al. 2020). Táž důkazní logika se zde uplatňuje na odlišnou jednotku: posuzování shody adresuje službu, která data drží na konci životního cyklu, kdežto rámec adresuje instituci a projekt, v nichž se přijímají rozhodnutí určující, co se k této službě vůbec dostane.

Reforma hodnocení výzkumu sdílí diagnózu, že počítání výstupů je necitlivé k tomu, co počítané objekty obsahují (CoARA 2022; DORA 2013). Reforma zachází s uznáním práce s daty jako s cílem, kdežto rámec zachází s uznáním jako s proměnnou podmiňující každou dimenzi prostřednictvím zpětné hrany z oddílu 12.2, a to na tom základě, že úsilí neviditelné pro institucionální hodnocení se neudrží. Reforma hodnocení je při tomto čtení předpokladem politiky správy výzkumných dat, a nikoli paralelní agendou, což je pozice falzifikovatelnější, než jakou reformní literatura obvykle zaujímá. Tabulka 15-1 uvádí čtyři vztahy a náklady v jejím čtvrtém sloupci nejsou vedlejší: nástroj, jenž nevydává jediné číslo, nemůže sloužit srovnávacím účelům, kvůli nimž instituce posouzení nejčastěji zadávají.

Tabulka 15-1. Vztah rámce ke čtyřem sousedním rodinám nástrojů, s náklady každé odchylky.

Rodina	Sdílené	Odchylka	Náklady odchylky
Modely způsobilosti a zralosti	Stupňovitý vývoj; institucionální profil	Úrovně se připojují k proměnným a každý práh pojmenovává pozorovatelnou skutečnost	Žádné jediné skóre, a tedy žádné pořadí
Nástroje posuzování FAIR	Čtení nezávislé na producentovi	Jednotkou je datová činnost a její záznam	Neautomatizovatelné; nákladnější než strojově vytěžené skóre
Certifikace repozitářů	Posouzení kustoda na základě záznamů	Jednotkou je rozhodující instituce	Žádný akreditační orgán a žádné vnější potvrzení
Reforma hodnocení výzkumu	Počítání je necitlivé k obsahu	Uznání je podmiňující proměnnou	Kritéria se stávají závislými na politice, kterou instituce nemusí ovládat

15.2 Fragmentace jako přechodný stav

Nejsilnější alternativa k diagnóze z kapitoly 1 tvrdí, že fragmentace tam popsaná je fází, a nikoli strukturou. Při tomto čtení nejsou nástroje propojeny proto, že obor je mladý, a infrastrukturní konvergence — federované katalogizování, společné metadatové modely, národní adresáře — dodá spojovací vrstvu, kterou kniha dodává analyticky (EOSC Association 2024; EOSC CZ 2025). Institucionální praxe by pak konvergovala i bez jakéhokoli rámce a rozhodovací struktura formulovaná nyní by dokumentovala přechodný stav. Toto čtení si zaslouží, aby bylo bráno vážně, neboť období 2024–2026 pokryté korpusem je obdobím, v němž taková konvergence viditelně probíhá.

Toto čtení je nicméně neúplné způsobem, jež lze přesně lokalizovat. Konvergence řeší konflikty formátu, slovníku a výměny, neboť jde o koordinační problémy, z jejichž řešení má prospěch každá strana. Neřeší konflikty z oddílu 4.5 instanciované v oddílu 13.6, protože ty vznikají mezi požadavky, z nichž je každý uplatněn správně a jejichž naplnění se vzájemně vylučuje. Společný metadatový model neurčuje, zda převáží retenční povinnost, nebo žádost o výmaz, ani zda zpřístupnění upravuje právní dovolení, nebo komunitní protokol.

Konvergence může druhou třídu konfliktů ve skutečnosti vyostřit, neboť tam, kde je uložení snadné a katalogizování federované, se rozhodnutí o zpřístupnění, o výběru k uchování a o řízeném přístupu přijímají častěji a rychleji. Mechanismus je tentýž, jaký byl pro automatizaci identifikován v oddílu 12.5: přesun činnosti do infrastruktury mění náklady záznamu, a nikoli umístění autority. Přechodné čtení se proto přijímá pro vrstvu interoperability a odmítá pro vrstvu rozhodování.

CHYBĚJÍCÍ DOKLAD

Opakované měření deseti oblastí rozhodování z tabulky 1-1 na pevné populaci institucí po dobu alespoň pěti let, zaznamenávající, zda podíl institucí s pojmenovaným vlastníkem rozhodnutí roste s infrastrukturní konvergencí — pozorování podírající přechodné čtení proti čtení strukturnímu.

15.3 Integrace abstrakcí

Oddíl 14.2 vznáší protihypotézu, jež by knihu připravila o její hlavní tvrzení, a zde se jí čelí, namísto aby byla ponechána tam, kde byla vyslovena. Stabilní

jádro může být stabilní proto, že je abstraktní, neboť požadavku pojmenovat stranu odpovědnou za výsledek lze vyhovět kdekoli právě tím, že se o činnosti této strany neřekne nic. Integrace napříč sedmi dimenzemi by pak byla artefaktem toho, že každá z nich je popsána na úrovni natolik obecné, aby rozdíly zmizely. Tuto hypotézu nelze zodpovědět prozkoumáním, protože abstraktní a robustní proměnná vypadají ve svých definicích shodně.

Dvě úvahy hovoří proti ní, aniž by ji vyvracely. První zní, že několik podmiňujících proměnných je definováno pozorovatelnou skutečností, v níž by podstatný podíl prostředí neobstál: průchodnost eskalace vyžaduje log se zaznamenaným výsledkem a intervalem a autoritativní kopie určení v registru, jež zavazuje, rozejdou-li se kopie navzájem. Proměnná není prázdná, chybí-li záznam, jež pojmenovává, často, a vlastní výklad rámce předpovídá, že tyto záznamy chybějí častěji, než jsou přítomny. Druhá zní, že pravidla produkují určení, jež některé strany odmítají, jak oddíl 13.6 dokládá u konfliktu mezi právním dovolením a komunitním protokolem, kdežto prázdná struktura neprodukuje žádné určení, proti němuž by kdokoli namítal.

Obě úvahy jsou slabší, než se jeví, neboť každá se opírá o vlastnost konstrukce, a nikoli o pozorování. Rozhodujícím testem je ten, jež pojmenovává oddíl 14.2: abstraktní proměnná nabývá v každém prostředí téhož stavu, kdežto robustní proměnná odděluje prostředí, jež se podle nezávislých důkazů liší. Dokud nebude rozložení stavů invariantních proměnných změřeno, zůstává protihypotéza živá a tvrzení o integraci je v této míře ponecháno otevřené.

CHYBĚJÍCÍ DOKLAD

Měření stavů invariantních proměnných z oddílu 14.2 na alespoň 20 institucích lišících se v nezávislých indikátorech výkonu ve správě výzkumných dat, s uvedením, zda tyto proměnné instituce od sebe oddělují, nebo zda všude nabývají téhož stavu — test odlišující robustnost od prázdnoty.

15.4 Zda rámec formalizuje stávající autoritu

Rozhodovací rámec, jenž přiděluje autoritu, může rozhodnutí zlepšit, anebo pouze zaznamenat, kdo rozhoduje již nyní, a druhá možnost není vzdálená. Každá datová činnost se přiřazuje jediné straně odpovědné za výsledek a každý konflikt směřuje k pojmenovanému příjemci, přičemž ve většině

institucí strany dostupné pro takové přiřazení již institucionální moc drží. Formalizace pak převádí neformální rozdělení na rozdělení zdokumentované a přidává legitimitu, aniž by cokoli přerozdělovala. Námitka je oprávněná a rámec žádný mechanismus přerozdělení neobsahuje.

Dodává namísto toho asymetrickou viditelnost. Rozhodnutí zaznamenané se svým základem, se svým zbytkovým rizikem a s rolí, jež toto riziko přijímá, je zpochybnitelné způsobem, jímž nezaznamenané rozhodnutí zpochybnitelné není, neboť jeho zpochybnění vyžaduje jiný záznam, a nikoli přístup k příležitosti, při níž bylo rozhodnutí přijato. Týmž směrem působí oddělení poradní funkce od rozhodovací, jež brání tomu, aby orientační posouzení fungovalo jako schválení, k jehož vydání se nikdo nepřihlásil. Ani jedna z obou vlastností autoritu nepřerozděluje a obě činí její výkon přičitatelným.

Nejostřejší podoba námítky se týká hodnotových soudů zabudovaných do samotného uspořádání pravidel. Pravidlo 3 staví právní nástroj nad komunitní protokol podle stupně důkazu a toto uspořádání kóduje postoj ke sporné otázce, a nikoli neutrální proceduru (Carroll et al. 2020). Dostupnou odpověď je zaznamenat zbytkové riziko a přijímající roli, namísto aby se uspořádání předkládalo jako technické, což činí soud viditelným a lokalizuje stranu, jež je za něj odpovědná. Viditelnost je slabší nápravou než neutralita a jedinou, jaká je dostupná struktura, která končí určením.

15.5 Zda nástroj slouží spíše auditu než praxi

Nástroje, jež čtou záznamy, jsou pro auditní funkce přitažlivé a vlastnosti posuzování rozvinuté zde činí rámec použitelným pro interního auditora, jenž nemá ve správě výzkumných dat žádnou způsobilost. Riziko je skutečné, nikoli hypotetické, protože struktura pobídek z oddílu 14.5 dopadá na tento rámec stejně jako na nástroje, jež kritizuje. Instituce, která by jej přijala pod tlakem auditu, by racionálně vyprodukovala registry, jež proměnné pojmenovávají, a praxi, kterou tyto registry stopují, by ponechala beze změny. Výsledek by na vyšší úrovni propracovanosti reprodukoval selhání spočívající v počítání plánů, diagnostikované v oddílu 7.7.

Návrh obsahuje jednu vlastnost, jež tomuto výsledku vzdoruje, a jednu, jež jej zve. Vzdorování plyne z definic úrovní, neboť úrovně 3 a 4 jsou dosažitelné pouze prostřednictvím záznamů vzniklých jako vedlejší produkt a prostřednictvím událostí, jejichž načasování si posuzovaná strana

nezvolila. Pozvání plyne z širší souboru proměnných, protože 55 proměnných napříč sedmi dimenzemi představuje zátěž, jíž se instituce nejlevněji zbaví tím, že bude s posouzením zacházet jako s dokumentačním cvičením. Která z obou tendencí převáží, závisí na tom, zda výběrový rámec stanovuje posuzovatel, což je podmínka užití, a nikoli vlastnost rámce, a tvrzení, že rámec zlepšuje rozhodnutí, tak jako tak dostupné není.

CHYBĚJÍCÍ DOKLAD

Srovnání záznamů o rozhodnutích a objemu dokumentace na institucích přijímajících rámec pro podporu rozhodování s institucemi přijímajícími jej pro interní audit, umožňující stanovit, zda způsob přijetí rozhoduje o tom, zda se mění praxe, nebo dokumentace.

15.6 Prokázaná, podepřená a hypotetická tvrzení

Odlišení převzatých faktů, syntéz a původních výsledků, udržované v celé knize, samo o sobě čtenáři neříká, které výsledky jsou jisté. Uvádí se zde proto druhé odlišení, jež to první protíná. Prokázaná tvrzení plynou z prozkoumání konstrukce nebo z korpusu v podobě, v jaké byl zmapován; podepřená tvrzení se opírají o rekonstrukci nebo o argument, a nikoli o měření; hypotetická tvrzení jsou předpovědi, jež rámec generuje a jež v knize nic netestuje. Tabulka 15-2 přiřazuje hlavní tvrzení ke třem třídám.

Tabulka 15-2. Hlavní tvrzení knihy, s důkazní třídou každého z nich a s tím, co by je posunulo do třídy silnější.

Tvrzení	Třída	Co by je posílilo
Deset oblastí rozhodování z tabulky 1-1 je pokryto částečně a nepropojeně	Prokázané pro zmapovaný korpus	Audit institucionálních politik
55 proměnných obstojí v testech přiřaditelnosti, doložitelnosti a uzavřenosti	Prokázané prozkoumáním	Nic; jde o vlastnost odvozenou
Pořadí podmíněnosti z tabulky 12-2 je asymetrické	Prokázané prozkoumáním	Měření kovariace napříč projekty
Pravidla 1–5 ponechávají tři třídy konfliktu z tabulky 13-2 otevřené	Prokázané konstrukcí	Registr autentických konfliktů
Proměnné jsou vyjádřitelné napříč čtyřmi obměňovanými podmínkami	Podepřené rekonstrukcí	Empirická aplikace na třech až pěti organizacích

Tvrzení	Třída	Co by je posílilo
Přenositelnost je asymetrická mezi podmiňujícími a podmíněnými dimenzemi	Podepřené analytickým soudem	Nezávislé zařazení s uvedeným koeficientem shody
Prahy mezi úrovněmi 2, 3 a 4 oddělují instituce, jež se od sebe liší	Podepřené argumentem z návrhu	Posouzení proti nezávislým indikátorům výkonu
Uznání podmiňuje, zda se úsilí ve správě výzkumných dat udrží	Hypotetické	Měření před a po vůči srovnávací skupině
Užití rámce mění rozhodnutí, a nikoli dokumentaci	Hypotetické	Longitudinální studie přijetí

Rozdíl mezi druhou a třetí třídou se při čtení ztrácí nejsnáze. Podepřené tvrzení přežilo styk s materiálem, jež rámec sám nevytvořil, a v tomto styku mohlo neobstát, kdežto hypotetické tvrzení je odvozeno z vlastní struktury rámce a nesetkalo se s ničím. Tvrzení druhého druhu opravňují k hedgingu *naznačuje*, tvrzení třetího druhu *k nasvědčuje tomu*, že, a hedging v částech III a IV byl přidělen na tomto základě.

15.7 Status důkazní základny

Rámec odvozený ze zmapovaného korpusu a posouzený proti rekonstruovaným prostředím podpírá slabší tvrzení než rámec validovaný proti institucionálním záznamům a rozdíl spočívá v tom, co se mohlo pokazit. Korpus byl sestaven účelovým výběrem, a nikoli podle zdokumentovaného vyhledávacího protokolu, a kodoval jej jediný analytik, takže není doložena ani jeho úplnost, ani reprodukovatelnost hranic mezi jeho dimenzemi (oddíly 4.3 a 4.4). Rekonstrukce nemůže rámci odporovat tak, jak to dokáže instituce, neboť materiál, z něhož je vybudována, je materiálem, z něhož byl rámec odvozen. To, o čem kniha referuje, je tedy studie návrhu s výsledkem o vnitřní konzistenci a se specifikovaným validačním programem, a nikoli validovaný nástroj.

Tento status určuje podobu každého tvrzení, k němuž je kniha oprávněna. Opravňuje k tvrzením o tom, co rámec určuje, co ponechává neurčené a které proměnné jsou čitelné bez interpretace, neboť jde o vlastnosti konstrukce. Neopravňuje k tvrzením o rozložení institucionální praxe, o spolehlivosti kteréhokoli posouzení ani o účinku přijetí, neboť to vše jsou

měření. Značky rozmístěné v textu tuto hranici registrují a jejich počet odráží důkazní základnu oboru stejnou měrou jako design této studie.

Pro užití výsledku z toho plyne jeden důsledek, a je to důsledek, nikoli varování. Rámec v tomto důkazním postavení lze aplikovat, zpochybňovat a testovat a nelze jej citovat jako opodstatnění. Instituce, jež by se jej dovolávala k osvědčení vlastního uspořádání, by zacházela s konstrukcí jako s autoritou, což je omyl, jež oddíl 3.6 identifikuje u užívání principiálních rámců jako nástrojů souladu.

Co z těchto výsledků plyne pro instituce, pro role držící rozhodnutí, pro infrastruktury a pro systémy, jež tuto práci uznávají, je předmětem následující kapitoly, v níž je každá implikace odvozena z uvedeného výsledku a označena tam, kde sám tento výsledek dosud prokázán není.

Č Á S T V I

Implikace

Kapitola 16. Implikace pro instituce, data stewardy a infrastruktury

Část V ustavuje, čím rámec je, co určuje a na čem jeho tvrzení v současnosti spočívají. Tato kapitola odvozuje, co z těchto výsledků plyne pro strany, jejichž rozhodnutí rámec popisuje, a odvození se v celém svém průběhu podržuje jediné kázni. Každá implikace níže se uvádí spolu s výsledkem, z něž plyne, a náleží-li tento výsledek k podepřené nebo hypotetické třídě z tabulky 15-2, přejímá implikace tento status a je odpovídajícím způsobem označena. Implikace, jejíž podpírající výsledek prokázán nebyl, je předpovědí o tom, co by plynulo, kdyby výsledek platil, a předložit ji jinak by studii návrhu proměnilo ve vodítko.

S hlavní výzkumnou otázkou to souvisí jako závěrečný krok argumentu, a nikoli jako jeho dodatek. Kapitola 1 klade otázku integrace v pojmech rozhodnutí obhajitelných na institucionální i projektové úrovni a rámec, jehož důsledky pro tyto úrovně zůstávají neuvedeny, ponechává otázku zodpovězenou zpola. Následující výklad proto identifikuje body, v nichž struktura rámce — a nikoli jeho určení — mění to, co instituce, role nebo služba dělá. Určení se nepřenáší, jak ustavuje oddíl 14.7, a žádné určení se zde nenabízí.

16.1 Co z pořadí podmíněnosti plyne pro výzkumné instituce

Pořadí podmíněnosti z tabulky 12-2 umísťuje governance bez předchůdce a test uzavřenosti z oddílu 4.6 je naplněn už jen z tohoto důvodu. Plyne z toho důsledek pro posloupnost, v níž institucionální úsilí nabývá účinku, a nikoli pro posloupnost, v níž se obvykle vynakládá. Úsilí vložené do podmíněných dimenzí dříve, než mají dimenze podmiňující určité stavy, produkuje proměnné, jejichž stavy nelze číst, neboť profil popisu, uspořádání dlouhodobého uchování nebo požadavek na reprodukovatelnost přijatý dříve, než jsou fixovány třída dat, autoritativní kopie a strana odpovědná za výsledek, je preferencí, a nikoli určením. Pořadí je vlastností závislostní struktury rámce a platí nezávisle na tom, zda se kovariace předpovězená v oddílu 12.2 potvrdí.

Instance s nejzávažnějšími důsledky se týká nástrojů. Kapitola 1 uvádí, že nabídka nástrojů a služeb pro správu výzkumných dat je nasycená, a nikoli vzácná, a oddíl 5.2 formuluje proměnné governance jako vlastnosti činností a záznamů, a nikoli systémů. Nástroj pořízený dříve, než je přidělena rozhodovací autorita, digitalizuje neurčený postup, takže záznamy, jež generuje, nenaplnují žádnou proměnnou, neboť proměnná se čte ze záznamu pojmenovávajícího roli, datum a výsledek. Pořízení instrumentace má tedy posouditelné účinky teprve po přiřazení, na němž její záznamy závisejí, což je výrok o pořadí závislosti, a nikoli o pořizování.

Dva další důsledky plynou ze zjištění kapitoly 13 a oba se týkají toho, co instituce drží poté, co bylo dosaženo určení. Mezera v incidentech z oddílu 13.4 umísťuje řešení incidentů do governance, a nikoli do samostatné funkce, neboť se týká toho, kdo rozhoduje, v jakém intervalu a s jakým záznamem, a odpovídajícím institucionálním důsledkem je, že registr incidentů náleží k eskalační struktuře, a nikoli k bezpečnostní službě, jež nad výzkumnými daty žádnou autoritu nedejří. Mlčení o nákladech v témže oddílu má důsledek ostřejší: rámec určuje, co je přiměřené riziku, a nikoli to, co je finančně únosné, takže instituce může držet správné určení, jež nedokáže provést. Která strana je příslušná přijmout neprovedené určení jako trvajícím zbytkové riziko, je otázkou governance, kterou rámec vznáší a nezodpovídá.

CHYBĚJÍCÍ DOKLAD

Srovnání institucí, jež přidělily rozhodovací autoritu dříve, než pořídily nástroje pro správu výzkumných dat, s institucemi, jež pořadí obrátily, s měřením podílu proměnných čitelných na úrovni 2 a výše v každé skupině, umožňující stanovit, zda má pořadí závislosti zde odvozený praktický účinek.

16.2 Rozhodovací práva data stewarda odvozená z rámce

Role odvozená z rámce se od role odvozené z popisu pracovního místa liší tím, co určuje její hranice. Oddíl 5.1 odděluje poradní vrstvu od vrstvy rozhodovací a zachází s rozhodovací hranicí jako s proměnnou, jež má vlastní záznam, z čehož plyne, že práva data stewarda jsou fixována prahy, na nichž se porada stává určením, a nikoli výčtem úkolů. Rozhodovací práva data stewarda pokrývají při tomto odvození čtení, jež rámec vyžaduje — stav proměnné, záznam, z něž se čte, zda vznikl konflikt a které pravidlo se uplatní —, a končí v bodě, kde určení alokuje riziko. Určení druhého druhu náleží straně odpovědné za výsledek dané činnosti, vlastníkovu registru

klasifikace nebo příjemci eskalace pojmenovanému v D1, jak je přiřazuje tabulka 12-3.

Zjištění, že poradní a rozhodovací funkci nelze sloučit bez ztráty, má důsledek, jež lze snadno číst jako omezení, ač je opakem. Oddíl 12.5 uvádí tutěž hranici pro automatizované výstupy: navržená klasifikace nebo doporučené určení přístupu jsou poradou a rozhodnutím je jejich přijetí stranou, která má autoritu. Data steward, jenž zároveň radí v otázce zpřístupnění a zpřístupnění schvaluje, zaujímá postavení, v němž žádná vnější strana nedokáže dodatečně odlišit posouzení od schválení, což odstraňuje zpochybnitelnost, kterou oddíl 15.4 identifikuje jako jedinou nápravu, již rámec vůči formalizované moci má. Oddělení tedy chrání posouzení data stewarda stejnou měrou, jakou omezuje jeho autoritu.

Co rámec na roli klade, je eskalační povinnost s určitou podmínkou splnění. Průchodnost eskalace je v tabulce 5-2 definována pojmenovaným příjemcem, očekávaným intervalem a zalogovaným výsledkem, takže povinnost je splněna směřováním a záznamem, a nikoli výsledkem, a eskalace, již se nedostane odpovědi, je zjištěním o instituci, a nikoli selháním data stewarda. Kapitola 13 přidává druhou povinnost téže podoby: tam, kde stav proměnné číst nelze, se zaznamenává důvod a odlišuje se nepřítomnost záznamu od nepřístupnosti záznamu. Obě povinnosti může naplnit role, jež nedrží vůbec žádnou autoritu, což je právě to, co je činí odvoditelnými z rámce, a nikoli z organizačního postavení.

Tabulka 16-1 uvádí hlavní implikace této kapitoly, výsledek, z něž je každá odvozena, a důkazní třídu, kterou implikace přejímá.

Tabulka 16-1. Implikace odvozené v této kapitole, s podpírajícím výsledkem a s třídou, kterou nesou podle tabulky 15-2.

Implikace	Odvozeno z	Přejatá třída
Úsilí v governance předchází úsilí v podmíněných dimenzích	Pořadí podmíněnosti, tabulka 12-2	Prokázané prozkoumáním
Pořízení nástroje má posouditelné účinky teprve po přidělení autority	Oddíly 1.2 a 5.2	Podepřené
Řešení incidentů náleží k eskalační struktuře	Mezera v incidentech, oddíl 13.4	Prokázané
Instituce může držet neproveditelné určení	Mlčení o nákladech, oddíly 13.4 a 14.6	Prokázané

Implikace	Odvozeno z	Přejatá třída
Práva data stewarda končí tam, kde určení alokuje riziko	Rozhodovací hranice, oddíl 5.1	Prokázané
Eskalační povinnost je splněna směřováním a záznamem, nikoli výsledkem	Průchodnost eskalace, tabulka 5-2	Prokázané
Institucionální posouzení vyžaduje strojově zpracovatelný stav ze strany repozitářů	Definice úrovní, oddíl 12.4; kustodie, oddíl 14.3	Podepřené
Uspořádání uznání podmiňuje, zda kritéria popisují financovanou činnost	Zpětná hrana, oddíl 12.2; oddíl 11.6	Hypotetické
Kurikula správy výzkumných dat zaměřená na rozhodovací strukturu překonávají kurikula zaměřená na nástroje	Oddíly 1.2 a 5.1	Hypotetické

16.3 Co institucionální posouzení vyžaduje, aby infrastruktury vystavovaly

Tvrzení rámce o posouzení se soustřeďuje mezi úrovněmi zralosti 2 a 3 a úroveň 3 je dosažitelná pouze prostřednictvím záznamu vzniklého jako vedlejší produkt činnosti (oddíl 12.4). Tam, kde se činnost odehrává uvnitř repozitáře, metadatového registru nebo národní služby, leží tento vedlejší záznam u služby, a nikoli u instituce, a oddíl 14.3 uvádí důsledek: proměnná je pak naplněna atestací namísto prozkoumání a její stupeň důkazu klesá. Instituce nemůže takovou proměnnou zvednout nad úroveň 1 žádným vlastním úsilím, neboť strop je vlastností toho, co služba vystavuje. Rozhraní, o něž jde, jsou tedy podmínkou institucionálního posouzení, a nikoli pohodlím.

Z již uvedených proměnných plyne pět tříd vystavení a každá z nich je odvozena, a nikoli navržena. Proměnná publikovaného objektu z tabulky 11-2 vyžaduje, aby identifikátor, verze, licence a role přispěvatelů byly získatelné pro každé uložení, a nikoli pouze zobrazené na vstupní stránce. Proměnné režimu přístupu a kontroly výstupu z prostředí z tabulky 9-3 vyžadují, aby režim nastavený u uložení byl čitelný pro ukládající instituci, neboť instituce, jež nedokáže přečíst stav přístupu k vlastním fondům, nedokáže posoudit vlastní přičtení omezení. Proměnná důkazu integrity vyžaduje, aby datované události kontroly bitové neměnnosti a migrace byly získatelné, a nikoli pouze tvrzené v dokumentu politiky, a proměnná ujednání o ukončení

vyžaduje, aby ustanovení o nástupnictví a cesta exportu metadat byly uvedeny v podobě, kterou může záznam o výběru citovat s datem ověření.

Třída rozhraní	Obsluhovaná proměnná rámce	Dosažitelná zralost	Náhrada, chybí-li
Strojově čitelný export metadat	popis; uznání	Úroveň 3 — měřeno	ruční výpis z rejstříku
Log rozhodnutí o přístupu	ochrana; governance	Úroveň 4 — ověřeno	periodické potvrzení
Zpráva o bitové neměnnosti	uchování	Úroveň 4 — ověřeno	namátková ruční kontrola
Rozklad PID a verzování	životní cyklus; publikování	Úroveň 3 — měřeno	lokální mapa identifikátorů
Vazba plán-záznam (maDMP)	plánování	Úroveň 5 — zlepšuje se	administrativní srovnání

Rozhraní, která repozitáře a rejstříky vystavují, aby institucionální posouzení bylo možné

Obr. 16-1. Vrstva rozhraní mezi institucionálním posouzením a vnějšími službami, zobrazující pro každou z pěti tříd vystavení proměnnou rámce, již slouží, úroveň zralosti, kterou činí dosažitelnou, a náhradní řešení tam, kde vystavení chybí; podklad: tabulky 9-3, 11-2 a 12-4. Vlastní zpracování autora.

Pátá třída se týká agregace a je tou, jíž současné služby vyhovují nejméně často. Posouzení na institucionální úrovni vyžaduje, aby výše uvedené stavy byly získatelné pro fondy instituce jako celek, a nikoli po jednotlivých uloženích, neboť výběrový rámec stanovený posuzovatelem předpokládá vyčíslitelnou populaci (oddíl 14.5). Tam, kde takový dotaz neexistuje, dodává výběrový rámec nutně posuzovaná instituce a první zranitelnost vůči vykazování místo konání z oddílu 14.5 zůstává strukturně otevřená. Národní metadatové adresáře zaujímají v tomto ohledu zvláštní postavení, neboť registrační povinnosti jim vyčíslitelnou populaci konstrukčně dávají (EOSC CZ 2025).

CHYBĚJÍCÍ DOKLAD

Přehled strojově zpracovatelných rozhraní vystavovaných široce užívanými generalistickými a oborovými repozitáři proti pěti zde odvozeným třídám vystavení, s uvedením maximální úrovně zralosti, jíž může vnější instituce dosáhnout, pro každou proměnnou tabulek 9-3 a 11-2.

16.4 Důsledek toho, že se s uznáním zachází jako s podmiňující proměnnou

Oddíl 12.2 připouští jeden vztah běžící proti pořadí podmíněnosti, v němž uznání podmiňuje každou dimenzi, protože úsilí neviditelné pro institucionální hodnocení se neudrží. To, že se s uznáním zachází jako s proměnnou, a nikoli jako s politickou aspirací, má důsledek pro status vlastních kritérií rámce. Tam, kde cesta uznání existuje, popisují kritéria posouzení činnost, kterou má instituce důvod udržet; tam, kde neexistuje, popisují tatáž kritéria nefinancované očekávání a nízký profil je pak zjištěním o systému hodnocení, a nikoli o způsobilosti instituce. Toto rozlišení je dostupné pouze proto, že uznání je neseno jako proměnná se záznamem, totiž jako pravidla hodnocení a záznamy o vykázání výsledku z tabulky 11-2.

Druhý důsledek se týká návrhu samotných uspořádání uznání. Oddíl 11.6 identifikuje nebezpečí, že nahrazení počtu článků počtem datových sad reprodukuje mechanismus, jež měla reforma napravit, a proměnné rámce lokalizují přesně to místo, kde k záměně dochází: počet uložení čte pouze proměnnou publikovaného objektu a nedotýká se žádné z proměnných popisu, výběru k uchování, přístupu ani dlouhodobého uchování, jež rozhodují o tom, jakou hodnotu uložení má. Uspořádání uznání, jež čte několik proměnných namísto jediné, je při tomto odvození vůči záměně méně náchylné, což je tvrzení strukturní, a nikoli změřené. České uspořádání z roku 2026 poskytuje instanci, v níž registrace v národním adresáři spojuje uznání s kvalitou popisu, a jeho účinky na chování jsou předmětem značky v kapitole 11, jež zůstává otevřená (Úřad vlády České republiky 2025).

CHYBĚJÍCÍ DOKLAD

Srovnání uspořádání uznání, jež čtou jediný počet uložení, s uspořádáními vážícími uznání na metadatové nebo kurátorské podmínky, s měřením úplnosti popisu a zajištění řízeného přístupu v obou případech, umožňující otestovat, zda víceproměnné uznání záměnnému mechanismu vzdoruje.

16.5 Co plyne pro výuku způsobilosti ve správě výzkumných dat

Rámec umisťuje selhání, jimiž se zabývá, do rozhodovací struktury, a nikoli do dovednosti, a toto umístění je odvozené, nikoli tvrzené. Kapitola 1 uvádí nasycenost nabídky nástrojů vedle přetrvávajícího selhávání, oddíl 5.1

připisuje opakující se kategoriální omyl slučování poradní a rozhodovací funkce a oddíl 13.5 ukazuje, že proměnné nesoucí posuzovací váhu rámce se čtou ze záznamů bez interpretace. Při tomto čtení adresuje kurikulum uspořádané kolem nástrojů a formátů vrstvu, v níž je selhání soustředěno nejméně. Odvození je silné ve svém prvním kroku a slabší v druhém, neboť soustředění selhání v rozhodovací struktuře je syntézou korpusu, a nikoli měřením, a implikace níže jsou z tohoto důvodu v tabulce 16-1 označeny jako hypotetické.

Platí-li toto umístění, plynou pro obsah vzdělávání ve správě výzkumných dat tři důsledky. Způsobilstv číst záznam podle toho, čeho je důkazem — odlišit plán od realizovaného opatření, tvrzení od záznamu vzniklého jako vedlejší produkt a atestaci od prozkoumání —, je operativním jádrem, neboť takovým čtením je každý stav proměnné. Způsobilstv směřovat namísto řešit plyne z eskalační povinnosti z oddílu 16.2, neboť schopnost identifikovat, které autoritě konflikt náleží, je oddělitelná od autority jej rozhodnout. A kazuistický materiál nesoucí konflikty, jež pravidla nerozhodnou, ze tří tříd tabulky 13-2 učí tomu, čemu kurikulum vypracovaných příkladů naučit nemůže: že obhajitelným výsledkem je někdy eskalace se zdokumentovaným konfliktem, a nikoli určení.

Čtvrtý důsledek se týká posuzování učení a zrcadlí vlastní rozlišení rámce. Kurikulum posuzované podle toho, co dokážou účastníci popsat, měří dokumentaci způsobilstvi stejným způsobem, jakým počty plánů měří dokumentaci správy výzkumných dat, a obdobným výsledkovým měřítkem je to, zda je určení přijaté následně ve vlastním projektu účastníka dohledatelné k proměnným rámce. Design takového měření je uveden v oddílu 4.6 a proveden nebyl. Právě jeho absencí se vysvětluje, proč jsou zde vzdělávací implikace nejslabší z celé kapitoly.

CHYBĚJÍCÍ DOKLAD

Aplikační test s 12–20 účastníky z alespoň dvou oborů, srovnávající kurikulum zaměřené na rozhodovací strukturu s kurikulem zaměřeným na nástroje prostřednictvím slepého bodového hodnocení artefaktu vytvořeného ve vlastním projektu účastníka po šesti až osmi týdnech, umožňující stanovit, zda zde odvozené umístění selhání platí.

Výše odvozené implikace sdílejí jednu vlastnost, jež rozhoduje o tom, jak se čtou. Každá plyne z uvedeného výsledku a nese důkazní třídu tohoto výsledku, takže nejsilnější z nich jsou stejně jisté jako prozkoumání

struktury rámce a nejslabší jsou předpověďmi čekajícími na měření pojmenovaná v jejich značkách. Co kniha ustavila, co ponechala otevřené a co si nyní výzkumný program žádá, je předmětem závěrečné kapitoly.

Kapitola 17. Závěr

17.1 Problém a podaná odpověď

Problémem, jímž se tato kniha zabývá, není nedostatečnost některého z nástrojů upravujících výzkumná data, nýbrž absence vrstvy, jež by je spojovala. FAIR se vypovídá o datové sadě, TRUST o repozitáři, plán správy dat o projektu, právo ochrany osobních údajů o operaci zpracování a reforma hodnocení o kariéře a skládání napříč těmito objekty nedefinuje žádný z nich (Wilkinson et al. 2016; Lin et al. 2020; European Parliament and Council 2016). Kapitola 1 ustavila, že toto částečné pokrytí je systematické, a nikoli nahodilé: nástroje jsou nejsilnější tam, kde lze objekt popsat izolovaně, a nejslabší tam, kde se rozhoduje za konkurujících si omezení, tedy právě tam, kde je praxe nejhustší. Otázka, jež z toho plyne, zní, zda lze částečné požadavky integrovat do jednoho rámce proměnných, rozhodovacích pravidel a kritérií posouzení, který přináší rozhodnutí obhajitelná na institucionální i projektové úrovni.

Odpověď, kterou kniha podává, je podmíněná a její podmínky se uvádějí, namísto aby byly pouze naznačeny. Integrace je dosažitelná na úrovni rozhodovací struktury a není dosažitelná na úrovni určení, protože proměnné nesoucí autoritu a posouditelnost drží napříč prostředími, kdežto proměnným nesoucím odpovědi fixuje stavy převážně právní režim, obor, velikost a infrastruktura. Rámec tohoto druhu je tedy exportovatelný jako struktura, a nikoli jako soubor odpovědí, což je zjištění oddílu 14.7 a důvod, proč kniha neobsahuje žádná doporučení. Tím, co integrace přináší, není při tomto čtení shoda o tom, co dělat, nýbrž postup končící určením přičitatelným pojmenované straně, se zaznamenaným zbytkovým rizikem a s uvedenou podmínkou návratu k přezkumu.

17.2 Přínos v pořadí, v němž byl budován

Argument byl budován v posloupnosti, v níž každý krok dodával to, co následující krok předpokládal. Část I ustavila problém, charakterizovala data stewardship jako institucionální způsobilost, a nikoli jako pracovní místo nebo servisní pult, a kriticky přečetla normativní vrstvu, aby přesně určila, kde FAIR, CARE a TRUST přestávají cokoli určovat (Carroll et al. 2020). Část II fixovala výzkumný design: korpus 23 domén stratifikovaný podle funkce,

kteřou každý zdroj plní, tříúrovňové odstupňování důkazní síly, kódovací postup o třech průchodech, z něž byly dimenze odvozeny, a nikoli mu vnuceny, a tři testy vnitřní konzistence požadující, aby se každá proměnná připojila k roli, pojmenovala záznam a směřovala své konflikty k autoritě. Ve stejném bodě bylo fixováno oddělení konstrukce od validace, jež řídí vše, co následuje.

Část III vyprodukovala v sedmi kapitolách samotný rámec. Šest dimenzí bylo rozvinuto jako soubory proměnných, z nichž každá je definována jako vlastnost datové činnosti a každá pojmenovává záznam, z něž se čte její stav: governance a odpovědnost za výsledek, životní cyklus a provenience, plánování jako závazek, popis a kvalita, ochrana a přístup, reprodukovatelnost a dlouhodobé uchování spolu s publikováním a uznáním. Kapitola 12 poté provedla skladbu, jež soubor proměnných převádí v rámec: doplnila úroveň indikátoru, ustavila pořadí podmíněnosti mezi sedmi dimenzemi, uvedla pět rozhodovacích pravidel pro konkurující si požadavky a předefinovala zralost jako pozorovatelný stav proměnné, a nikoli jako aspirační stupeň instituce. Záznam o zbytkovém riziku s šesti poli, mezi nimi s rolí přijímající riziko a s podmínkou, za níž se rozhodnutí vrací k přezkumu, je výstupem, jenž odlišuje určení od preference.

Část IV uvedla konstrukci do styku s materiálem, jež sama nevytvořila, a oddělila dvě operace, které vývoj rámců běžně slučuje. Čtyři prostředí byla rekonstruována tak, aby obměňovala velikost, obor, právní režim a infrastrukturu, a rekonstrukce vrátily smíšený výsledek uvedený bez zjemňování: rámec je vyjádřitelný a jeho postup se uzavírá, přičemž je neúplný ve věci incidentů, mlčí o nákladech a je neurčitý pro tři charakterizovatelné třídy konfliktu. Kapitola 14 ustavila, že citlivost na prostředí má vzorec, a nikoli difuzní podobu, a že sleduje pořadí podmíněnosti, což je asymetrie, kterou opakuje oddíl 17.1. Část V umístila výsledek vůči modelům zralosti, nástrojům posuzování FAIR, certifikaci repozitářů a reformě hodnocení, vzala alternativní vysvětlení vážně a přiřadila každé hlavní tvrzení k jedné ze tří důkazních tříd.

Část VI odvodila důsledky těchto výsledků pro instituce, pro role držící rozhodnutí, pro infrastrukturu, pro hodnocení výzkumu a pro vzdělávání, a odvodila je pod jedinou kázní: každá implikace nese důkazní třídu výsledku, z něž plyne. Tato kázeň produkuje nesourodý soubor, v němž jsou důsledky pořadí podmíněnosti stejně jisté jako prozkoumání závislostní struktury rámce, kdežto důsledky pro výuku jsou předpověďmi čekajícími na aplikační

test. Nesourodost se uvádí, namísto aby byla zarovnána, protože kapitola implikací předložených v jednotné síle by byla nerozlišitelná od souboru doporučení.

Přínos nárokový v kapitole 1 lze tedy posoudit vůči tomu, co bylo dodáno. Integrovaný návrh existuje a jeho vnitřní vlastnosti jsou založeny prozkoumáním. Operacionalizace existuje, s podmínkami, prioritním uspořádáním, pojmenovanou autoritou a zbytkovým rizikem, a její vnější vlastnosti — rozlišovací síla, spolehlivost, oddělovací síla — jsou měřeními, jež provedena nebyla. Validací je rekonstrukce doprovázená uvedenou hranicí přenositelnosti, což je méně než validace a více než ilustrace.

17.3 Meze výsledku

Meze výsledku jsou jiného druhu než meze rámce a toto rozlišení má v závěru svou váhu. Kapitola 14 ohraničuje, co rámec určuje: neurčuje, co je finančně únosné, ani zda bylo rozhodnutí správné, ani vědeckou významnost souboru dat, ani etickou přijatelnost účelu. Zde jde o něco užšího, totiž o status vlastních tvrzení knihy. Rámec odvozený z účelově sestaveného korpusu, kódovaný jediným analytikem a posouzený proti rekonstrukcím podpírá tvrzení o konstrukci, a nikoli tvrzení o institucích, a žádná věta předchozích kapitol nebyla napsána tak, aby nesla více.

Dva důsledky tohoto statusu stojí za to zopakovat vlastními slovy. První zní, že hustota důkazních značek v celém textu je vlastností oboru stejnou měrou jako této studie, neboť srovnávací měření ve správě výzkumných dat zůstává natolik vzácné, že by na totéž omezení narazil jakýkoli rámec srovnatelného rozsahu. Druhý zní, že protihypotéza z oddílu 14.2 — totiž že stabilní jádro rámce je stabilní proto, že je abstraktní, a nikoli proto, že je robustní — zůstává živá a tvrzení knihy o integraci je přesně v této míře ponecháno otevřené. Závěr, jenž by je uzavřel pouhým tvrzením, by předvedl právě to selhání, jež protihypotéza namítá.

Ještě jedna mez se týká užití, jehož je výsledek v tomto postavení schopen. Konstrukci, jejíž podmínky selhání jsou uvedeny, lze aplikovat na prostředí, zpochybňovat tam, kde jsou její určení nevítaná, a testovat tam, kde jsou její předpovědi specifické, a nemůže sloužit jako opodstatnění uspořádání kterékoli instituce, jež ji přijme. Jde o totéž rozlišení, které kapitola 3 vede mezi principiálním rámcem čteným jako omezení návrhu a tímž rámcem čteným jako nástroj souladu, a na tento rámec dopadá se stejnou silou. Jeho

přijetí ustavuje, že instituce má strukturu, a neustavuje nic o rozhodnutích uvnitř ní přijímaných.

17.4 Výzkumný program, jež kniha ponechává otevřený

Důkazní značky rozmístěné v předchozích kapitolách nejsou seznamem výhrad, nýbrž specifikací empirického programu, jež si tento rámec žádá, a spadají do čtyř skupin. První skupina je popisná a založila by stav předmětu: audit institucionálních politik pro výzkumná data napříč vymezenou populací, zaznamenávající pro každou oblast rozhodování, který nástroj je citován, zda je pojmenován vlastník rozhodnutí a zda je uvedena eskalační cesta. Druhou skupinou je měření samotného nástroje, zahrnující empirickou aplikaci na třech až pěti organizacích se sjednaným přístupem, výběrový rámec stanovený posuzovatelem, nezávislé kódování všech 55 proměnných a uvedené rozložení úrovní zralosti. Obě skupiny převádějí rekonstrukce ve zjištění o institucích, což je krok, jež část IV učinit odmítá.

Třetí skupinu zaujímá shoda hodnotitelů a její postavení je předrženo ostatním. Škála zralosti, na níž dva způsobí posuzovatelé přiřazují téže proměnné odlišné úrovně, není měřením a srovnáním mezi institucemi ani v čase pak nemají žádnou interpretaci. Oddíl 14.4 předpovídá v neshodě strukturu, a nikoli její rovnoměrnou míru — vysokou shodu na hranicích mezi úrovněmi 0, 1 a 2, nižší mezi úrovněmi 2, 3 a 4 a nejnižší u interpretativních proměnných z oddílu 13.5 —, a studie vykazující jediný koeficient napříč všemi proměnnými by právě tuto strukturu zakryla. Design je uveden: alespoň tři posuzovatelé, týž projektový materiál, shoda vykázaná pro každou proměnnou a pro každou hranici mezi úrovněmi proti prahu stanovenému předem, přičemž proměnné pod tímto prahem se uvádějí jako rozhodčí určení, a nikoli jako posouzení.

Čtvrtá skupina je longitudinální a nese otázku, na níž nakonec závisí praktická hodnota rámce. Přijetí kteréhokoli nástroje, jenž určuje zdroje, mění chování, které tento nástroj měří, a nejlevněji vyprodukovanou změnou je změna dokumentační. Zda přijetí rámce mění rozhodnutí, jež instituce přijímají, anebo pouze záznamy, jež vedou, lze zodpovědět pozorováním stavů proměnných, registrů konfliktů a realizovaných opatření v odstupu po dobu několika let, přičemž instituce přijímající rámec pro podporu rozhodování se odliší od institucí přijímajících jej pod tlakem auditu. Předpověď, kterou rámec generuje, je dostatečně specifická, aby mohla

selhat: mění-li přijetí pouze dokumentaci, rozdíl mezi profily ze sebesouzení a profily ověřenými posuzovatelem se neuzavře a pozorování na úrovni 4 zůstanou soustředěna v naplánovaných událostech.

Dvě další otázky stojí mimo tyto skupiny a rámec by spíše rozšířily, než otestovaly. Proměnná incidentu identifikovaná v oddílu 13.4 si žádá formulaci a posouzení jako součást dimenze governance a mlčení o nákladech v témže oddílu značí hranici, na níž se rámec tohoto typu setkává s otázkou, k jejímuž zodpovězení nebyl vybudován, totiž s otázkou, jak naložit s určením, jež je přiměřené riziku a přitom finančně neúnosné pro instituci, která je drží. Zda kapacita náleží do rámce jako proměnná se záznamem, anebo zůstává kontextovou podmínkou, vůči níž se určení posuzují, je otevřenou otázkou návrhu, a nikoli vadou k napravení doplněním.

CHYBĚJÍCÍ DOKLAD

Studie s opakovaným měřením po dobu alespoň pěti let na institucích přijímajících rámec a na párovaných institucích, jež jej nepřijaly, zaznamenávající stavy proměnných, registry konfliktů, realizovaná opatření a rozdíl mezi profily ze sebesouzení a profily ověřenými, umožňující stanovit, zda přijetí mění rozhodnutí, nebo dokumentaci.

17.5 Čím výsledek je

Podmínku, na níž závisí vše v této knize, lze nejúsporněji vyslovit jako vlastnost záznamů. Proměnná je použitelná v posuzování tehdy, čte-li se její stav ze záznamu, jež činnost vyprodukovala z vlastních důvodů, protože takový záznam může odporovat tomu, co o sobě instituce uvádí, a žádné tvrzení, žádná politika ani žádný plán tuto vlastnost nemají. Proto rámec nepočítá nic, co může instituce vyprodukovat tím, že se dobře popíše, proto je úroveň zralosti 3 oddělena od úrovně 2 záznamem vzniklým jako vedlejší produkt a úroveň 4 událostí, kterou instituce nenaplánovala, a proto se s počty plánů v celé knize zachází jako se statistikami pokrytí bez výsledkové interpretace. Rozlišení mezi dokumentací a praxí není tématem knihy, nýbrž jejím operativním kritériem.

Druhým, čím výsledek je, je lokalizace toho, čeho je v oboru nedostatek. Správu výzkumných dat adresuje několik set nástrojů a služeb a jejich počet dále roste, zatímco rozhodnutí určující, jakou hodnotu data mají poté, co projekt skončí, zůstávají ve velkém podílu prostředí nepřidělena. Instrumentace je hojná a vzácná je struktura, do níž lze instrumentaci umístit:

přidělená autorita, práh, na němž se porada stává určením, eskalace, jež se uzavírá, a záznam rizika, které někdo přijal. Rámec, jenž takovou strukturu dodává, rozhodnutí nezlepšuje, nýbrž je činí přičitatelnými, což je vlastnost slabší a ta, na níž závisí jakékoli následné zlepšení.

Třetí a nejméně pohodlný bod se týká toho, co rámec dělá tam, kde jsou požadavky ve skutečném konfliktu. Konflikt nerozpouští a ve třech identifikovaných třídách nevrací žádné určení vůbec, nýbrž eskaluje se zdokumentovaným konfliktem. Takové chování není nedostatkem konstrukce, nýbrž poctivým koncovým bodem každé struktury, jež respektuje skutečnost, že retenční povinnost i žádost o výmaz, stejně jako právní dovolení i komunitní protokol, jsou uplatněny správně. Rámce, jež v takových případech určení vydávají, skrývají hodnotový soud uvnitř procedury, a přínosem tohoto rámce je, že soud, role, která jej přijímá, a riziko, jež ponechává, jsou v záznamu, kde je lze zpochybnit.

References

Aksenova, Anna, Anoop Johny, Tim Adams, Phil Gribbon, Marc Jacobs, and Martin Hofmann-Apitius. 2024. "Current State of Data Stewardship Tools in Life Science." *Frontiers in Big Data* 7: 1428568. <https://doi.org/10.3389/fdata.2024.1428568>.

Anscombe, F. J. 1973. "Graphs in Statistical Analysis." *The American Statistician* 27 (1): 17–21. <https://doi.org/10.1080/00031305.1973.10478966>.

Barker, Michelle, Neil P. Chue Hong, Daniel S. Katz, Anna-Lena Lamprecht, Carlos Martinez-Ortiz, Fotis Psomopoulos, Jennifer Harrow, Leyla Jael Castro, Morane Gruenpeter, Paula Andrea Martinez, and Tom Honeyman. 2022. "Introducing the FAIR Principles for Research Software." *Scientific Data* 9: 622. <https://doi.org/10.1038/s41597-022-01710-x>.

Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, Rodrigo Sara, Jennifer D. Walker, Jane Anderson, and Maui Hudson. 2020. "The CARE Principles for Indigenous Data Governance." *Data Science Journal* 19 (1): 43. <https://doi.org/10.5334/dsj-2020-043>.

CoARA (Coalition for Advancing Research Assessment). 2022. *Agreement on Reforming Research Assessment*. Brussels: CoARA. <https://coara.eu/agreement/the-agreement-full-text/>.

Consultative Committee for Space Data Systems. 2012. *Reference Model for an Open Archival Information System (OAIS)*. Recommended Practice CCSDS 650.0-M-2 (Magenta Book). Washington, DC: CCSDS Secretariat. <https://public.ccsds.org/pubs/650x0m2.pdf>. Superseded for standardization purposes by ISO 14721:2025.

CoreTrustSeal Standards and Certification Board. 2022. *CoreTrustSeal Requirements 2023–2025* (Trustworthy Data Repositories Requirements). Version 01.00. Geneva: Zenodo. <https://doi.org/10.5281/zenodo.7051012>.

CoreTrustSeal Standards and Certification Board. 2025. *CoreTrustSeal Requirements 2026–2028*. Version 01.00. Geneva: Zenodo. <https://doi.org/10.5281/zenodo.17660463>.

Cousijn, Helena, Patricia Feeney, Daniella Lowenberg, Eleonora Presani, and Natasha Simons. 2019. "Bringing Citations and Usage Metrics Together

to Make Data Count." *Data Science Journal* 18 (1): 9. <https://doi.org/10.5334/dsj-2019-009>.

Cox, Andrew M., Mary Anne Kennan, Liz Lyon, Stephen Pinfield, and Laura Sbaffi. 2019. "Maturing Research Data Services and the Transformation of Academic Libraries." *Journal of Documentation* 75 (6): 1432–1462. <https://doi.org/10.1108/JD-12-2018-0211>.

Crowston, Kevin, and Jian Qin. 2011. "A Capability Maturity Model for Scientific Data Management: Evidence from the Literature." *Proceedings of the American Society for Information Science and Technology* 48 (1): 1–9. <https://doi.org/10.1002/meet.2011.14504801036>.

DataCite Metadata Working Group. 2024. *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs*. Version 4.6. Hannover: DataCite e.V. <https://doi.org/10.14454/mzv1-5b55>.

DataCite Metadata Working Group. 2026. *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs*. Version 4.7. Hannover: DataCite e.V. <https://doi.org/10.14454/qdd3-ps68>.

DORA. 2013. *San Francisco Declaration on Research Assessment*. <https://sfdora.org/read/>.

EOSC Association. 2024. *Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC)*. Version 1.3. Brussels: EOSC Association. <https://doi.org/10.5281/zenodo.17582648>.

EOSC CZ. 2025. "Datové sady jako nový druh výzkumného výsledku" [Datasets as a new type of research output; result type T "Digitální kolekce dat"]. EOSC CZ, December 16, 2025. <https://www.eosc.cz/novinky-akce/clanky/vysledek-t>.

EOSC CZ. 2026. "Národní metadatový adresář (NMA)" [National Metadata Directory]. EOSC CZ. Accessed August 13, 2026. <https://www.eosc.cz/sluzby/vyhledavani/nma>.

European Data Protection Board. 2026. *Guidelines 02/2026 on Anonymisation*. Version for public consultation. Brussels: EDPB. https://www.edpb.europa.eu/public-consultations/guidelines-022026-on-anonymisation_en.

European Parliament and Council. 1996. "Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal

Protection of Databases." *Official Journal of the European Communities* L 77 (27 March 1996): 20–28. <https://eur-lex.europa.eu/eli/dir/1996/9/oj>.

European Parliament and Council. 2016. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)." *Official Journal of the European Union* L 119 (4 May 2016): 1–88. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

European Parliament and Council. 2024. "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)." *Official Journal of the European Union* L 2024/1689 (12 July 2024). <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

Federer, Lisa. 2018. "Defining Data Librarianship: A Survey of Competencies, Skills, and Training." *Journal of the Medical Library Association* 106 (3): 294–303. <https://doi.org/10.5195/jmla.2018.306>.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92. <https://doi.org/10.1145/3458723>.

Gelman, Andrew, and Eric Loken. 2014. "The Statistical Crisis in Science: Data-Dependent Analysis—a 'Garden of Forking Paths'—Explains Why Many Statistically Significant Comparisons Don't Hold Up." *American Scientist* 102 (6): 460–465. <https://doi.org/10.1511/2014.111.460>.

Higgins, Sarah. 2008. "The DCC Curation Lifecycle Model." *International Journal of Digital Curation* 3 (1): 134–140. <https://doi.org/10.2218/ijdc.v3i1.48>.

International Organization for Standardization. 2025. *ISO 14721:2025. Space Data System Practices — Reference Model for an Open Archival Information System (OAIS)*. Geneva: ISO. <https://www.iso.org/standard/87471.html>.

ISO/IEC (International Organization for Standardization and International Electrotechnical Commission). 2025. *ISO/IEC 18670:2025. Information Technology — SoftWare Hash IDentifier (SWHID) Specification V1.2*. Geneva: ISO. <https://www.iso.org/standard/89985.html>.

Jetten, Mijke, Marjan Grootveld, Annemie Mordant, Mascha Jansen, Margreet Bloemers, Margriet Miedema, and Celia W. G. van Gelder. 2021. *Pro-*

Professionalising Data Stewardship in the Netherlands: Competences, Training and Education. Dutch Roadmap towards National Implementation of FAIR Data Stewardship. Version 1.1. Amsterdam: Dutch National Programme Open Science. <https://doi.org/10.5281/zenodo.4320504>.

Library of Congress. 2015. *PREMIS Data Dictionary for Preservation Metadata*. Version 3.0. Prepared by the PREMIS Editorial Committee. Washington, DC: Library of Congress. <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>.

Liggins, Libby, Māui Hudson, and Jane Anderson. 2021. "Creating Space for Indigenous Perspectives on Access and Benefit-Sharing: Encouraging Researcher Use of the Local Contexts Notices." *Molecular Ecology* 30 (11): 2477–2482. <https://doi.org/10.1111/mec.15918>.

Lin, Dawei, Jonathan Crabtree, Ingrid Dillo, Robert R. Downs, Rorie Edmunds, David Giarretta, Marisa De Giusti, Hervé L'Hours, Wim Hugo, Reyna Jenkyns, Varsha Khodiyar, Maryann E. Martone, Mustapha Mokrane, Vivek Navale, Jonathan Petters, Barbara Sierman, Dina V. Sokolova, Martina Stockhause, and John Westbrook. 2020. "The TRUST Principles for Digital Repositories." *Scientific Data* 7: 144. <https://doi.org/10.1038/s41597-020-0486-7>.

Local Contexts. n.d. "About the Labels." Accessed August 13, 2026. <https://localcontexts.org/labels/about-the-labels/>.

Miksa, Tomasz, Paul Walk, and Peter Neish. 2020. *RDA DMP Common Standard for Machine-actionable Data Management Plans*. Version 1.1. Research Data Alliance, DMP Common Standards Working Group. <https://doi.org/10.15497/RDA00039>.

Miksa, Tomasz, Stephanie Simms, Daniel Mietchen, and Sarah Jones. 2019. "Ten Principles for Machine-Actionable Data Management Plans." *PLOS Computational Biology* 15 (3): e1006750. <https://doi.org/10.1371/journal.pcbi.1006750>.

Mitchell, Margaret, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. "Model Cards for Model Reporting." In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*, 220–229. New York: Association for Computing Machinery. <https://doi.org/10.1145/3287560.3287596>.

National Academies of Sciences, Engineering, and Medicine. 2019. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25303>.

National Digital Stewardship Alliance. 2026. *Levels of Digital Preservation*. Version 2.1. NDSA. <https://www.ndsa.org/publications/levels-of-digital-preservation/> (materials at <https://osf.io/rcs38>).

National Institute of Standards and Technology. 2010. *Contingency Planning Guide for Federal Information Systems*. NIST Special Publication 800-34, Rev. 1, by Marianne Swanson, Pauline Bowen, Amy Wohl Phillips, Dean Gallup, and David Lynes. Gaithersburg, MD: NIST. <https://doi.org/10.6028/NIST.SP.800-34r1>.

National Institute of Standards and Technology. 2023. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1. Gaithersburg, MD: NIST. <https://doi.org/10.6028/NIST.AI.100-1>.

National Institute of Standards and Technology. 2024. *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. NIST AI 600-1. Gaithersburg, MD: NIST. <https://doi.org/10.6028/NIST.AI.600-1>.

Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. 2018. "The Preregistration Revolution." *Proceedings of the National Academy of Sciences* 115 (11): 2600–2606. <https://doi.org/10.1073/pnas.1708274114>.

Paulk, Mark C., Bill Curtis, Mary Beth Chrissis, and Charles V. Weber. 1993. "Capability Maturity Model, Version 1.1." *IEEE Software* 10 (4): 18–27. <https://doi.org/10.1109/52.219617>.

Rauber, Andreas, Ari Asmi, Dieter van Uytvanck, and Stefan Pröll. 2015. *Data Citation of Evolving Data: Recommendations of the Working Group on Data Citation (WGDC)*. Research Data Alliance. <https://doi.org/10.15497/RDA00016>.

RDA FAIR Data Maturity Model Working Group. 2020. *FAIR Data Maturity Model: Specification and Guidelines*. Research Data Alliance. <https://doi.org/10.15497/rda00050>.

Sansone, Susanna-Assunta, Peter McQuilton, Philippe Rocca-Serra, Alejandra Gonzalez-Beltran, Massimiliano Izzo, Allyson L. Lister, and Milo Thurston. 2019. "FAIRsharing as a Community Approach to Standards, Repositories and Policies." *Nature Biotechnology* 37 (4): 358–367. <https://doi.org/10.1038/s41587-019-0080-8>.

Science Europe. 2021. *Practical Guide to the International Alignment of Research Data Management — Extended Edition*. Brussels: Science Europe. <https://doi.org/10.5281/zenodo.4915861>.

Seawright, Jason, and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61 (2): 294–308. <https://doi.org/10.1177/1065912907313077>.

Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22 (11): 1359–1366. <https://doi.org/10.1177/0956797611417632>.

Skills4EOSC. 2025. *Data Steward: Minimum Viable Skills Profile*. Version 3, by Karsten Kryger Hansen, Angus Whyte, Laurence Horton, Bojana Koteska, Päivi Rauste, and Nida van Leersum. Skills4EOSC project. Zenodo. <https://doi.org/10.5281/zenodo.16760455>.

Smith, Arfon M., Daniel S. Katz, Kyle E. Niemeyer, and the FORCE11 Software Citation Working Group. 2016. "Software Citation Principles." *PeerJ Computer Science* 2: e86. <https://doi.org/10.7717/peerj-cs.86>.

Soiland-Reyes, Stian, Peter Sefton, Mercè Crosas, Leyla Jael Castro, Frederik Coppens, José M. Fernández, Daniel Garijo, Björn Grüning, Marco La Rosa, Simone Leo, Eoghan Ó Carragáin, Marc Portier, Ana Trisovic, RO-Crate Community, Paul Groth, and Carole Goble. 2022. "Packaging Research Artefacts with RO-Crate." *Data Science* 5 (2): 97–138. <https://doi.org/10.3233/DS-210053>.

Steegen, Sara, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel. 2016. "Increasing Transparency through a Multiverse Analysis." *Perspectives on Psychological Science* 11 (5): 702–712. <https://doi.org/10.1177/1745691616658637>.

UK Data Service. 2025. "What Is the Five Safes Framework?" UK Data Service. Accessed August 13, 2026. <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 160018. <https://doi.org/10.1038/sdata.2016.18>.

World Wide Web Consortium. 2009. *SKOS Simple Knowledge Organization System Reference*. W3C Recommendation, 18 August 2009, edited by Alistair Miles and Sean Bechhofer. <https://www.w3.org/TR/skos-reference/>.

World Wide Web Consortium. 2013. *PROV-O: The PROV Ontology*. W3C Recommendation, 30 April 2013, edited by Timothy Lebo, Satya Sahoo, and Deborah McGuinness. <https://www.w3.org/TR/prov-o/>.

World Wide Web Consortium. 2024. *Data Catalog Vocabulary (DCAT) — Version 3*. W3C Recommendation, 22 August 2024, edited by Riccardo Albertoni, David Browning, Simon Cox, Alejandra Gonzalez-Beltran, Andrea Perego, and Peter Winstanley. <https://www.w3.org/TR/vocab-dcat-3/>.

Wu, Mingfang, Stephen M. Richard, Chantelle Verhey, Leyla Jael Castro, Baptiste Cecconi, and Nick Juty. 2023. "An Analysis of Crosswalks from Research Data Schemas to Schema.org." *Data Intelligence* 5 (1): 100–121. https://doi.org/10.1162/dint_a_00186.

Úřad vlády České republiky, Rada pro výzkum, vývoj a inovace. 2025. *Metodika hodnocení výzkumných organizací (Metodika 2025+)* [Methodology for the Evaluation of Research Organisations 2025+]. Approved by Government Resolution No. 458 of 18 June 2025. Praha: Úřad vlády ČR. <https://vyzkum.gov.cz/FrontClanek.aspx?idsekce=1064052>.

Verification notes

Year corrections (manuscript key – verified):

- CoreTrustSeal (2023) → **2022**. The *CoreTrustSeal Requirements 2023–2025* were published 5 September 2022 (v01.00, DOI 10.5281/zenodo.7051012); "2023–2025" is the validity period, not the publication year.
- EOSC CZ (2026) → **2025** for the article on result type T ("Datové sady jako nový druh výzkumného výsledku," published 16 December 2025). The NMA service page itself is undated and is cited as EOSC CZ 2026 with an access date (the new NMA version launched 19 January 2026), so the key 2026 remains usable only for the NMA entry.
- European Commission (2016) → author corrected to **European Parliament and Council** (GDPR, Regulation (EU) 2016/679); year 2016 is correct.
- W3C (2023) → **2024** for DCAT 3: it became a W3C Recommendation on 22 August 2024 (in 2023 it was only a working draft/candidate).

- NIST backup/contingency (SP 800-34 Rev. 1) → **2010**, not 2023/2024. The NIST (2023) key matches AI 100-1 (AI RMF 1.0, January 2023) and NIST (2024) matches AI 600-1 (Generative AI Profile, July 2024); if the manuscript's backup/contingency citation points to SP 800-34, it must carry the year 2010 (entry included above).
- EDPB (2026): verified as *Guidelines 02/2026 on Anonymisation* — currently the public-consultation version; if a final adopted version appears before the book goes to press, update the entry. (The companion *Guidelines 01/2025 on Pseudonymisation* exist should the manuscript also need a 2025 EDPB key.)
- NDSA (2026): verified — Version 2.1 of the Levels of Digital Preservation was announced 23 March 2026.
- ISO 14721:2025 (OAIS revision): verified as published; the CCSDS (2012) Magenta Book CCSDS 650.0-M-2 remains the correct source for the 2012 key.
- EOSC Association (2024): verified as SRIA v1.3 (dated 31 October 2024, deposited on Zenodo 11 December 2024); the formal corporate author on Zenodo is the "Horizon Europe Co-programmed Partnership for the European Open Science Cloud," commonly cited under EOSC Association.
- DataCite Metadata Working Group (2024) = Schema 4.6 (released 5 December 2024) and (2026) = Schema 4.7 (released 3 March 2026) — both verified with DOIs.
- Skills4EOSC (2025): verified as *Data Steward: Minimum Viable Skills Profile*, v3, 7 August 2025 (extract from the annex to deliverable D2.6).
- UK Data Service (2025): the Five Safes page is an undated webpage; cited with the manuscript's year and an access date.
- Federer (2018): verified as the JMLA data-librarianship competencies survey (not the PLOS ONE data-reuse study, which matches the "data librarian / data services" description less well).

UNVERIFIED entries: none.

All [CITATION REQUIRED] markers in the manuscript were resolved on 13 August 2026; fourteen entries above were added in that pass and verified against Crossref, Zenodo, ISO and publisher records.

Glosář

Uvedené definice jsou pracovní a platí pro účely této knihy. Tam, kde je termín vymezen právním předpisem nebo platnou normou, je rozhodující primární zdroj. Hesla jsou řazena abecedně podle českého termínu; anglický ekvivalent se uvádí v závorce, protože odborná literatura k tématu je převážně anglická.

Anonymizace (anonymization). Zpracování, po němž osoba již není identifikovatelná prostředky, jejichž použití lze rozumně předpokládat. Anonymizovaná data stojí mimo působnost právní úpravy ochrany osobních údajů; o tom, zda je tento práh dosažen, rozhodují testy vyčlenění, propojitelnosti a odvození.

Bitová neměnnost (fixity). Vlastnost spočívající v tom, že se bitový obsah objektu nezměnil; ověřuje se kontrolními součty.

Data steward. Profesní role, která koordinuje odpovědnou správu dat, radí výzkumníkům a eskaluje riziko. Rámec vyžaduje, aby poradní funkce této role zůstala oddělena od rozhodovací autority.

Data stewardship. Institucionální způsobilost, již jsou výzkumná data řízena napříč svým životním cyklem; praxe, kterou tato kniha modeluje. Správa výzkumných dat v operativním smyslu označuje činnost probíhající uvnitř této způsobilosti.

Dimenze (dimension). Nejvyšší úroveň hierarchie rámce: jedna ze sedmi oblastí — governance, životní cyklus a provenience, plánování, popis a kvalita, ochrana, reprodukovatelnost, dlouhodobé uchování a uznání —, napříč nimiž se rozhodnutí o správě dat přijímají.

Dlouhodobé uchování (preservation). Řízená činnost, již vybrané objekty zůstávají použitelné v čase. Odlišuje se od zálohování, které slouží provoznímu obnovení, a od publikování, které slouží přístupu.

Důvěryhodné výzkumné prostředí (trusted research environment). Kontrolované technické a organizační prostředí pro analýzu citlivých dat, s posouzenými osobami, projekty a výstupy.

Eskalace (escalation). Zaznamenaný přesun rozhodnutí od role, která na ně narazila, k autoritě oprávněné je přijmout; spouští jej stanovený práh.

FAIR. Findable, Accessible, Interoperable, Reusable — cílové vlastnosti digitálních výzkumných objektů a jejich metadat. FAIR neznamená otevřený.

Identifikátor konceptu (concept identifier). Persistentní identifikátor, který se rozkládá na všechny verze objektu, na rozdíl od identifikátoru jedné konkrétní verze.

Indikátor (indicator). Nejnižší úroveň hierarchie rámce: záznam, operace na něm prováděná, výběrová jednotka a datum pozorování, jimiž se stav proměnné pozoruje.

Metadata. Strukturované výroky popisující identitu, význam, původ, technické vlastnosti, práva nebo uchování objektu.

Odpovědnost za výsledek (accountability). Vlastnost datové činnosti spočívající v tom, že za její výsledek odpovídá jediná pojmenovaná role. Odlišuje se od odpovědnosti za provedení, která určuje, kdo danou činnost vykonává.

Persistentní identifikátor (PID) (persistent identifier). Identifikátor, jehož rozkládání na příslušný objekt je udržováno nezávisle na změnách umístění nebo kustodie.

Plán správy dat (DMP) (data management plan). Verzovaný záznam rozhodnutí přijatých o vytváření, organizaci, ochraně, sdílení a dlouhodobém uchování výzkumných dat. V této knize se s plánem zachází jako se souborem ověřitelných závazků, a nikoli jako s administrativním výstupem.

Práh verzování (versioning threshold). Stanovená změna souborů, metody nebo významu, na niž se z pracovní revize stává nová verze s vlastním persistentním identifikátorem.

Proměnná (variable). Jednotka rámce mezi komponentou a indikátorem: pojmenovaná vlastnost institucionální nebo projektové praxe, jejíž stav lze pozorovat a srovnávat.

Provenience (provenance). Zaznamenaná historie agentů, činností a entit, jež vytvořily objekt v jeho současné podobě. V rámci jde o vlastnost, která činí kvalitu, reprodukovatelnost a práva posouditelnými.

Pseudonymizace (pseudonymization). Nahrazení přímých identifikátorů takové povahy, že přiřazení k osobě vyžaduje odděleně uchovávanou informaci. Pseudonymizovaná data zůstávají osobními údaji.

Reprodukovatelnost (reproducibility). Získání konzistentních výsledků z týchž dat, téhož kódu a týchž analytických podmínek. Odlišuje se od replikovatelnosti, která se týká nových dat nebo nové studie.

Řízený přístup (controlled access). Režim přístupu, v němž jsou metadata veřejná, kdežto přístup k datům vyžaduje posouzenou žádost. Režim pro data, jež nejsou ani otevřená, ani uzavřená.

Řízený slovník (controlled vocabulary). Vyjmenovaná množina přípustných hodnot pro metadatový prvek, udržovaná identifikovatelnou autoritou.

Strojově zpracovatelný (machine-actionable). Strukturovaný tak, že jiný systém dokáže obsah přečíst a jednat na jeho základě. Silnější vlastnost než strojová čitelnost, která vyžaduje pouze to, aby obsah bylo možné rozebrat.

Stupeň důkazu (evidence grade). Klasifikace výroku ze zdroje podle důkazní síly: právní nebo normativní požadavek; platná norma či certifikované schéma; případové důkazy, konsenzuální dokument nebo nastupující praxe.

Třída dat (data class). Kategorie ochrany přiřazená datovému objektu, jež určuje jeho přípustný režim uložení, přístupu a sdílení.

Úroveň zralosti (maturity level). Jeden z pěti pozorovatelných stavů proměnné nebo dimenze rámce, odlišený od úrovně nižší uvedenou pozorovatelnou skutečností, a nikoli aspirační nálepkou.

Výběr k uchování (appraisal). Posouzení, jímž se objekty vybírají k dlouhodobému uchování nebo se u nich stanoví vyřazení. Rozhodnutí o výběru k uchování má vlastníka a zaznamenaný důvod.

Záznam o zbytkovém riziku (residual risk record). Šestipolní záznam, který produkuje každé uplatnění rozhodovacích pravidel: proměnné v konfliktu, přijatá varianta, opuštěná varianta, riziko, jež zůstává ponecháno, role, která je přijímá, a podmínka, za níž se rozhodnutí vrací k přezkumu.

Příloha A. Rejstřík chybějících empirických dokladů

Tato příloha zachycuje v rejstříku každé místo, na němž rukopis uvádí empirický požadavek, jež dosavadní výzkumný program dosud nenaplnil. Je generována ze značek vložených do jednotlivých kapitol, a je proto úplná už svou konstrukcí. Každá položka pojmenovává měření nebo sběr dat, bez jehož provedení nelze související tvrzení uvést jako výsledek. Citační značky přítomné v dřívějších verzích rukopisu byly 13. srpna 2026 vypořádány proti ověřeným zdrojům a osm obrázků bylo vykresleno; obojí se zde proto již neuvádí.

Tabulka A-1. Chybějící empirické doklady podle kapitol (45 položek).

Kap.	Požadavek
1	audit institucionálních politik pro výzkumná data napříč vymezenou populací výzkumných organizací, zaznamenávající pro každou z deseti oblastí rozhodování v tabulce 1-1, který nástroj je citován, zda je pojmenován vlastníkem rozhodnutí a zda je uvedena eskalační cesta — na podporu tvrzení, že pokrytí je částečné a nepropojené, nikoli pouze nerovnoměrné
1	párové srovnání obsahu plánu správy dat při podání s realizovanou praxí při ukončení projektu na vzorku projektů, umožňující stanovit rozdíl mezi dokumentovým souladem a realizovanou praxí
2	srovnání více institucí s konstantní velikostí instituce a oborovým složením, měřící výsledkové indikátory (doba do vyřešení, úplnost metadat při uložení, podíl projektů s provedeným plánem dlouhodobého uchování) napříč centralizovaným, distribuovaným a federovaným uspořádáním, aby se otestovalo, zda organizační model predikuje výsledek po zohlednění zdrojů
2	retrospektivní sledování vzorku ukončených projektů s cílem zjistit, zda pojmenovaná role držela kustodii dat a rozhodnutí o přístupu 12 měsíců po skončení projektu, na podporu tvrzení, že vazbou omezující celý systém je kontinuita odpovědnosti za výsledek, nikoli její počáteční přidělení
3	analýza shody mezi automatizovanými skóre posouzení FAIR a nezávislým expertním auditem vhodnosti k použití na stratifikovaném vzorku datových sad, uvádějící podíl vysoce bodovaných datových sad shledaných nepoužitelnými pro jejich deklarovaný účel, na podporu tvrzení, že automatizovaná skóre neměří opětovnou použitelnost

Kap. Požadavek

- | | |
|---|--|
| 3 | zdokumentovaný soubor konfliktních případů z institucionální praxe, kódovaný podle typu konfliktu, přijatého řešení, rozhodující autority a zaznamenaného zbytkového rizika, umožňující stanovit rozložení typů konfliktu a otestovat, zda jsou řešení uvnitř instituce konzistentní |
| 4 | data na úrovni institucí o skutečném provádění deklarovaných datových politik — doby odezvy služeb, četnost eskalací, podíl projektů s implementovaným ujednáním o dlouhodobém uchování —, umožňující stanovit, zda dokumentový závazek predikuje provozní praxi. |
| 4 | zdokumentovaný vyhledávací protokol pro aplikovanou vrstvu — dotazované databáze a registry, dotazovací řetězce, datum provedení, počet záznamů prověřených a vyloučených v každém kroku — postačující k tomu, aby nezávislý badatel korpus rekonstruoval. Předložený korpus byl sestaven účelovým výběrem proti kritériím v tabulce 4-1 a jeho úplnost proto není doložena. |
| 4 | nezávislé překódování náhodného vzorku nejméně 100 vytěžených formulací požadavků druhým analytikem s uvedeným koeficientem shody, umožňující stanovit, zda jsou hranice dimenzí v tabulce 4-2 reprodukovatelné, a nikoli idiosynkratické. |
| 4 | provedení expertního přezkumu konstruktů — počet a oborové složení posuzovatelů, přezkoumávaná tvrzení a naložení s každou kritikou připomínkou — na podporu jakéhokoli tvrzení, že rámec je technicky správný. |
| 4 | provedení aplikačního testu s 12–20 účastníky nejméně ze dvou oborů, včetně výchozího posouzení, zaslepeného bodování vytvořeného artefaktu podle rubriky a následné kontroly přenosu do vlastního projektu účastníka po šesti až osmi týdnech, na podporu jakéhokoli tvrzení, že rámec mění praxi, a nikoli znalosti. |
| 5 | srovnávací výsledková data napříč institucemi provozujícími odlišná organizační uspořádání — doba od žádosti k vyřešení, podíl eskalovaných žádostí, podíl eskalací uzavřených zaznamenaným rozhodnutím —, aby bylo možné otestovat, zda typ uspořádání predikuje výkon governance po zohlednění kvality eskalace. |
| 5 | longitudinální data o datových sadách, jejichž odpovědná osoba opustila instituci — podíl sad s dokumentovaným nástupcem, podíl sad, u nichž byly následné žádosti o přístup zodpovězeny, a doba do první nezodpověditelné žádosti —, aby bylo možné stanovit praktickou rychlost rozpadu nepřidělené odpovědnosti za výsledek. |
| 5 | párování institucionální data o indikátorech založených na plánech a o výsledkových indikátorech pro tutéž množinu projektů, aby bylo možné otestovat tvrzení této studie, že se obojí rozchází a že počty plánů stav dat nadhodnocují. |
| 6 | pro vzorek publikovaných analytických datových sad pokus o rekonstrukci pouze z uloženého záznamu provenience, měřící podíl obnovitelných transformačních kroků a podíl neobnovitelných kroků, jež byly provedeny ručně — na podporu tvrzení, že ruční přechody tvoří většinu rozdílu mezi zaznamenaným a skutečným odvozením. |

Kap. Požadavek

- | | |
|----|--|
| 6 | opakované testy spuštění na definované množině uložených výpočetních datových sad v pevných intervalech po uložení, zaznamenávající podíl úspěšných regenerací a příčiny selhání — pro stanovení toho, jak rychle se rozpadá předpoklad regenerovatelnosti, o němž se výběrové uchovávaní opírá. |
| 7 | srovnání položek plánu vzniklých za použití textových šablon, strukturovaných formulářů a znalostních systémů u srovnatelných projektů, skórováných podle rubriky konkrétnosti a proveditelnosti, aby bylo možné otestovat, zda nástroj ovlivňuje konkrétnost závazků, nebo pouze úplnost polí. |
| 7 | longitudinální záznam verzí plánu napříč kohortou projektů, kódovaný podle spouštěče revize, schválení a zdůvodnění, aby bylo možné otestovat tvrzení této studie, že revize spouštěná událostmi koreluje s realizovanými opatřeními silněji než úplnost plánu při podání. |
| 8 | párování srovnání úplnosti popisu a žádostí o vyjasnění ze strany těch, kdo data opětovně používají, mezi datovými sadami popsány během sběru a datovými sadami popsány při uložení, při kontrole oboru, složitosti dat a repozitáře, aby bylo možné otestovat, zda časování vzniku popisu predikuje náklady opětovného použití. |
| 8 | měření úsilí těch, kdo data opětovně používají — doba do první analýzy, počet žádostí o vyjasnění, podíl opuštěných pokusů o opětovné použití — v závislosti na úplnosti aplikačního profilu, stanovující velikost nákladů přesunutých neúplným popisem. |
| 9 | testování reidentifikace na vzorku uložených výzkumných datových sad prohlášených za anonymní, s uvedením podílu sad neobstojících v testech vyčlenění, propojitelnosti a odvození, aby bylo možné stanovit, jak často se odstranění přímých identifikátorů pokládá za dostačující |
| 9 | audit výzkumných projektů zaznamenávající, zda jsou cíle obnovy stanoveny pro jednotlivé třídy dat a zda byl ve stanoveném období proveden a zdokumentován úplný test obnovy, aby bylo možné stanovit rozdíl mezi poskytovaným zálohováním a doloženou obnovitelností |
| 10 | srovnání předem registrovaných analytických plánů s odpovídajícími publikovanými analýzami na vzorku studií, se zaznamenáním podílu odchylek vykázaných jako odchylky, aby bylo možné stanovit, zda předem stanovený plán v praxi funguje jako oddělovač druhů tvrzení |
| 11 | institucionální audit zaznamenávající pro vymezenou populaci ukončených projektů, zda byla stanovena retenční lhůta, zda bylo zdokumentováno rozhodnutí o výběru k uchování a jaký podíl projektových dat zůstal po ukončení projektu na nespravovaném úložišti |
| 11 | srovnávací měření opětovného použití, citovanosti a dlouhodobé dostupnosti datových sad uložených v oborových oproti certifikovaným generalistickým repozitářům, při kontrole oboru, aby bylo možné stanovit, zda má kompromis mezi oborovou vhodností a certifikovanou důvěryhodností měřitelné důsledky |
-

Kap. Požadavek

- | | |
|----|---|
| 11 | měření objemu uložení, úplnosti metadat a poskytování řízeného přístupu na českých výzkumných organizacích před zavedením druhu výsledku datová kolekce a po něm, oproti srovnávací skupině bez takového uznání, aby bylo možné otestovat, zda uznání mění chování ve správě dat, a nikoli chování ve vykazování |
| 12 | průřezové měření stavů proměnných uvnitř souboru projektů, testující, zda podmíněné proměnné kovariují se svými podmiňujícími proměnnými tak, jak tvrdí tabulka 12-2 — tedy pozorování, jež by uspořádání falzifikovalo. |
| 12 | uplatnění pravidel 1–5 na pevný soubor zdokumentovaných konfliktních případů několika nezávisle pracujícími posuzovateli, s uvedením podílu případů, jimž se dostane téhož určení a téhož přiřazení zbytkového rizika. |
| 12 | posouzení 55 proměnných na institucích vybraných tak, aby se lišily v nezávislých indikátorech výkonu ve správě výzkumných dat, testující, zda prahy mezi úrovněmi 2 a 3 a mezi úrovněmi 3 a 4 oddělují instituce, jež nezávislé důkazy rozlišují. |
| 12 | kódování formulací požadavků z korpusu governance a dokumentace umělé inteligence vůči 55 proměnným, s uvedením podílu formulací přiřaditelných stávající proměnné a s charakteristikou případného zbytku — tedy test toho, zda je tabulka 12-5 vyčerpávající, a nikoli ilustrativní. |
| 13 | empirická aplikace na třech až pěti výzkumných organizacích vybraných podle kritérií tabulky 13-1, zahrnující sjednaný přístup k projektovým záznamům, výběrový rámec datových činností stanovený posuzovatelem, a nikoli institucí, nezávislé kódování všech 55 proměnných dvěma posuzovateli, uvedený koeficient shody, rozložení úrovní zralosti a registr konfliktů s uplatněným pravidlem a se zaznamenaným zbytkovým rizikem — studie, jež by níže uvedené rekonstrukce převedla ve zjištění o institucích. |
| 13 | aplikace rámce na probíhající projekt s více partnery po celou dobu jeho trvání, se stavy proměnných zaznamenanými ve třech bodech a s párovým srovnáním závazků plánu s realizovanými záznamy provedeným při ukončení projektu, umožňující stanovit, zda určení rámce zůstávají stabilní, mění-li se podmínky projektu. |
| 13 | nezávislé přiřazení stavů a úrovní zralosti u všech 55 proměnných alespoň třemi posuzovateli pracujícími z téhož projektového materiálu, se shodou vykázanou zvlášť pro obě výše identifikované skupiny, umožňující stanovit, které proměnné jsou použitelné v posuzování a které vyžadují rozhodčí určení. |
| 13 | registr autentických konfliktních případů shromážděných z výzkumných projektů, kódovaných podle třídy a podle pravidla, jež je vyřešilo, s uvedením podílu případů spadajících do tří tříd tabulky 13-2 — měření, na němž závisí jakékoli tvrzení o pokrytí pravidel. |
-

Kap. Požadavek

- | | |
|----|--|
| 14 | nezávislé zařazení všech 55 proměnných do tří výše uvedených tříd dvěma analytiky pracujícími pouze z definic, s uvedeným koeficientem shody, spolu s rozložením stavů invariantních proměnných napříč prostředními — pozorování, jež odděluje robustnost od abstraktnosti. |
| 14 | studie shody s alespoň třemi posuzovateli nezávisle přiřazujícími úrovně zralosti všem 55 proměnným u téhož projektového materiálu, s vykázanou shodou pro každou proměnnou a pro každou hranici mezi úrovněmi proti prahu stanovenému předem, přičemž proměnné pod tímto prahem se uvádějí jako rozhodčí určení, a nikoli jako posouzení. |
| 14 | párové sebesposouzení a ověření posuzovatelem u týchž proměnných v týchž organizacích, s uvedením směru a velikosti rozdílu podle tříd proměnných, umožňující stanovit, nakolik lze na vykázanou zralost spoléhat a kde je nezávislé ověření nepostradatelné. |
| 15 | opakované měření deseti oblastí rozhodování z tabulky 1-1 na pevné populaci institucí po dobu alespoň pěti let, zaznamenávající, zda podíl institucí s pojmenovaným vlastníkem rozhodnutí roste s infrastrukturní konvergencí — pozorování podpírající přechodné čtení proti čtení strukturnímu. |
| 15 | měření stavů invariantních proměnných z oddílu 14.2 na alespoň 20 institucích lišících se v nezávislých indikátorech výkonu ve správě výzkumných dat, s uvedením, zda tyto proměnné instituce od sebe oddělují, nebo zda všude nabývají téhož stavu — test odlišující robustnost od prázdnoty. |
| 15 | srovnání záznamů o rozhodnutích a objemu dokumentace na institucích přijímajících rámec pro podporu rozhodování s institucemi přijímajícími jej pro interní audit, umožňující stanovit, zda způsob přijetí rozhoduje o tom, zda se mění praxe, nebo dokumentace. |
| 16 | srovnání institucí, jež přidělily rozhodovací autoritu dříve, než pořadily nástroje pro správu výzkumných dat, s institucemi, jež pořadí obrátily, s měřením podílu proměnných čitelných na úrovni 2 a výše v každé skupině, umožňující stanovit, zda má pořadí závislosti zde odvozený praktický účinek. |
| 16 | přehled strojově zpracovatelných rozhraní vystavovaných široce užívanými generalistickými a oborovými repozitáři proti pěti zde odvozeným třídám vystavení, s uvedením maximální úrovně zralosti, již může vnější instituce dosáhnout, pro každou proměnnou tabulek 9-3 a 11-2. |
| 16 | srovnání uspořádání uznání, jež čtou jediný počet uložení, s uspořádáními vážícími uznání na metadatové nebo kurátorské podmínky, s měřením úplnosti popisu a zajištění řízeného přístupu v obou případech, umožňující otestovat, zda víceproměnné uznání záměnnému mechanismu vzdoruje. |
| 16 | aplikační test s 12–20 účastníky z alespoň dvou oborů, srovnávající kurikulum zaměřené na rozhodovací strukturu s kurikulem zaměřeným na nástroje prostřednictvím slepého bodového hodnocení artefaktu vytvořeného ve vlastním projektu účastníka po šesti až osmi týdnech, umožňující stanovit, zda zde odvozené umístění selhání platí. |

Kap.	Požadavek
------	-----------

- | | |
|----|--|
| 17 | studie s opakovaným měřením po dobu alespoň pěti let na institucích přijímajících rámec a na párovaných institucích, jež jej nepřijaly, zaznamenávající stavy proměnných, registry konfliktů, realizovaná opatření a rozdíl mezi profily ze sebeposouzení a profily ověřenými, umožňující stanovit, zda přijetí mění rozhodnutí, nebo dokumentaci. |
|----|--|
-

Rozložení položek napříč kapitolami je samo o sobě vypovídající. Koncentrace v metodologické kapitole a v části IV odráží skutečnost, že rámec je v současnosti konstrukcí podepřenou zmapovaným korpusem, a nikoli validovaným nástrojem, což je postavení, jež mu přisuzuje kapitola 15. Vypořádání položek tabulky A-1 je podmínkou, za níž by tvrzení zařazená v tabulce 15-2 jako podepřená mohla být přerazena mezi prokázaná.

Příloha B. Protokol externí recenze a validace

Druh výsledku, podle něž se tato kniha posuzuje, předpokládá externí odbornou recenzi dokončeného rukopisu. Posouzení knižního návrhu ze strany nakladatele tomuto předpokladu nevyhovuje, protože hodnotí záměr, a nikoli provedený argument. Protokol uvedený níže upravuje recenzi tohoto rukopisu a produkuje dokumentaci, kterou posouzení vyžaduje.

B.1 Protokol recenze

1. Verze rukopisu předaná k recenzi se zmrazí a zaznamená se kontrolní součet souboru, takže recenzovaný text je později identifikovatelný.
2. Recenzenti se vybírají mimo pracoviště autora a jejich odbornost i nezávislost se dokumentují. Vyžádány jsou dvě recenze: jedna z oblasti data stewardship a otevřené vědy, jedna z oblasti metodologie nebo institucionálního řízení.
3. Každý recenzent obdrží úplný rukopis, posudkový formulář a výslovnou žádost o odborný posudek vědecké monografie.
4. Každá recenze se vrací jako datovaný a podepsaný dokument obsahující doporučení publikovat, publikovat po zapracování úprav, nebo nepublikovat.
5. Vzniká tabulka připomínek zaznamenávající každou připomínku recenzenta, odpověď autora a provedenou změnu, s odkazem na místo této změny.
6. Porecenzní verze se zmrazí s vlastním kontrolním součtem a recenze, odpovědi i rozdílový protokol proti recenzované verzi se archivují.
7. Teprve po těchto krocích se finální rukopis uvolňuje do výroby.

B.2 Uchovávaná dokumentace

Tabulka B-1. Záznamy produkované recenzním procesem.

Záznam	Obsah	Kde je uchován
Zmrazený předrecenzní rukopis	soubor a kontrolní součet	projektový archiv
Dokumentace k recenzentům	důvod výběru, afiliace, prohlášení o nezávislosti	projektový archiv
Podepsané recenze	datované doporučení a úplný text	projektový archiv
Tabulka připomínek	připomínka, odpověď, změna, místo	příloha k archivu; souhrn citovatelný v knize
Zmrazený porecenzní rukopis	soubor, kontrolní součet, rozdílový protokol	projektový archiv

B.3 Vztah ke značkám chybějících dokladů

Recenze, kterou tato příloha upravuje, posuzuje argument knihy. Nevyporádává značky chybějících dokladů zachycené v příloze A, jež pojmenovávají empirickou práci, kterou žádná recenze nenahradí. Obě linie jsou záměrně oddělené: značky se vypořádávají výzkumem, rukopis recenzí a posouzení výsledku vyžaduje obojí.

Příloha C. Registr práv k obrázkům a tabulkám

Ke každému obrázku a ke každé tabulce se vede registr práv, protože citace zdroje sama o sobě nezakládá svolení k reprodukci cizího materiálu. Současný stav registru je jednoduchý: každý obrázek i každá tabulka v tomto rukopise jsou vlastním dílem autora, vytvořeným či sestaveným z jeho vlastní syntézy, a žádný cizí obrázek, snímek obrazovky, logo ani mapa se zde nereprodukuje. Registr se prochází znovu vždy, když se materiál doplňuje.

Tabulka C-1. Registr práv.

ID	Položka	Původ	Držitel práv	Režim	Doklad
FIG 1-1	Dvouúrovňová rozhodovací struktura	vlastní schéma	autor	vlastní dílo	zdrojový soubor v projektovém archivu
FIG 4-1	Analytický řetězec od zdroje k proměnné	vlastní schéma	autor	vlastní dílo	zdrojový soubor v projektovém archivu
FIG 5-1	Matice přiřazení pro datovou činnost	vlastní schéma	autor	vlastní dílo	zdrojový soubor v projektovém archivu
FIG 6-1	Revize, verze a vydání	vlastní schéma	autor	vlastní dílo	zdrojový soubor v projektovém archivu
FIG 7-1	Struktura ověřitelné položky plánu	vlastní schéma	autor	vlastní dílo	zdrojový soubor v projektovém archivu
FIG 8-1	Vrstevnatá architektura popisu	vlastní schéma	autor	vlastní dílo	zdrojový soubor v projektovém archivu
FIG 10-1	Řetězec reprodukovatelnosti a jeho body zlomu	vlastní schéma	autor	vlastní dílo	zdrojový soubor v projektovém archivu

ID	Položka	Původ	Držitel práv	Režim	Doklad
FIG 16-1	Vrstva rozhraní pro institucionální posouzení	vlastní schéma	autor	vlastní dílo	zdrojový soubor v projektovém archivu
TAB (vše)	Tabulky 0-1 až C-1	vlastní kompilace	autor	vlastní dílo	zdroje rukopisu

Schémata překreslující cizí originál, adaptovaný materiál a jakýkoli materiál dříve publikovaný v kvalifikační práci nebo v článku autora vyžadují samostatný řádek se získaným svolením a s jeho dokladem, a to před odevzdáním rukopisu.